ITC 4/53

Information Technology and Control
Vol. 53 / No. 4/ 2024
pp. 1152-1168
DOI 10.5755/j01.itc.53.4.36565

Smartphone-Based Psychological Sensing: A Large-Scale Study on the Impact of Extreme Isolation

Received 2024/03/06 | Accepted after revision 2024/06/02

HOW TO CITE: Yuan, Z., Cao, Y., Chen, Z. (2024). Smartphone-Based Psychological Sensing: A Large-Scale Study on the Impact of Extreme Isolation. *Information Technology and Control,* 53(4), 1152-1168. https://doi.org/10.5755/j01.itc.53.4.36565

Smartphone-Based Psychological Sensing: A Large-Scale Study on the Impact of Extreme Isolation

Zhaohui Yuan, Yuqing Cao
School of Information and Software Engineering, East China Jiaotong University, Nanchang, China;
e-mails: yuanzh@whu.edu.cn, CAOYuqqing@163.com

Zhong Chen
School of Software Engineering, East China Jiaotong University, Nanchang, China;
e-mail: chenzhong100@outlook.com

Corresponding authors: Zhaohui Yuan, e-mail: yuanzh@whu.edu.cn; Yuqing Cao, e-mail: CAOYuqqing@163.com

The COVID-19 pandemic and associated isolation measures have greatly impacted mental health, especially among students. Previous attempts at using mobile sensors to analyze users' emotional states faced barriers including insufficient data and limited modalities. This study aims to address these limitations and derive insights on psychological changes under extreme isolation. We collected a large-scale multivariate dataset from 725 undergraduate students during the complete COVID-19 campus lockdown period. To our knowledge, this is the largest dataset on this population during an extended isolated period. Features were engineered from mobile sensor data to capture modalities including physical activity, sleep patterns, and social interaction. Additionally, self-reported assessments related to mental health conditions were compiled. This rich dataset was leveraged to develop a machine learning model based on autoencoders to detect emotional states. Comprehensive experiments indicate the model can accurately predict mental health changes using mobile sensor data. Our work has unique contributions in collecting large-scale isolated data, engineering informative modalities for modeling mental health, and providing a validated detection method. This can support rapid screening and intervention for mental health crises, especially those emerging from extreme events. The dataset and models open promising directions for big data analytics in mobile health and psychological research.

KEYWORDS: Big Data, Mobile Sensing, Mental Health, Machine Learning, Auto Encoder.

1. Introduction

According to a report from the U.S. Department of Health & Human Services, the global population is currently facing a severe mental health crisis. In 2021, statistics showed that the proportion of young people aged 18 to 25 with severe schizophrenia in the United States reached 11.4%, and over 4.8% of adults had serious suicidal thoughts, with 1.4% having attempted or planned suicide [1]. Mental health issues such as depression and anxiety have become major factors affecting academic performance, and daily life, especially among young students. The COVID-19 pandemic has exacerbated these issues, with students experiencing increased academic burdens, time constraints, and learning difficulties, as well as social isolation due to campus lockdowns, leading to social anxiety, loneliness, helplessness, and depression [29]. During the COVID-19 pandemic, the university where the authors are located experienced two campus isolations lasting more than six weeks each. From the period of campus isolations until now, four heart-breaking student suicides occurred, and many more students may still be struggling with extreme anxiety and helplessness.

It is noteworthy that early detection of low mood is crucial in preventing more serious mental health consequences [9, 20]. However, traditional screening methods require professional psychologists or medical personnel to conduct scientific analysis through interviews or questionnaires. These methods are not only time-consuming and require a large number of professionals but are also difficult to promote on a large scale due to privacy concerns and other factors. Moreover, users may not be able to fully reflect their emotional state through one or two interviews, especially in extreme environments such as confinement or isolation where ordinary users with severe psychological problems may find it difficult to receive timely and effective intervention and help.

To address this problem, researchers have attempted to use sensors on mobile phones or wearable devices that people carry with them every day to collect activity information and analyze users' emotional states [51, 16, 33, 18, 35, 48, 19, 28]. This method has achieved a series of meaningful research results as a new supplement to traditional periodic depression screening methods and psychological health expert assessment systems. The results show that real-time continuous psychological health assessment using mobile devices can monitor depression and anxiety states and provide suggestions, thereby reducing or preventing the likelihood of users developing severe depression. This method has the advantages of real-time and continuous assessment, non-customized devices, easy operation, and complete automation, which can partially compensate for the shortcomings of insufficient interview experience of interviewers and the influence of unstable emotional factors of the evaluated persons, and has great application prospects.

However, using sensors on mobile devices for psychological sampling faces a series of barriers. Firstly, accurate psychological monitoring models require a large amount of sampling data support to establish correlations between various symptoms and negative emotions. However, the number of volunteers participating in current research is relatively small, generally only tens of people, which makes it difficult to accurately reflect users' psychological states and changes and train universal accurate monitoring models. Secondly, human emotions such as depression and anxiety are complex emotional experiences that can affect all aspects of human activity and physiology [27]. However, the modalities of sampling data involved in ordinary wearable device sensors are relatively simple or even unitary, such as only collecting data on activity range and sleep quality. The sampling data is not comprehensive enough to accurately reflect various indicators of users' emotional states. Thirdly, the utilization of wearable smart devices for psychological monitoring often necessitates the extensive collection of personal information from users. Our research critically depends on the analysis of user behavioral data to enhance predictive accuracy, potentially implicating privacy concerns and augmenting the uncertainty and risk associated with psychological detection. Fourthly, there are few studies that can collect data on the psychological changes of people in extreme situations such as being isolated in lockdown campus during the pandemic period. It is also difficult to discover the subsequent psychological effects brought about by the extreme isolations. Furthermore, during the period of pandemic isolation, it has been challenging for us to obtain timely feedback from participants, which has posed a significant challenge to further refining our data collection. Finally, most previous similar studies used a single psychological detection questionnaire as a judgment basis. However, there are many categories of mental health problems with different symptoms and manifestations, and a single psychological detection questionnaire cannot fully reflect users' psychological states [26].

To address these problems and make full use of the opportunity window when many students were locked down, we collected large-scale data samples from 725 undergraduate students at our university for four weeks during the complete campus lockdown period of COVID-19 pandemic. We urgently developed a mobile application program to collect sensor data from their wearable smart devices for psychological monitoring. To our knowledge, this is the largest-scale intelligent device psychological data sampling experiment so far with more than ten times the number of participants than previous research works. At the same time, our sampling was conducted in an extreme situation where a large number of students were confined to campus and could not see the outside world, which is expected to become an important reference for predicting changes in people's psychological states in other extreme situations in the future.

To comprehensively reflect users' emotional states, we redesigned the modalities and correlations of sampling data. The collected data includes location, environmental noise level, steps taken, battery usage status, screen lock time, phone usage time, network usage traffic, sleep time, application usage status, etc., covering almost all data related to credit psychological conditions that can be sampled by smartphones.

We analyzed the relationship between symptoms and related psychological conditions in PHQ-9 [21] and SAS [54] psychological assessment questionnaires and studied the trend of psychological changes in extreme situations. We designed a machine learning-based detection network using an unsupervised autoencoder as input to detect students' emotional states and evaluated its performance on our sampling data. The sampling process followed ethical and voluntary rules in scientific research, and we plan to share all sampled data with other researchers in the next step.

Our contributions include:

- During the entire campus lockdown period of the 2019 corona virus disease (COVID-19) pandemic, we conducted psychological monitoring and collected the largest-scale data sample of college students. These data can facilitate in-depth academic research on psychological health and behavioral issues during the epidemic, thereby providing a more scientific basis for relevant policies and intervention measures.
- We redesigned the sampling data pattern and correlations to comprehensively reflect users' emotional states, thereby enhancing the accuracy and reliability of the research results.
- We analyzed the correlation between the sampled data and the scores of psychological state assessment questionnaires, providing empirical data support for our experiment.

Based on our analysis of the psychological changes of students in closed campus isolation during the epidemic, we predicted changes in psychological states in other extreme situations, and people designed a location detection network based on machine learning, using unsupervised autoencoders as input to detect students' Emotional state.

2. Related Work

Mobile sensing data is particularly suitable for assessing human behavior and mental healthy, allowing us to track humans in studies and gain insights in a non-intrusive manner [3].

