An Initial Investigation of Mental Well-being Monitoring through Personal Healthcare Devices

Silvia M. Massa*
Eleonora Porcu*
Daniele Riboni*
silviam.massa@unica.it
e.porcu29@studenti.unica.it
riboni@unica.it
Department of Mathematics and Computer Science, University of Cagliari
Cagliari, Italy

ABSTRACT

Smart sensory devices, such as smart watches, scales, and blood pressure gauges, are increasingly adopted by individuals aiming to improve their health and fitness. Those devices gather extensive data about cardiovascular parameters, physical activities, sleep quality, and behavior. Thanks to data analytics and artificial intelligence algorithms, they provide insights into the health status of individuals. Derived data is used to support self-care interventions and to provide practitioners with additional health information acquired on a continuous basis. However, most of the current solutions focus on the physical dimension of health, while the mental dimension is often neglected. In this paper, we present the initial investigation of a system to recognize a wide range of psychological parameters, including behavioral inhibition/activation, anxiety, and stress, leveraging data acquired from personal healthcare devices. We experimented with the application of different supervised learning algorithms on features extracted from heart, sleep, and inertial sensor data acquired from a cohort of 21 individuals over 24 hours each. Our preliminary findings suggest that our method may yield promising outcomes in recognizing different aspects of mental well-being. However, due to the limited size of the used dataset, a more comprehensive experimental evaluation, with a broader number of participants and carried out over an extended monitoring period, is imperative to substantiate the results.

CCS CONCEPTS

• **Computer systems organization** → *Embedded and cyber-physical systems*; • **Human-centered computing** → *Ubiquitous and mobile computing*; *Personal digital assistants*; • **Applied computing** → *Health care information systems.*

*All authors contributed equally to this research.

UMAP Adjunct '24, July 01–04, 2024, Cagliari, Italy

ACM ISBN 979-8-4007-0466-6/24/07
https://doi.org/10.1145/3631700.3665240

KEYWORDS

Pervasive healthcare, Smart sensory devices, Self-care, Mental well-being, Machine learning.

ACM Reference Format:
Silvia M. Massa, Eleonora Porcu, and Daniele Riboni. 2024. An Initial Investigation of Mental Well-being Monitoring through Personal Healthcare Devices. In *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization (UMAP Adjunct '24), July 01–04, 2024, Cagliari, Italy*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3631700.3665240

1 INTRODUCTION

As emphasized by the World Health Organization, mental well-being is a fundamental factor of overall health, enabling individuals to harness their potential in learning, working, and fully realizing their abilities [10]. Numerous factors intertwine to influence mental well-being, including psychological and biological factors, social relationships, and exposure to particular environmental circumstances. In this regard, multiple studies have underscored a notable surge in the prevalence of psychological disorders, such as anxiety, stress, and depression, both during and following the COVID-19 outbreak, along with its consequential societal impacts [7–9]. Thus, there is an urgent need to devise effective measures for early detection of psychological issues and prompt intervention. Digital mental health tools [16], including mental health apps, have demonstrated moderate effectiveness in addressing various mental health conditions, notably anxiety and depression [15]. However, users often fail to sustain long-term engagement with mental well-being apps. Additionally, these apps frequently lack tools for objectively measuring shifts in the user's psychological well-being.

Nowadays, personal healthcare devices, such as smartwatches, smart scales, and other sensorized devices, are increasingly adopted by people to track various aspects of their physical health and well-being [3, 6, 12]. In this work, we explore the integration of similar technologies into a comprehensive framework for supporting mental health. By leveraging data from personal healthcare devices, users can be provided with a comprehensive understanding of how their daily activities, behaviors, and health-related data impact their mental well-being. This integration could not only promote sustained engagement with mental health interventions but also enable users to track their progress more effectively. Moreover, therapists may exploit additional data to enhance their services and provide personalized support to their patients for tailoring interventions to their individual needs.

In this paper, we illustrate our system architecture, the proposed algorithms, and the experimental evaluation carried out with a real-world dataset. We designed algorithms for preprocessing data acquired from personal healthcare devices, extracting statistical features, and training machine learning regressors for automatically assessing different kinds of psychological parameters, including behavioral inhibition/activation, stress, and anxiety. We considered data about sleep patterns (including sleep duration, interruptions, and fragmentation), heart rate, and physical activity. According to Papa *et al.* [11], sensor intrusiveness and comfort have a significant impact on perceived usefulness and ease of use. Among the most commonly used sensors, cameras, smartphones, and wearable sensors [4] are often perceived as intrusive and uncomfortable. For this reason, we decided to rely on unobtrusive smartwatches and sleep monitors. To evaluate our method, we conducted experiments using a public dataset acquired from a cohort of 21 individuals. Each individual was continuously monitored for 24 hours and was asked to fill out several questionnaires to evaluate different psychological aspects. The goal of our experiments was to investigate the feasibility of predicting psychological questionnaire scores based on the data acquired from personal healthcare devices.

We assessed the accuracy of our algorithms by evaluating the mean average error (MAE); i.e., the average absolute difference between the predicted and the actual scores. We observe that the MAE is relatively small compared to the domain of possible scores. However, the statistical significance of our experiments is severely limited by the small number of participants. Moreover, each subject was monitored for 24 hours only. This time frame is insufficient for fully capturing the intricacies of mental well-being. Indeed, mental health indicators can exhibit substantial fluctuations over longer durations, and a brief time frame may not offer a comprehensive perspective. Experiments with data acquired from a larger trial carried out for a longer time period are needed to substantiate our results. Having access to a larger training dataset will also provide us with the opportunity to explore more sophisticated machine learning techniques, including deep learning models, which may be better suited to handle the complexities inherent in psychological aspects. Summarizing, the main contributions of this work are the following:

- we present the main components of a system aimed at monitoring psychological parameters based on sensor data acquired from personal healthcare devices;
- we illustrate basic feature extraction methods and machine learning algorithms to predict the scores of psychological questionnaires based on sensor data;
- we present the results of a preliminary experimental evaluation with a real-world dataset.

The rest of the paper is structured as follows. Section 2 illustrates the architecture and algorithms proposed in this work, and the used dataset. Section 3 reports our experimental evaluation. Section 4 concludes the paper, discussing the limitations of the current work, and illustrating future research directions.

2 MATERIALS AND METHODS

In this section, we describe our system architecture, the used dataset, and the proposed techniques.

2.1 Architecture and dataset

Figure 1 illustrates our system architecture. Users wear a smartwatch that continuously acquires heart rate and activity data. Moreover, they use a sleep monitor to detect events like going to/going out of bed, sleep time, awakenings, and movements during sleep. These data are provided, on a daily basis, to our data integration platform, which preprocesses the data to remove noise. De-noised data are then processed to extract features of interest, as explained in Section 2.2. The resulting feature vectors are used by a regression algorithm to predict different psychological parameters, reported in Section 2.3, which are communicated to the individual and his/her therapists through different interfaces. The regression model is trained based on a labeled dataset acquired from a group of volunteers.

We evaluated our methods using the MMASH dataset [13], collaboratively collected by BioBeats and the University of Pisa, aims to explore the correlation between physical activities, sleep, and psychological states. After obtaining informed consent, data were acquired from 22 healthy male volunteers with homogeneous characteristics for 24 consecutive hours each through several wearable sensors. Acquired data include sleep duration and quality, beat-to-beat data, triaxial wrist accelerometer data, and scores from several questionnaires monitoring psychological and behavioral aspects. However, the dataset does not report sleep data for participant 11. Since sleep data are necessary for our system, we excluded participant 11 from the analysis. Hence, we conducted the experiments with the remaining 21 participants' data. The study was approved by the Ethical Committee of the University of Pisa. The dataset is available online[1].

2.2 Data preprocessing and feature extraction

The extraction of representative features from sensor data is essential to improve the prediction effectiveness of an ML algorithm. In this work, we used a sliding-window-based technique to extract statistical features from preprocessed data. We partitioned each volunteer's cardiac rhythm data and physical activity data into one-hour sliding windows. For each window, we computed the following features.

- Mean, standard deviation, kurtosis, skew, and entropy of the beat-to-beat intervals. According to [5], the maximum time between two beats should be about 1 sec; however, this value can fluctuate due to various factors. For this reason, before calculating features, interbeat intervals with a value greater than 3 seconds were deleted, as they were considered measurement errors.
- The total number of steps walked.
- The mean and standard deviation of both heart rate and movement vector magnitude.