2.1. Mental Healty and Mobile Sensing

In terms of mobile sensing and mental health, there have been numerous studies on depression, anxiety, stress, and mood. The StudentLife study [46] investigated the relationship between passive mobile sensing behavior and mental health outcomes in 48 college students over a 10-week period. The authors found a significant correlation between passive mobile sensing data and the PHQ-9 depression scale for the first time.

Canzan et al. designed a mobile phone application to monitor individuals affected by depressive mood disorders by analyzing only their mobility patterns from GPS traces, and they found that there exists a significant correlation between mobility trace characteristics and the depressive moods [10]. Northwestern University's study [43] showed that mobility and phone usage features extracted from mobile data were significantly correlated with the severity of depression symptoms measured by the PHQ-9. They recruited 40 participants from a general community for a two-week period. The results showed that participants' phone usage time and frequency were significantly correlated with depression symptom severity. They also reproduced their initial study using the StudentLife dataset [46] and replicated their results using different datasets [47], indicating that mobility features can be widely applied to depression sensing in different communities. These findings provide us with more information about depression symptoms and phone usage and offer assistance.

Prof. Andrew T. Campbell at Dartmouth College is one of the leaders in mobile sensing mental health. His group has been conducting research on smartphone-Based psychological sensing for years. His work involves developing methods to use smartphones to detect and monitor changes in a person's mental and emotional state [48]. The goal is to create tools that can help individuals manage their mental health and well-being. Their research involves using sensors on the smartphone to collect data on a person's behavior, such as their activity level, social interactions, and sleep patterns [5]. The work has the potential to revolutionize the field of mental health by providing individuals with real-time feedback and support to manage their mental health [49, 52].

However, these research findings require further investigation, including sampling larger scale people, exploring the relationship between different phone usage features and further exploring the impact of phone features on mental health.

2.2. Mobile Mental Health Sensing and Covid-19

Related work in the field of Smartphone-Based Psychological Sensing and monitoring human mental issues during the COVID-19 pandemic has been gaining attention in recent years.

Campbell's group collected data and then analyzed to identify patterns and changes that may indicate changes in mental or emotional state [28]. One study conducted by Wang et al. [44] explored the feasibility of using smartphones to monitor mental health during the COVID-19 pandemic. The study found that smartphone-based psychological sensing can effectively monitor mental health and well-being, and can serve as an early warning system for mental health issues.

Huckins et al. conducted a study using mobile sensing technology on undergraduate students, comparing their behavior from the previous semester to their behavior in the first semester affected by COVID-19 [17]. The authors found that students faced various psychological stressors and challenges during the early stages of the pandemic, resulting in reduced physical activity and decreased mobility. Another article reported in a subsequent study that an increase in anxiety and depression was significantly correlated with an increase in Google searches for coronavirus and COVID fatigue-related terms among a group of undergraduate students [25].

Sañudo et al. conducted a 7-day study on N=20 young people during the COVID lockdown period in Spain, finding that their average daily step count decreased by 68% and sleep time increased by 7% during the lockdown [32]. Other studies also support the conclusion that physical activity decreased during the COVID-19 pandemic [33, 7]. A study from the UK reported that participants' physical activity gradually decreased during the early stages of the pandemic, decreasing by 47% in the first week [13].

2.3. Psychological Sensing and Prediction

The smartphone sampling, learning, and predicting mental health states has been explored the potential of utilizing these devices to collect data and predict mental health outcomes.

One study conducted by Saeb et al. [31] explored the use of smartphone sensors to monitor mood in real-time. The study found that using smartphone sensors can provide accurate and reliable data on mood, which can be used to develop personalized interventions for mental health issues. Another study by Saeb et al. [30] used machine learning algorithms to predict depression symptoms based on smartphone usage patterns. The study found that features such as the number of phone calls made and received, the duration of phone calls, and the number of text messages sent and received were significant predictors of depression symptoms.

Researchers in [45] trained a random forest machine learning model on daily weather, personality, and smartphone sensing data to classify stress levels among N=117 students living in university dormitories, achieving an accuracy rate of 72%. They also found increasing evidence linking activity ability to depression [37, 6]. Finally, Xu et al. [52] conducted a study on N=188 undergraduate students, capturing daily behavior and behavioral pattern differences between the depression group and the non-depression group.

Another study conducted by Li et al. [23] developed a smartphone-based system that uses machine learning algorithms to detect depression and anxiety symptoms in real-time. The system was found to be effective in detecting symptoms of depression and anxiety, and could potentially be used to provide timely interventions for those experiencing mental health issues.

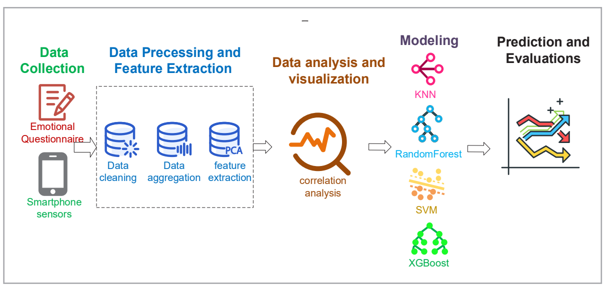
Overall, the use of smartphones for sampling, learning, and predicting mental health states shows great promise for improving mental health outcomes. However, more research is needed to fully understand the potential of these technologies and to develop effective interventions for a range of mental health conditions.

3. Data Collection

During the COVID-19 pandemic, our school was locked down twice for approximately six weeks each time, during which all students were restricted to the campus and unable to leave. As this extreme situation is rare, there is little data on the psychological changes experienced by individuals in such circumstances. To address this, we quickly developed the MentalTest smartphone application and recruited 725 undergraduate student volunteers to collect sensor data and complete daily depression and mood questionnaires for four weeks (with some missing data in the final two weeks). The gender ratio of volunteers is approximately 7:3, with the majority of individuals aged between 18 and 22 years old, ensuring the representativeness and diversity of the research findings. This

Figure 1

Overview of the research process



data collection process provided more comprehensive and detailed information than previous studies, resolving the issue of insufficient data. Figure 1 shows the flow of our research process, we present detailed methodology for data collection and analysis as well as prediction and evaluation later.

3.1. The Mental Collection Apps

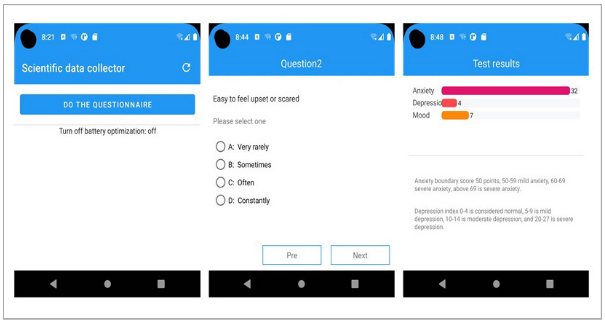
The MentalTest app was designed to access various smartphone sensors, including accelerometer readings, microphone-collected ambient sound, battery charging time, Bluetooth connection frequency and duration, gyroscope readings, ambient light levels, GPS location information, magnetometer readings, network connection time and frequency, normal physical activity levels, application pings (data points sampled per minute when an application is running), screen lock status and duration, step count, tone volume, and Wi-Fi usage.

This application accesses mobile sensors used for data collection through the universal APIs. MentalTest continuously records mobile sensor data in the background without affecting the user's normal use of the phone. We used a flat design with clear and recognizable elements to improve user experience. When users launched the app, they were prompted to complete the PHQ-9 [21] and SAS [54] psychological questionnaires to self-assess their emotional states. The questionnaire questions were straightforward and did not include excessive professional terminology or difficult-to-understand abbreviations. Users could also view their results and gain further insight into their emotional states. The smartphones used in the study were primarily based on the Android operating system, with suppliers including Samsung, Xiaomi, and others. Figure 2 shows the interface of the MentalTest application.

In the traditional sense, users tend to be sensitive about disclosing personal attributes, behaviors, habits, and other data to digital health tools such as smartphone applications. Most users are unwilling to disclose their personal behavioral data to such applications. However, our research heavily relies on the analysis of user behavioral data to enhance predictive accuracy. To address this, we have increased transparency (i.e., ensuring that users have a full understanding of the application's functionality and data usage). As a result, users are willing to share their personal behavioral data in pursuit of better, more refined services.

Figure 2

Our data collection application



During the data collection period, we strictly followed relevant laws and regulations. Users had to voluntarily grant access permissions for their data to be collected, and all information was uploaded anonymously with a timestamp of the user's first login. No user identity information was recorded, ensuring privacy.