In the case of data lack for a whole hour, the corresponding features were filled using the average between the features of the previous

[1] https://physionet.org/content/mmash/1.0.0/

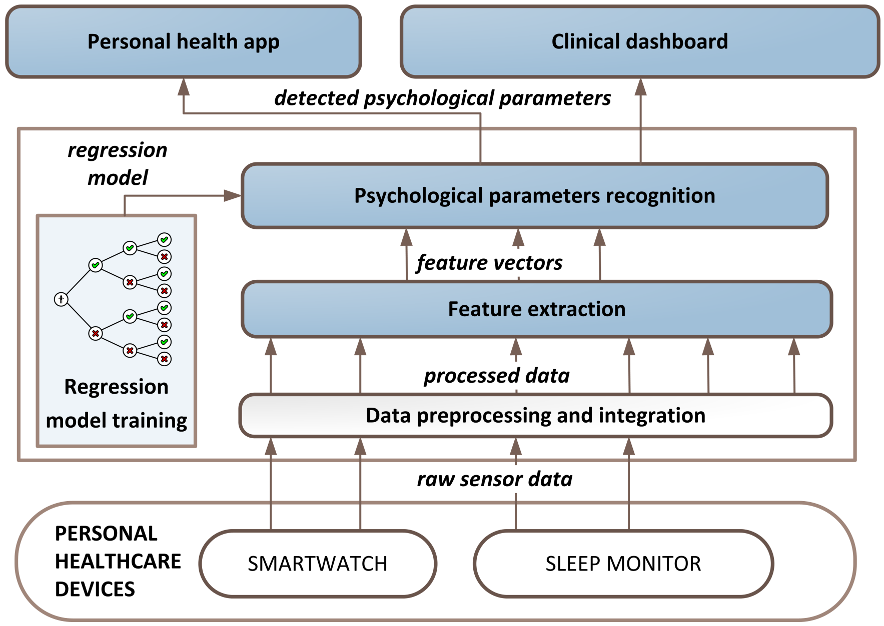

Figure 1: Our system architecture

and next hour, if present in the dataset, or using the features of the nearest hour.

We also extracted features related to sleep for the whole 24 hours, without division into sliding windows. Specifically, we considered these features: (i) in bed time, (ii) out of bed time, (iii) onset time, (iv) latency (time taken to get asleep), (v) total sleep time, (vi) total minutes in bed, (vii) efficiency (time asleep / time in bed), (viii) wake after sleep onset, (ix) number of awakenings, (x) average awakening length, (xi) fragmentation index, (xii) movement index, and (xiii) sleep fragmentation index. For participant 1, we merged the two sleep periods eliminating the waking period in the middle.

Considering that data were acquired from 9:00 of one day to 9:59 of the next day, each feature vector is built from 25 sliding windows. The structure of the feature vectors used to predict the different psychological conditions includes the sleep features, the features extracted from the 24-hour sliding windows, and the ground truth identified as explained in Section 2.3. As a consequence, every participant in the dataset was represented by an individual feature vector composed of 263 features plus the class label.

2.3 Ground truth on psychological parameters

The ground truth corresponds to the scores of the following psychological questionnaires. Each participant filled the DSI questionnaire before going to sleep, while the other ones were filled at the start of the data recording.

- The BIS/BAS questionnaire measures how individuals react to negative events (BIS) or to attractive stimuli (BAS). BAS is subdivided into: "drive" which describes the individual's perseverance in achieving goals; "fun-seeking" which is related to impulsivity and immediate reward due to sensational stimuli or risky situations; and "reward responsiveness" which indicates the predisposition to experience positive effects from reward-related stimuli [2]. In the Italian version employed for MMASH dataset acquisition, the BIS/BAS questionnaires comprise 20 scoring items, each rated on a 5-point Likert scale. Of these, seven items pertain to BIS (Inhibition), five items to Reward Responsiveness, four items to Drive, and four items to Fun-seeking dimensions.
- Daily Stress Inventory (DSI) measures an individual's stress level based on the frequency of stressors encountered in daily life. It is composed of 58 items, each assessing the intensity of a stressful event experienced during the last 24 hours. The intensity of each stressful event is rated through a 7-point Likert scale [1].
- STAI-Y questionnaires measure state and trait anxiety in adult subjects. The former refers to the current perception of anxiety in response to a temporary situation, while the latter represents a stable tendency to experience anxiety irrespective of current situations. Each scale consists of 20 items measuring the intensity of anxious feelings rated on a 4-point Likert scale [14].

3 EXPERIMENTAL EVALUATION

The goal of our experimentation is to evaluate the ability to automatically infer the scores of the psychological tests reported in Section 2.3 based on the observation of sensor data acquired from the personal healthcare devices of our platform.