3.2. Psychological Assessment Scales and the Groundtruth

In our study, we used the PHQ-9 and SAS questionnaires as ground truth indicators of psychological states. The PHQ-9 is a self-evaluation tool developed by the American Psychiatric Association that combines multiple factors to assess depression levels [21]. The SAS anxiety self-test measures anxiety levels and is widely used as an effective measurement method [54]. They are a simple and user-friendly self-assessment scale, suitable for respondents to complete on their own. Clinically validated through experimental trials, they demonstrate high reliability and validity in assessing depressive symptoms, providing objective assessment outcomes. After collecting all participants' questionnaire scores, we analyzed the score distributions. As shown in Table 1, the PHQ-9 scores ranged from 0 to 27 and were divided into five depression levels: no depression to mild depression (0-4), mild depression (5-9), moderate depression (10-14), and severe depression (15-27). The SAS scores ranged from 20 to 80 and were divided into four anxiety levels: normal (0-50), mild (51-59), moderate (60-69), and severe anxiety (69 or higher).

Table 1

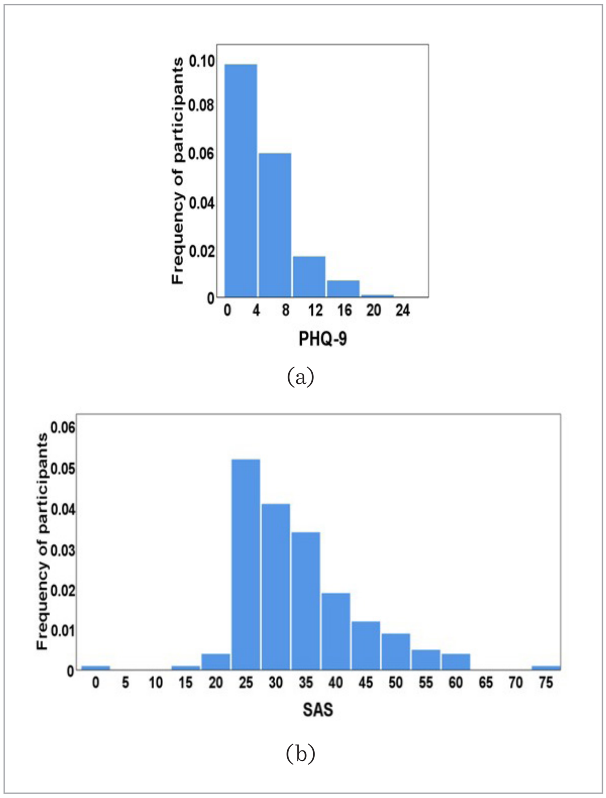
PHQ-9 and SAS Assessment Standard

LEVEL	normal	mild	Near moderate	moderate	severe
PHQ-9	0-4	5-9	10-14	15-19	20-27
SAS	0-50	51-59	Not applicable	60-69	Over 69

Figure 3 shows the distribution of scale scores for PHQ-9 and SRS, respectively. Our analysis showed that although most participants scored within the normal range on both questionnaires, many students had already experienced mental health problems. Therefore, our study provided valuable insights into the psychological effects of extreme situations such as pandemic-related lockdowns.

Figure 3

PHQ-9 score distribution. (b) SAS score distribution



4. Data Processing Methods

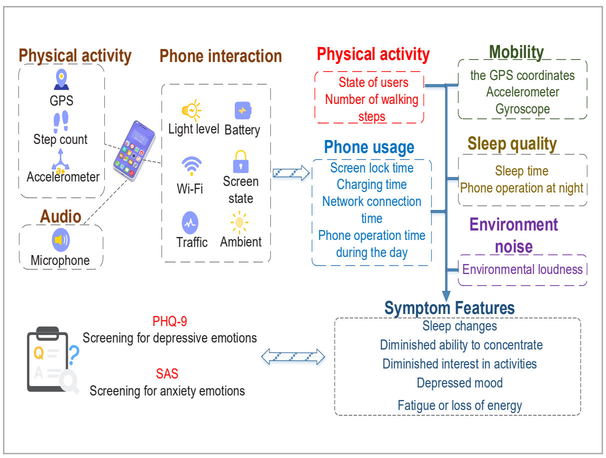
Traditional data collection modes based on smartphones are relatively limited, such as sleep or activity trajectory, which cannot fully utilize the sampling capabilities of smartphones and cannot accurately monitor changes in users' psychological states. To address this issue, we have increased the types of sampled data as much as possible and attempted to identify the relationship between different sampling data and psychological states. We perform data cleaning on the collected emotional self-evaluation results and smartphone sensor data, removing invalid or incomplete data, filling in missing data, and standardizing the data format. After data fusion and feature extraction, we conduct Pearson correlation analysis and covariance analysis, and use visualization tools such as heat maps and box plots to intuitively display the correlation and distribution of the data.

4.1. Symptom Features and Data Collection

Figure 4 gives the association between psychological symptoms and our collected data which we amassed as multi-modality as possible. Research has shown that depression and anxiety may lead to symptoms such as sleep changes, lack of concentration, loss of interest and pleasure in activities, low mood, fatigue, or loss of energy [27]. Based on these negative emotional symptoms, we have selected corresponding sensor data as possible manifestations of symptoms and explored the relationship between user mobile sensor data and their depression and anxiety questionnaire scores.

Figure 4

We continuously collect mobile sensing data from Android smartphones to seek symptom features from mobile sensing data. The symptom is characterized by mapping smartphones to five depressive symptoms: altered sleep, decreased concentration, decreased interest in activities, low mood, and fatigue or low energy



Lack of sleep can have negative effects on both physical and mental health [39] Sleep deprivation can lead to physical exhaustion, weakened immunity, increased risk of disease, and can also affect mood and cognitive function. Moreover, lack of sleep may exacerbate depressive symptoms such as anxiety, sadness, and low mood. We use four smartphone sensors - ambient light, microphone, activity, and screen switch - to infer users' sleep status.

A decrease in interest or pleasure in activities may be related to depression and anxiety [41]. During the pandemic, students may experience restrictions on work, study, entertainment in school, increased social distance, etc., which may lead to emotional instability and increased stress. These factors may cause a decrease in interest in entertainment activities. Although playing with phones is a common form of entertainment, people with depression may lose interest in leisure activities. We calculate the user's phone lock duration, charging duration, network connection duration, and daytime phone usage time to explore their interest in using phones.

Depression and anxiety are common mental health problems that can have negative effects on a person's physical and mental health, including causing fatigue. Physiologically, depression and anxiety may cause physical fatigue and lack of motivation [42]. Studies have shown that depressive symptoms can cause the body to produce more cortisol, a hormone that suppresses the sleep-wake cycle of the body, making people feel tired and unable to concentrate. Depressive symptoms may also cause the body to lack energy and endurance, making people feel tired. We use an accelerometer to detect users' states, including stationary, walking or running. We also access the phone's built-in pedometer through an application and upload step count data. These sensors can be integrated into the phone's internal chip and send sensor data to the backend application for processing and identification, reflecting the user's current state.

Geolocation and mobility are also related to depression symptoms according to evidence [15]. We extract the user's current location, which is represented by time reference, longitude value, and latitude value. We use a series of defined mobility indicators to aggregate and process user mobility information and extract useful features from their mobility trajectory using statistical methods.

Noise can have negative effects on people's mental health, leading to problems such as depression and anxiety [38]. Noise may exacerbate some depressive symptoms such as loneliness, loss of interest, and decreased self-worth. Noise can also cause alertness and anxiety. We collect phone sound information through a microphone and use an algorithm to calculate environmental loudness.

4.2. Feature Set Construction

Accurate and diverse data can help us extract more effective features to improve the predictive performance of the final model. Outdated data sources, single sources, insufficient sample sizes or improper sampling methods can all affect the reliability and accuracy of analysis results. We exclude missing, duplicate, or erroneous data from data analysis and finally decide to aggregate complete and continuous data within the main time periods each day. The aggregated amount of data must be equal to the amount of data generated within 240 minutes.

Over four weeks, we extracted several passive and continuous sensor data streams from undergraduate students' phones. According to the above rules, we first clean up the raw data by removing useless information such as missing values, duplicate values, and outliers from the data and converting it into a standardized format to ensure data quality and completeness. Then we process and extract features from the filtered data. We extract useful features from the original data that are useful for model prediction to make them easier for models to learn and understand.