We conducted our experiments using a leave-one-subject-out cross-validation approach: we use a participant's data for testing, and the data of the remaining participants for training, iterating the test on each participant. We executed those experiments using different regression algorithms of the Weka toolkit. We report the achieved results in terms of Mean Absolute Error (MAE). That metric computes the average magnitude of the difference between

Regressor	Inhibition (min 7 - max 35)	Reward responsiveness (min 5 - max 25)	Drive (min 4 - max 20)	Fun-seeking (min 4 - max 20)
Decision Table	2.71	2.81	1.74	3.36
Gaussian Processes	3.13	2.68	**1.59**	2.76
Linear Regression	3.12	2.51	1.71	2.75
M5 Rules	3.03	2.26	2.36	2.80
Random Forest	2.52	2.27	1.66	2.86
Random Tree	2.90	3.76	2.19	4.43
Simple Linear Regr.	**2.41**	**1.90**	2.35	3.46
SVM Regression	3.17	3.04	1.60	**2.64**

Table 1: Mean Absolute Error (MAE) across regression algorithms applied to BIS/BAS questionnaires. For each scale, we report the minimum and maximum score. We highlight in bold the MAE of the algorithms that achieved the best results for each scale.

Regressor	DSI (min 0 - max 406)	STAI-Y State (min 20 - max 80)	STAI-Y Trait (min 20 - max 80)
Decision Table	13.79	7.75	5.42
Gaussian Processes	10.23	8.01	5.95
Linear Regression	11.14	8.86	5.80
M5 Rules	16.56	9.12	6.84
Random Forest	12.38	7.33	**4.34**
Random Tree	20.43	13.95	6.45
Simple Linear Regression	**9.35**	**7.22**	5.20
SVM Regression	9.92	8.02	6.43

Table 2: Mean Absolute Error (MAE) across regression algorithms applied to stress (DSI) and anxiety (STAI-Y) questionnaires. For each scale, we report the minimum and maximum score. We highlight in bold the MAE of the algorithms that achieved the best results for each scale.

the actual score and the score predicted by the regression algorithm based on the feature vectors reported in Section 2.2.

Table 1 shows the results obtained with the recognition of BIS/BAS questionnaire scores. The table reports the four dimensions considered by BIS/BAS, together with their respective possible minimum and maximum scores. Among the tested algorithms, Simple Linear Regression obtained the best results for Inhibition and Reward responsiveness, while Gaussian Processes and Support Vector Machines (SVM) regressors obtained the best results for Drive and Fun-seeking dimensions, respectively. The Random Forest regressor obtained results close to the ones achieved by the best regressors in the recognition of all BIS/BAS scores.

In Table 2, we report the results obtained in the recognition of stress-related (DSI) and anxiety-related (STAI-Y) questionnaires. For DSI, we obtained the best results using Simple Linear Regression, which achieved the best results also in the recognition of STAI-Y State anxiety score. In the recognition of the STAI-Y Trait anxiety score, the algorithm achieving the best result was the Random Forest regressor. Moreover, the former obtained results very close to the ones achieved by the best regressors in the recognition of STAI-Y State anxiety score.

To the best of our knowledge, this is the first work that investigates the prediction of BIS/BAS, DSI, and STAI-Y psychological scores based on the observation of sensor data acquired from personal healthcare devices. Hence, a direct comparison with previous approaches is unfeasible. To further verify the performance of our models and compare the results on the different psychological aspects, we can rely on the normalized MAE (NMAE), which considers not only the average prediction error but also the range of possible ground truth values. To this aim, we divided the best result achieved for each psychological parameter, which is highlighted in bold in Table 1 and Table 2, by the corresponding value range (max-min). The NMAE value can vary between 0 and 1, where 0 indicates that the value is predicted without any errors, and 1 indicates the opposite. The NMAE with the lowest value was obtained on the prediction of the DSI scores (0.023), while the highest NMAE was obtained on Fun-seeking (0.165). Except for Fun-seeking and STAI-Y State, all other psychological states were detected with an NMAE of less than 0.1, which is a relatively small value.

4 CONCLUSION, LIMITATIONS, AND FUTURE WORK

Continuous monitoring of mental well-being can enhance treatment outcomes and improve overall health. In this paper, we presented an initial investigation of the basic components of a pervasive system monitoring various psychological aspects based on healthcare device data and machine learning algorithms. While preliminary experimentation indicated promising outcomes, this study is subject to different limitations. The significance of the experimental evaluation is limited by the small number of participants, and by the short monitoring period. Hence, additional experiments with data acquired from a larger trial are needed to substantiate the results. Moreover, due to the limited size of the training set, we were unable to explore deep learning algorithms, which seem better suited to predict complex psychological conditions than the basic regression algorithms employed in this work. Several research directions should be investigated in the future. We believe that integrating additional devices would allow us to increase the accuracy of the system through the extraction of more expressive features. We also intend to investigate feature selection techniques and collaborate with mental health professionals to investigate the monitoring of additional psychological conditions.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their insightful comments and suggestions. We acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.5 - Call for tender No.3277 published on December 30, 2021 by the Italian Ministry of University and Research (MUR) funded by the European Union – NextGenerationEU. Project Code ECS0000038 – Project Title eINS Ecosystem of Innovation for Next Generation Sardinia – CUP F53C22000430001-Grant Assignment Decree No. 1056 adopted on June 23, 2022 by the Italian Ministry of University and Research (MUR).