Finally, according to the requirements of data aggregation, we segment these time series. We label these features as basic features, including recording duration, mean value, median value, and variance. We can divide these basic features into five categories: audio features, physical activity features, mobile interaction features, mobility features, and sleep quality features as shown in Figure 4.

5. Questionnaire Correlation and Data Correlation Analysis

In this section, we first used Pearson correlation analysis [12] to evaluate the relationship between user mobile sensing characteristics and PHQ-9 and SAS questionnaire scores, in order to validate our hypothesis that emotion characteristics derived from mobile sensing can represent the degree of depression and anxiety in college students.

PHQ-9 score is an effective measure of depression severity, while SAS score is an effective measure of anxiety severity. Correlation analysis showed the relationship between mobile sensing and the severity of depression and anxiety in users. We report the correlation coefficient and p-value [40]. In addition to correlation analysis, we used Analysis of variance (ANOVA) [36] to test whether there were significant differences in the mean values of mobile sensing characteristics between non-depressed and depressed groups, and non-anxious and anxious groups. ANOVA is a statistical model widely used to analyze differences in group means. We report the F statistics [50] and p-value from ANOVA. The F statistic represents the ratio of between-group variability to within-group variability, while the p-value indicates whether there is a significant difference in group means.

5.1. Correlation Between Symptom Features and Psychological State Scores

During isolation, people face various uncertainties and pressures, which may lead to exacerbation of negative emotions. Symptoms such as depression and anxiety are often accompanied by changes in sleep, negative thinking, inability to concentrate, fatigue, and lack of energy, which can affect people's daily activities and interests. In this mood, even things that people usually enjoy, such as playing mobile games or using social media, may become boring, uninteresting or even unappealing. Table 2 shows the results of Pearson correlation analysis.

Charging time (the time it takes to power the mobile phone through the power supply), lock time (the time the phone screen is in an inactive state), nighttime usage time (time spent using mobile phones from 7 pm to 7 am), daytime usage time (time spent using mobile phone from 7 a.m. to 7 p.m.), and online time (time spent using web browsers, apps or other online services) are proxy features for measuring mobile phone usage. Students with higher PHQ-9 scores may spend less time charging their phones ($r=-0.116$, $p=0.0.024$).

Students with more abnormal emotional states tend to have longer phone lock times (PHQ-9: $r=0.129$, $p=0.012$; SAS: $r=1.152$, $p=0.003$) and shorter phone

Table 2
Pearson correlation results

	PHQ-9		SAS	
	correlation	p-value	correlation	p-value
Loudness	0.131	0.010*	0.041	0.424
State	-0.347	<0.001**	-0.367	<0.001**
StepTo	-0.100	0.052	-0.259	<0.001**
Mobility	0.074	0.152	0.190	<0.001**
Charing Time	-0.116	0.024*	-0.078	0.128
LockTime	0.129	0.012*	0.152	0.003*
Nigth OpenTime	-0.103	0.044*	-0.105	0.040*
OpenTime	-0.132	0.010*	-0.148	0.004*
SleepTime	-0.082	0.110	-0.029	0.577
Traffic	-0.051	0.317	-0.050	0.332
WifiTime	-0.126	0.014*	-0.201	<0.001**

*p<0.01,**p<0.05

usage times (PHQ-9: r=-0.132, p=0.010; SAS: r=-0.148, p=0.004). Additionally, due to the inconvenience of accessing indoor time, we utilized Wi-Fi usage as a proxy. Subsequent experimental results confirmed that students with higher questionnaire scores are less likely to use Wi-Fi networks (PHQ-9: r=-0.126, p=0.014; SAS: r=-0.201, p=0.001), which may be related to their discomfort with online methods and reluctance to stay in relatively enclosed spaces. In addition, during the epidemic, people may feel lonely and anxious due to social isolation and restrictions. Playing with mobile phones can help people stay in touch with others, but virtual socializing cannot fully meet people's social needs, which may further exacerbate depression or anxiety.

Sleeping time is an indicator of changes in sleep. Sleep time is defined using environmental parameters such as lighting and current mobile phone usage, and we did not find a correlation between sleeping time and PHQ-9 or SAS scores. However, we found that students with more changes in sleeping time have higher scores on PHQ-9 (r=-0.082, p=0.110) and SAS (r=-0.029, p=0.577). The results indicate that students with more irregular sleep patterns are more likely to experience depressive and anxious emotions. However, another set of data shows that students who use their phones for shorter periods at night have higher questionnaire scores (PHQ-9: r=-0.103, p=0.044; SAS: r=-0.105, p=0.040), further reflecting students' aversion to mobile phones.

We found a strong correlation between questionnaire scores and user's running, walking, and stationary state (PHQ-9: r=-0.347, p<=0.001; SAS: r=-0.367, p<=0.001). People with higher PHQ-9 and SAS scores tend to remain stationary and take fewer steps (PHQ-9: r=-0.100, p=0.052; SAS: r=-0.259, p<=0.001), indicating that they are more likely to feel tired or lack energy. People in a state of depression or anxiety may prefer to remain stationary because this state can help them alleviate emotional pain and fatigue, and they may avoid any activity that makes them feel uneasy or anxious. In a state of depression, patients may experience symptoms such as low mood, lack of interest, and lack of motivation, which make them feel fatigued and lack energy.

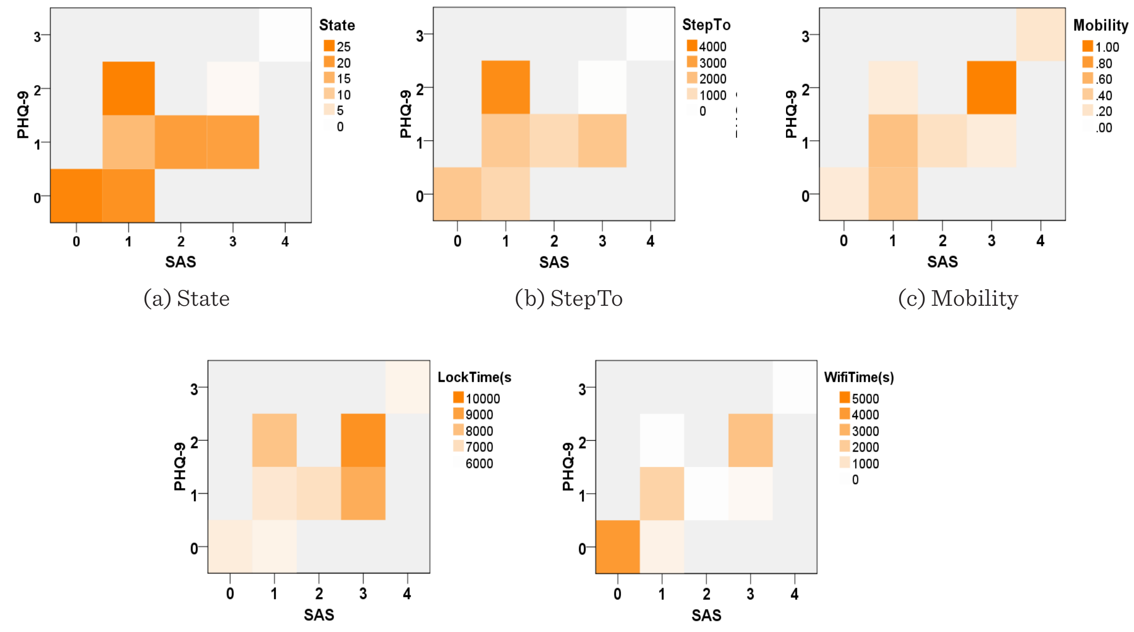
Environmental noise levels (measured using built-in microphone and associated app) showed no correlation with anxiety levels but were positively correlated with depression levels (r=0.131, p=0.010). Noise may have a negative impact on people's mental health, leading to depression and other emotional problems. Long-term noise pollution can have adverse effects on attention, mood, and sleep, deepening depressive emotions. In addition, noise may also increase stress levels, thereby affecting psychological state.

For latitude and longitude coordinates collected from GPS sensors, effective mobile information cannot be extracted through simple aggregation. We used knowledge from related fields and defined mobile indicators to further process users' GPS information. The data showed a strong correlation between anxiety levels and user mobility (r=0.190, p<=0.001). Studies have shown that anxiety can lead people to move more to seek distraction and alleviate anxiety. Anxiety can cause physiological and psychological changes such as increased heart rate, rapid breathing, muscle tension, discomfort, and stress reactions that affect a person's mobility.

According to the users' self-reported scores and questionnaire score standards, we divided PHQ-9 scores into five categories labeled 0-4 corresponding to dif-

Figure 5

Correlation between different sensing data and PHQ-9, SAS scoring levels



(a) State (b) StepTo (c) Mobility

(d) LockTime (e) WifiTime

ferent levels of depression severity; higher category numbers indicate more severe psychological problems. SAS questionnaire scores were divided into four categories labeled 0-3; higher category numbers indicate greater anxiety while lower numbers indicate less anxiety. We studied the relationship between PHQ-9 and SAS scores and proposed sensor features, presenting the results in the form of a heatmap which showed a clear correlation between the proposed sensing data and questionnaire scores in Figure 5.

5.2. Comparison Between the Depressed and Non-depressed Groups, and the Anxious and Non-anxious Groups

We divided the students into depressed (>4) and non-depressed groups (≤4), and anxious (>50) and non-anxious groups (≤50), and performed inter-group variance analysis on the data. The Analysis of variance (ANOVA) group comparison results are shown in the Table 3. We will discuss the differences between the groups in terms of environmental loudness, activity status, screen lock time, and daytime phone usage time.

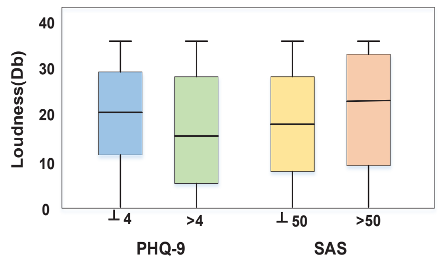
Table 3

Analysis of variance (ANOVA) results

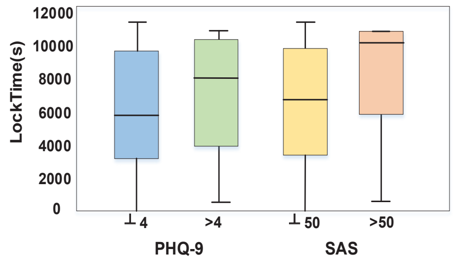
	PHQ-9		SAS	
	F	p-value	F	p-value
Loudness	17.805	<0.001	8.623	0.004
State	13.099	<0.001	32.970	<0.001
StepTo	5.026	0.026	0.040	0.842
Mobility	8.014	0.005	4.943	0.027
CharingTime	0.002	0.965	1.054	0.305
LockTime	0.056	0.813	8.720	0.003
Nigth OpenTime	0.009	0.925	6.301	0.012
OpenTime	0.039	0.844	10.986	0.001
SleepTime	1.999	0.158	0.248	0.618
Traffic	0.356	0.551	2.523	0.113
WifiTime	5.489	0.020	5.914	0.015

Figure 6

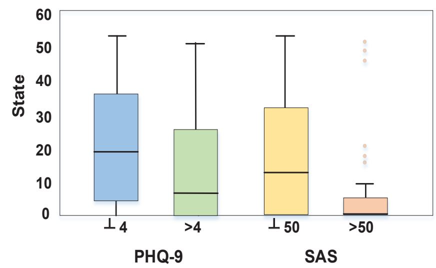
Distribution of PHQ-9 non depression group and depression group, SAS non anxiety group and anxiety group



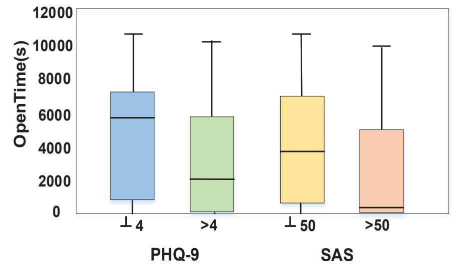
(a) Loudness distribution



(b) Locktime distribution



(c) State distribution



(d) Opentime distribution

Figure 6(a) shows the environmental loudness data of students in the depressed and anxious groups, which was subjected to variance analysis with PHQ-9 (F=17.805, p<=0.001) and SAS (F=8.623, p=0.004), with statistical significance. As shown in the figure, students in the depressed group had lower environmental loudness than those in the non-depressed group. We analyzed that this might be because students with depressive tendencies prefer to choose quiet places to be alone. On the other hand, students in the anxious group had slightly higher environmental loudness, and noise may be one of the reasons for their anxiety.

Figure 6(b) shows that students in the depressed and anxious groups had longer screen lock times, with significant intra-group differences between the anxious and non-anxious groups (F=8.720, p=0.003). This indicates that students in the anxious group were less willing to use their phones, while there was no significant difference in screen lock time between the non-depressed and non-anxious groups.

In terms of user status, there were significant intra-group differences (PHQ-9: F=13.099, p<=0.001; SAS: F=32.970, p<=0.001), with statistical significance. Figure 6(c) shows that students in the depressed and anxious groups were more likely to maintain a relatively static state, which is a common phenomenon. The reasons for this may be that depression and anxiety have negative effects on a person's emotions, making them feel tired, depressed, helpless, and other negative emotions, leading to a feeling of exhaustion and lack of motivation. Depressed and anxious people often feel lonely and misunderstood, which may make them lose the motivation to interact with others, possibly leading them to avoid social activities and intimate relationships and preferring to maintain a static state.

There was a significant difference (F=10.986, p=0.001) in daytime phone usage time between students in the anxious group and those in the non-anxious group. Figure 6(d) shows that students in the anxious group reduced their phone usage time. With news about COVID-19 constantly emerging, the epidemic is a challenge for everyone, which may cause people to worry about their own health and that of their family and friends, as well as their career development, leading to anxiety.

6. Mental Assessment and Prediction

In this section, we put the feature set into different machine learning models such as XGboost model [11], SVM model [37], RandomForest model [4], and KNN model [14] to predict PHQ-9 and SAS scores and evaluate them.

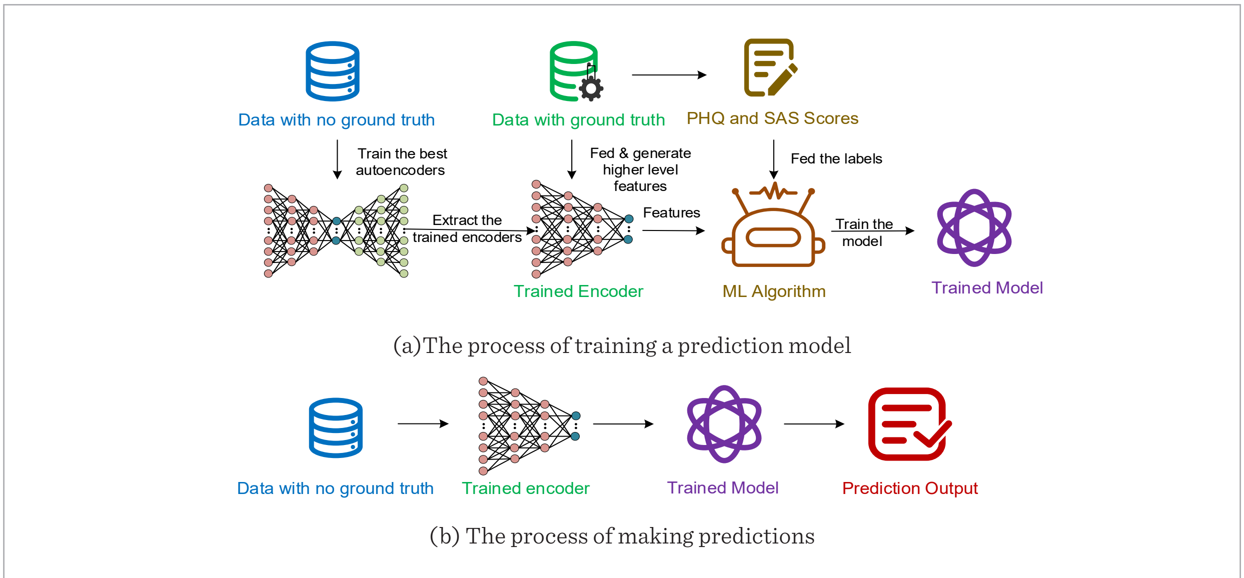
6.1. The Machine Learning Framework

The fundamental question of this study is whether it is possible to predict and diagnose depression and anxiety by observing various sensor data collected from personal smartphones. To answer this question, we first convert the raw data into features that can be understood and used by machine learning models. We cleaned the collected data to ensure its quality and completeness, including removing useless tags, filling in missing values, and handling outliers. We then extracted basis features from different mobile phone sensors: motion state (sitting, walking, running), walking steps, screen lock time, data flow usage size, charging time, WIFI connection time, daytime phone usage time, sleep time, phone usage time at night, and environmental loudness based on adjacent GPS coordinate points in the aggregated time series. These features were divided into audio, physical activity, and mobile interaction features.