REFERENCES

[1] Phillip J Brantley, Craig D Waggoner, Glenn N Jones, and Neil B Rappaport. 1987. A daily stress inventory: Development, reliability, and validity. *Journal of behavioral medicine* 10 (1987), 61–73.

[2] Charles S Carver and Teri L White. 1994. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of personality and social psychology* 67, 2 (1994), 319.

[3] Erica Corda, Silvia M Massa, and Daniele Riboni. 2024. Context-Aware Behavioral Tips to Improve Sleep Quality via Machine Learning and Large Language Models. *Future Internet* 16, 2 (2024), 46.

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

[5] Tomas I Gonzales, Justin Y Jeon, Timothy Lindsay, Kate Westgate, Ignacio Perez-Pozuelo, Stefanie Hollidge, Katrien Wijndaele, Kirsten Rennie, Nita Forouhi, Simon Griffin, et al. 2023. Resting heart rate is a population-level biomarker of cardiorespiratory fitness: The Fenland Study. *Plos one* 18, 5 (2023), e0285272.

[6] William J Gordon, Adam Landman, Haipeng Zhang, and David W Bates. 2020. Beyond validation: getting health apps into clinical practice. *NPJ digital medicine* 3, 1 (2020), 14.

[7] Mariah T Hawes, Aline K Szenczy, Daniel N Klein, Greg Hajcak, and Brady D Nelson. 2022. Increases in depression and anxiety symptoms in adolescents and young adults during the COVID-19 pandemic. *Psychological medicine* 52, 14 (2022), 3222–3230.

[8] Karel Kostev, Kerstin Weber, Steffi Riedel-Heller, Christian von Vultée, and Jens Bohlken. 2023. Increase in depression and anxiety disorder diagnoses during the COVID-19 pandemic in children and adolescents followed in pediatric practices in Germany. *European Child & Adolescent Psychiatry* 32, 5 (2023), 873–879.

[9] Ida Kupcova, Lubos Danisovic, Martin Klein, and Stefan Harsanyi. 2023. Effects of the COVID-19 pandemic on mental health, anxiety, and depression. *BMC psychology* 11, 1 (2023), 108.

[10] World Health Organization et al. 2022. World mental health report: Transforming mental health for all. (2022).

[11] Armando Papa, Monika Mital, Paola Pisano, and Manlio Del Giudice. 2020. E-health and wellbeing monitoring using smart healthcare devices: An empirical investigation. *Technological Forecasting and Social Change* 153 (2020), 119226.

[12] Ivan Miguel Pires, Gonçalo Marques, Nuno M Garcia, Francisco Flórez-Revuelta, Vasco Ponciano, and Salome Oniani. 2020. A research on the classification and applicability of the mobile health applications. *Journal of personalized medicine* 10, 1 (2020), 11.

[13] Alessio Rossi, Eleonora Da Pozzo, Dario Menicagli, Chiara Tremolanti, Corrado Priami, Alina Sîrbu, David A Clifton, Claudia Martini, and Davide Morelli. 2020. A public dataset of 24-h multi-levels psycho-physiological responses in young healthy adults. *Data* 5, 4 (2020), 91.

[14] Charlers D Speilberger, RLt Gorsuch, R Lushene, PR Vagg, and GA Jacobs. 1983. Manual for the state-trait anxiety inventory. *Palo Alto, CA: Consulting Psychologists* (1983).

[15] Ashley Wu, Matthew A Scult, Emily D Barnes, Jessica A Betancourt, Avital Falk, and Faith M Gunning. 2021. Smartphone apps for depression and anxiety: a systematic review and meta-analysis of techniques to increase engagement. *NPJ digital medicine* 4, 1 (2021), 20.

[16] Samaneh Zolfaghari, Elham Khodabandehloo, and Daniele Riboni. 2022. TraMiner: Vision-based analysis of locomotion traces for cognitive assessment in smart-homes. *Cognitive Computation* 14, 5 (2022), 1549–1570.

Received 10 April 2024; revised XXX; accepted XXX