According to the user self-test scores and scale rating standards, we divided PHQ-9 scores into five categories labeled 0-4, corresponding to different levels of depression. SAS scale scores were divided into four categories labeled 0-3, with higher category numbers indicating greater anxiety levels.

We randomly split 80% of the data into a training set and 20% into a test set. Unlike traditional machine learning algorithms used in previous studies, we chose a gradient boosting decision tree algorithm based on XGboost for this study because it has achieved good results in recent years and has efficient, flexible, and portable characteristics. It is capable of handling complex non-linear relationships, effectively capturing the underlying patterns between smartphone usage patterns and psychological questionnaire scores to provide accurate predictive outcomes. Furthermore, it can provide feature importance assessment, thereby identifying key factors that significantly impact psychological questionnaire scores. To further optimize model performance, we used grid search [24] combined with K-fold cross-validation [2]. We used the parameter combination with the highest predictive performance score as the optimal parameter setting in the final prediction model, Figure 7 shows the training and predicting process.

Figure 7
The process of our prediction scheme



Considering the imbalance of sample data category distribution, we used weighted model evaluation indicators for performance evaluation: weighted precision [22], recall [8], and F-1 scores [55]. We compared XGboost algorithm with other commonly used machine learning algorithms in previous studies, and the results are shown below.

Figure 8

Performance comparison between different algorithms

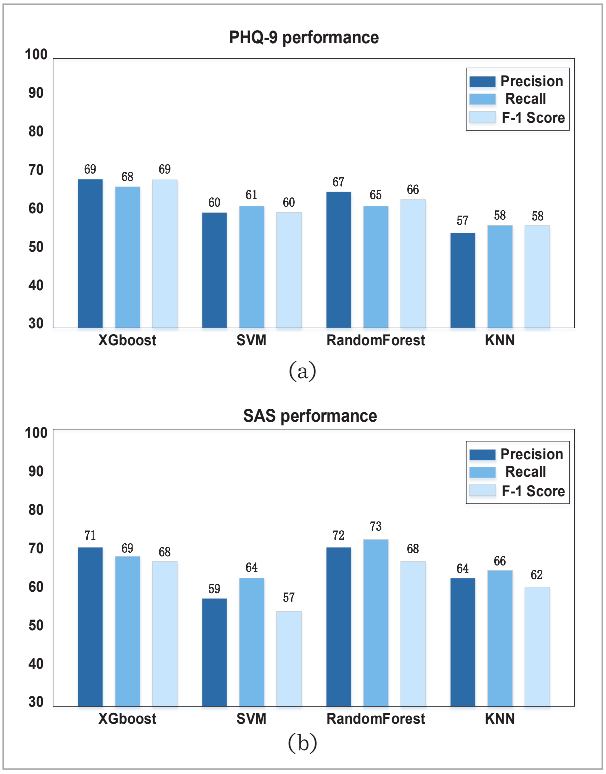


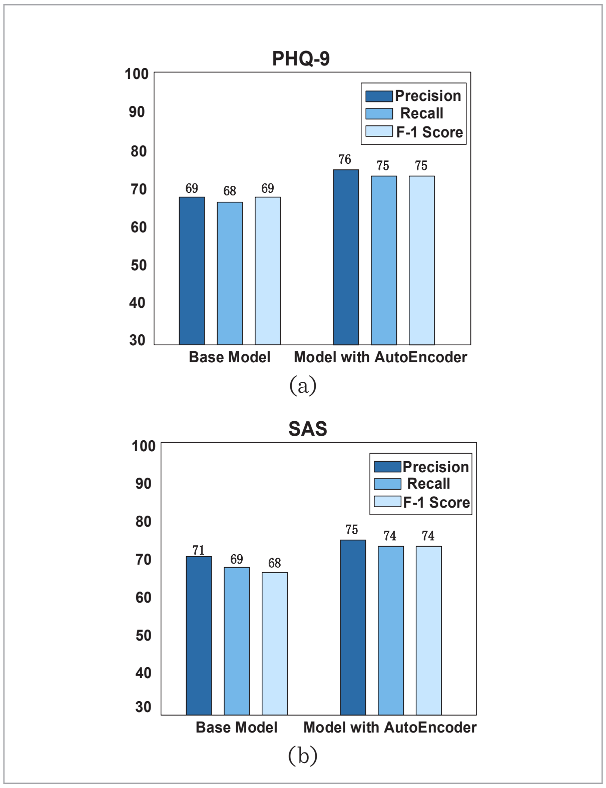
Figure 8 is the performance of different prediction model, from the experimental results, it can be seen that XGboost model has higher weighted average F-1 scores than other traditional machine learning algorithm models in SAS classification results but has the same F-1 score performance as Random Forest model. However, in PHQ-9 classification results, XGboost's F-1 score performance is higher than that of all other algorithm models' F-1 scores. Therefore, it can be concluded that XGboost model's overall performance is better than that of other models. The model achieved 69% and 68% F1-scores in PHQ-9 and SAS classifications, respectively.

6.2. Auto Encoder

On the basis of the established prediction model, we introduced an autoencoder to further improve its predictive performance. We hoped to extract effective features related to behavior patterns from mobile sensor data through autoencoder to further improve predictive model performance. The figure 9 shows the entire prediction framework structure after adding an encoder, which includes two parts: Figure a includes the training process of predicting student psychological stress model; Figure b represents the final prediction process of the model. First, we input training set data into an encoder for model training. Then we extract the trained encoding part to extract effective features from mobile sensor data using autoencoder and input them into machine learning models together with PHQ-9 and SAS category labels associated with these features as training data sets. After preparing dependent variables and transfer features, we train machine learning models. We use the trained

Figure 9

The performance of autoencoder



encoding network to extract effective features from test data sets and input them into psychological stress prediction models to generate prediction results.

After conducting comparative experiments with different models, we used the XGboost model as the final predictive model. We compared the performance of the predictive models constructed using features extracted by the encoder and manually created features, respectively. Weighted precision and recall were used as the evaluation metrics to assess the performance of the model. We trained a predictive model using a feature set that combined low-level features that closely resemble raw mobile sensing data and handcrafted features as the base model. The features extracted from the low-level features by the autoencoder were referred to as high-level features, and the predictive model trained using a feature set that combined both low-level and high-level features was referred to as the boosted model. The hidden layer number and the number of extracted features were set to 2 and 4, respectively, with the activation function set to Relu and the optimizer set to Adam. The performance comparison results of the models are shown in the figure.

The comparison results show that the stress prediction model trained using features extracted by the autoencoder outperformed the stress prediction model trained using manually created features. The F-1 score metrics for PHQ-9 and SAS improved by 6% and 4%, respectively, relative to the base model. Although the performance improvement was limited, it confirmed the feasibility of using an encoder to extract features from mobile sensing data.

7. Conclusion

In this paper, we have presented a comprehensive study on using mobile devices and wearable sensors for psychological monitoring in extreme situations, specifically during the COVID-19 pandemic. We collected the largest-scale data samples from undergraduate students at our university during the complete campus lockdown period and redesigned the modalities and correlations of sampling data to comprehensively reflect users' emotional states. We analyzed the correlation between sampling data and psychological state evaluation questionnaire scores and predicted changes in people's psychological states. Furthermore, we designed a machine learning-based detection network using an unsupervised autoencoder as input to detect students' emotional states with high accuracy. However, our approach is based on a general modeling method, which involves training a classification model for all test samples. Given the significant variability in individual behaviors, we will develop personalized models to further enhance model performance. We hope that our findings can contribute to the development of effective and accessible mental health monitoring systems that can help prevent severe mental health consequences.

Acknowledgement

This work was partly supported by the National Science Foundation of China, No.62463009, and the Province Science Foundation of Jiangxi, No.20224BAB202030.

References

1. Abuse and M. H. S. Administration. Key Substance Use and Mental Health Indicators in the United States: Results From the 2021 National Survey on Drug Use and Health, 2021 NSDUH Annual National Report, 2022, PEP22-07-01-005.
2. Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L., Ridella, S. The 'K' in K-Fold Cross Validation. In ESANN, 2012, 441-446.
3. akker, D., Kazantzis, N., Rickwood, D., Rickard, N. Mental Health Smartphone Apps: Review and Evidence-Based Recommendations for Future Developments. JMIR Mental Health, 2016, 3(1), 4984. DOI: 10.2196/mental.4984. https://doi.org/10.2196/mental.4984
4. Belgiu, M., Dragut, L. Random Forest in Remote Sensing: A Review of Applications and Future Directions. ISPRS Journal of Photogrammetry and Remote Sensing, 2016, 114, 24-31. https://doi.org/10.1016/j.isprsjprs.2016.01.011
5. Ben-Zeev, D., Chander, A., Tauscher, J., Buck, B., Nepal, S., Campbell, A., Doron, G. A Smartphone Intervention for People with Serious Mental Illness: Fully Remote Randomized Controlled Trial of Core. Journal of Medical Internet Research, 2021, 23(11), 29201. https://doi.org/10.2196/29201
6. Biddle, S. J., Ciaccioni, S., Thomas, G., Vergeer, I. Physical Activity and Mental Health in Children and

Adolescents: An Updated Review of Reviews and an Analysis of Causality. Psychology of Sport and Exercise, 2019, 42, 146-155. https://doi.org/10.1016/j.psychsport.2018.08.011

7. Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., Rubin, G. J. The Psychological Impact of Quarantine and How to Reduce It: Rapid Review of the Evidence. The Lancet, 2020, 395(10227), 912-920. https://doi.org/10.1016/S0140-6736(20)30460-8
8. Buckland, M., Gey, F. The Relationship Between Recall and Precision. Journal of the American Society for Information Science, 1994, 45(1), 12-19. https://doi.org/10.1002/(SICI)1097-4571(199401)45:1<12::AID-ASI2>3.0.CO;2-L
9. Calear, A. L., Christensen, H. Systematic Review of School-Based Prevention and Early Intervention Programs for Depression. Journal of Adolescence, 2010, 33(3), 429-438. https://doi.org/10.1016/j.adolescence.2009.07.004
10. Canzian, L., Musolesi, M. Trajectories of Depression: Unobtrusive Monitoring of Depressive States by Means of Smartphone Mobility Traces Analysis. Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2015, 1293-1304. https://doi.org/10.1145/2750858.2805845
11. Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T. XGBoost: Extreme Gradient Boosting. R Package Version, 2015, 0.4-2, 1-4.
12. Cleophas, T. J., Zwinderman, A. H. Bayesian Pearson Correlation Analysis. Modern Bayesian Statistics in Clinical Research, 2018, 111-118. https://doi.org/10.1007/978-3-319-92747-3_11
13. Flint, S. W. Stigmatizing Media Portrayal of Obesity During the Coronavirus (COVID-19) Pandemic. Frontiers in Psychology, 2020, 2124. https://doi.org/10.3389/fpsyg.2020.02124
14. Guo, G., Wang, H., Bell, D., Bi, Y., Greer, K. KNN Model-Based Approach in Classification. On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE, Springer, 2003, 986-996. https://doi.org/10.1007/978-3-540-39964-3_62
15. Harrison, V., Proudfoot, J., Wee, P. P., Parker, G., Pavlovic, D. H., Manicavasagar, V. Mobile Mental Health: Review of the Emerging Field and Proof of Concept Study. Journal of Mental Health, 2011, 20(6), 509-524. https://doi.org/10.3109/09638237.2011.608746
16. Hartley, K., Hoffman, B., Andujar, A. Smartphones and Learning: Evaluating the Focus of Recent Research. European Journal of Investigation in Health, Psychology and Education, 2023, 13(4), 748-758. https://doi.org/10.3390/ejihpe13040056
17. Huckins, J., DaSilva, A., Wang, W., Rogers, C., Wu, J., Murphy, E., Meyer, M., Wagner, D., Holtzheimer, P., Campbell, A. Mental Health and Behavior During the Early Phases of the COVID-19 Pandemic: A Longitudinal Mobile Smartphone and Ecological Momentary Assessment Study in College Students (Preprint). Journal of Medical Internet Research, 2020. https://doi.org/10.2196/preprints.20185
18. Jacobson, N. C., Summers, B., Wilhelm, S. Digital Biomarkers of Social Anxiety Severity: Digital Phenotyping Using Passive Smartphone Sensors. Journal of Medical Internet Research, 2020, 22(5), 16875. https://doi.org/10.2196/16875
19. Jafarlou, S., Azimi, I., Lai, J., Wang, Y., Labbaf, S., Nguyen, B., Qureshi, H., Marcotullio, C., Borelli, J. L., Dutt, N. D., Rahmani, A. M. Objective Monitoring of Loneliness Levels Using Smart Devices: A Multi-Device Approach for Mental Health Applications. medRxiv, 2023, 2023-06. https://doi.org/10.1101/2023.06.12.23291296
20. Kadison, R., DiGeronimo, T. F. College of the Overwhelmed: The Campus Mental Health Crisis and What to Do About It. Jossey-Bass, San Francisco, CA, 2004, 6.
21. Kroenke, K., Spitzer, R. L., Williams, J. B. The PHQ-9: Validity of a Brief Depression Severity Measure. Journal of General Internal Medicine, 2001, 16(9), 606-613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x
22. Lee, J., Kim, C., Kang, S., Shin, D., Kim, S., Yoo, H.-J. UNPU: An Energy-Efficient Deep Neural Network Accelerator with Fully Variable Weight Bit Precision. IEEE Journal of Solid-State Circuits, 2018, 54(1), 173-185. https://doi.org/10.1109/JSSC.2018.2865489
23. Li, Y., Li, G., Liu, L., Wu, H. Correlations Between Mobile Phone Addiction and Anxiety, Depression, Impulsivity, and Poor Sleep Quality Among College Students: A Systematic Review and Meta-Analysis. Journal of Behavioral Addictions, 2020, 9(3), 551-571. https://doi.org/10.1556/2006.2020.00057
24. Liashchynskyi, P., Liashchynskyi, P. Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS. ArXiv Preprint arXiv, 1912.06059, 2019.
25. Mack, D. L., DaSilva, A. W., Rogers, C., Hedlund, E., Murphy, E. I., Vojdanovski, V., Plomp, J., Wang, W., Nepal, S. K., Holtzheimer, P. E., Wagner, D. D., Jacobson, N.

C., Meyer, M. L., Campbell, A. T., Huckins, J. F. Mental Health and Behavior of College Students During the COVID-19 Pandemic: Longitudinal Mobile Smartphone and Ecological Momentary Assessment Study, Part II. Journal of Medical Internet Research, 2021, 23(6), 28892. https://doi.org/10.2196/28892

26. MacMillan, D. L., Reschly, D. J. Issues of Definition and Classification. In Ellis' Handbook of Mental Deficiency, Psychological Theory and Research, Routledge, 2012, 47-74.

27. Moschos, M. M. Physiology and Psychology of Vision and Its Disorders: A Review. Medical Hypothesis, Discovery and Innovation in Ophthalmology, 2023, 3(3), 83.

28. Nepal, S., Wang, W., Vojdanovski, V., Huckins, J. F., DaSilva, A., Meyer, M., Campbell, A. COVID Student Study: A Year in the Life of College Students During the COVID-19 Pandemic Through the Lens of Mobile Phone Sensing. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, 2022, 1-19. https://doi.org/10.1145/3491102.3502043

29. Panchal, U., Salazar de Pablo, G., Franco, M., Moreno, C., Parellada, M., Arango, C., Fusar-Poli, P. The Impact of COVID-19 Lockdown on Child and Adolescent Mental Health: Systematic Review. European Child and Adolescent Psychiatry, 2023, 32(7), 1151-1177. https://doi.org/10.1007/s00787-021-01856-w

30. Saeb, S., Lattie, E. G., Schueller, S. M., Kording, K. P., Mohr, D. C. The Relationship Between Mobile Phone Location Sensor Data and Depressive Symptom Severity. PeerJ, 2016, 4, 2537. https://doi.org/10.7717/peerj.2537

31. Saeb, S., Zhang, M., Kwasny, M., Karr, C. J., Kording, K., Mohr, D. C. The Relationship Between Clinical, Momentary, and Sensor-Based Assessment of Depression. 2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth), IEEE, 2015, 229-232. https://doi.org/10.4108/icst.pervasivehealth.2015.259034

32. Sanudo, B., Fennell, C., Sanchez-Oliver, A. J. Objectively-Assessed Physical Activity, Sedentary Behavior, Smartphone Use, and Sleep Patterns Pre-and During-COVID-19 Quarantine in Young Adults from Spain. Sustainability, 2020, 12(15), 5890. https://doi.org/10.3390/su12155890

33. Schuch, F. B., Bulzing, R. A., Meyer, J., Vancampfort, D., Firth, J., Stubbs, B., Grabovac, I., Willeit, P., Tavares, V. D. O., Calegaro, V. C., Deenik, J., López-Sánchez, G. F., Veronese, N., Caperchione, C. M., Sadarangani, K. P., Abufaraj, M., Tully, M. A., Smith, L. Associations of Moderate to Vigorous Physical Activity and Sedentary Behavior with Depressive and Anxiety Symptoms in Self-Isolating People During the COVID-19 Pandemic: A Cross-Sectional Survey in Brazil. Psychiatry Research, 2020, 292, 113339. https://doi.org/10.1016/j.psychres.2020.113339

34. Schuldt, C., Laptev, I., Caputo, B. Recognizing Human Actions: A Local SVM Approach. Proceedings of the 17th International Conference on Pattern Recognition, 2004, ICPR 2004, vol. 3, IEEE, 2004, 32-36. https://doi.org/10.1109/ICPR.2004.1334462

35. Sriwatanathamma, P., Sirivesmas, V., Simatrang, S., Bhowmik, N. H. Gamifying Cognitive Behavioral Therapy Techniques on Smartphones for Bangkok's Millennials with Depressive Symptoms: Interdisciplinary Game Development. JMIR Serious Games, 2023, 11(1), 41638. https://doi.org/10.2196/41638

36. St, L., Wold, S. Analysis of Variance (ANOVA). Chemometrics and Intelligent Laboratory Systems, 1989, 6(4), 259-272. https://doi.org/10.1016/0169-7439(89)80095-4

37. Stubbs, B., Vancampfort, D., Rosenbaum, S., Firth, J., Cosco, T., Veronese, N., Salum, G. A., Schuch, F. B. An Examination of the Anxiolytic Effects of Exercise for People with Anxiety and Stress-Related Disorders: A Meta-Analysis. Psychiatry Research, 2017, 249, 102-108. https://doi.org/10.1016/j.psychres.2016.12.020

38. Sygna, K., Aasvang, G. M., Aamodt, G., Oftedal, B., Krog, N. H. Road Traffic Noise, Sleep, and Mental Health. Environmental Research, 2014, 131, 17-24. https://doi.org/10.1016/j.envres.2014.02.010

39. Tarokh, L., Saletin, J. M., Carskadon, M. A. Sleep in Adolescence: Physiology, Cognition and Mental Health. Neuroscience Biobehavioral Reviews, 2016, 70, 182-188. https://doi.org/10.1016/j.neubiorev.2016.08.008

40. Thiese, M. S., Ronna, B., Ott, U. P-Value Interpretations and Considerations. Journal of Thoracic Disease, 2016, 8(9), 928. https://doi.org/10.21037/jtd.2016.08.16

41. Torous, J., Chan, S. R., Tan, S. Y.-M., Behrens, J., Mathew, I., Conrad, E. J., Hinton, L., Yellowlees, P., Keshavan, M. Patient Smartphone Ownership and Interest in Mobile Apps to Monitor Symptoms of Mental Health Conditions: A Survey in Four Geographically Distinct Psychiatric Clinics. JMIR Mental Health, 2014, 1(1), 4004. https://doi.org/10.2196/mental.4004

42. Turgoose, D., Maddox, L. Predictors of Compassion Fatigue in Mental Health Professionals: A Narrative Review. Traumatology, 2017, 23(2), 172. https://doi.org/10.1037/trm0000116

43. Wang, A., Chen, G., Yang, J., Zhao, S., Chang, C.-Y. A Comparative Study on Human Activity Recognition Using Inertial Sensors in a Smartphone. IEEE Sensors Journal, 2016, 16(11), 4566-4578. https://doi.org/10.1109/JSEN.2016.2545708

44. Wang, H., Da, L., Yang, L., Chu, S., Yang, F., Yu, S., Jiang, C. Colorimetric Fluorescent Paper Strip with Smartphone Platform for Quantitative Detection of Cadmium Ions in Real Samples. Journal of Hazardous Materials, 2020, 392, 122506. https://doi.org/10.1016/j.jhazmat.2020.122506

45. Wang, J., Qin, J., Zhong, H., Rajagopal, R., Xia, Q., Kang, C. Reliability Value of Distributed Solar + Storage Systems Amidst Rare Weather Events. IEEE Transactions on Smart Grid, 2018, 10(4), 4476-4486. https://doi.org/10.1109/TSG.2018.2861227

46. Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., Zhou, X., Ben-Zeev, D., Campbell, A. T. StudentLife: Assessing Mental Health, Academic Performance, and Behavioral Trends of College Students Using Smartphones. Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2014, 3-14. https://doi.org/10.1145/2632048.2632054

47. Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., Zhou, X., Ben-Zeev, D., Campbell, A. T. StudentLife: Using Smartphones to Assess Mental Health and Academic Performance of College Students. Mobile Health: Sensors, Analytic Methods, and Applications, 2017, 7-33. https://doi.org/10.1007/978-3-319-51394-2_2

48. Wang, R., Wang, W., DaSilva, A., Huckins, J. F., Kelley, W. M., Heatherton, T. F., Campbell, A. T. Tracking Depression Dynamics in College Students Using Mobile Phone and Wearable Sensing. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2018, 2(1), 1-26. https://doi.org/10.1145/3191775

49. Wang, W., Nepal, S., Huckins, J. F., Hernandez, L., Vojdanovski, V., Mack, D., Plomp, J., Pillai, A., Obuchi, M., DaSilva, A., Murphy, E., Hedlund, E., Rogers, C., Meyer, M., Campbell, A. First-Gen Lens: Assessing Mental Health of First-Generation Students Across Their First Year at College Using Mobile Sensing. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2022, 6(2), 1-32. https://doi.org/10.1145/3543194

50. Weir, B. S., Hill, W. G. Estimating F-Statistics. Annual Review of Genetics, 2002, 36(1), 721-750. https://doi.org/10.1146/annurev.genet.36.050802.093940

51. Wood, G., Goodyear, V., Adab, P., Al-Janabi, H., Fenton, S., Jones, K., Michail, M., Morrison, B., Patterson, P., Sitch, A. J., Wade, M., Pallan, M. Protocol: Smartphones, Social Media and Adolescent Mental Well-Being: The Impact of School Policies Restricting Daytime Use-Protocol for a Natural Experimental Observational Study Using Mixed Methods at Secondary Schools in England (SMART Schools Study). BMJ Open, 2023, 13(7). https://doi.org/10.1136/bmjopen-2023-075832

52. Xu, X., Chikersal, P., Doryab, A., Villalba, D. K., Dutcher, J. M., Tumminia, M. J., Althoff, T., Cohen, S., Creswell, K. G., Creswell, J. D., Mankoff, J., Dey, A. K. Leveraging Routine Behavior and Contextually-Filtered Features for Depression Detection Among College Students. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2019, 3(3), 1-33. https://doi.org/10.1145/3351274

53. Xu, X., Liu, X., Zhang, H., Wang, W., Nepal, S., Sefidgar, Y., Seo, W., Kuehn, K. S., Huckins, J. F., Morris, M. E., Nurius, P. S., Riskin, E. A., Patel, S., Althoff, T., Campbell, A., Dey, A. K., Mankoff, J. GlobeM: Cross-Dataset Generalization of Longitudinal Human Behavior Modeling. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2023, 6(4), 1-34. https://doi.org/10.1145/3569485

54. Yue, T., Li, Q., Wang, R., Liu, Z., Guo, M., Bai, F., Zhang, Z., Wang, W., Cheng, Y., Wang, H. Comparison of Hospital Anxiety and Depression Scale (HADS) and Zung Self-Rating Anxiety/Depression Scale (SAS/SDS) in Evaluating Anxiety and Depression in Patients with Psoriatic Arthritis. Dermatology, 2020, 236(2), 170-178. https://doi.org/10.1159/000498848

55. Zhang, D., Wang, J., Zhao, X. Estimating the Uncertainty of Average F1 Scores. Proceedings of the 2015 International Conference on the Theory of Information Retrieval, 2015, 317-320. https://doi.org/10.1145/2808194.2809488