February 2023

WeHeart: A Personalized Recommendation Device for Physical Activity Encouragement and Preventing "Cold Start" in Cardiac Rehabilitation

Rosa VAN TUIJN [a] Tianqin LU [a] Emma DRIESSE [a] Koen FRANKEN [a]
Pratik GAJANE [a] and Emilia BARAKOVA [a]

[a] *Eindhoven University of Technology*

ORCiD ID: Rosa van Tuijn https://orcid.org/0000-0003-1731-3743, Tianqin Lu https://orcid.org/0000-0001-8033-4024, Emma Driesse https://orcid.org/0000-0003-3099-2242, Koen Franken https://orcid.org/0000-0003-3534-616X, Pratik Gajane https://orcid.org/0000-0002-8087-5661, Emilia Barakova https://orcid.org/0000-0001-5688-4878

Abstract. This paper highlights the importance of physical activity in cardiac rehabilitation as a means of reducing morbidity and mortality rates associated with cardiovascular disease. However, forming physical activity habits is a challenge, and the approach varies depending on individual preferences. We introduce WeHeart, a personalized recommendation device that aims to gradually increase physical activity levels and avoid a "cold start". WeHeart employs a Random Forest classification model that combines both measured and self-reported data to provide personalized recommendations. The system also uses Explainable AI to improve transparency and foster trust. Our study showcases the potential of Machine Learning in providing personalized recommendations for physical activity, and we propose a reinforcement learning approach to improve the system's personalization over time. Overall, this study demonstrates the potential of WeHeart in encouraging physical activity and preventing "cold start" in cardiac rehabilitation.

Keywords. Cardiac Rehabilitation, Physical Activity, Supervised Learning, Habit Formation, Recommendation System, XAI (Explainable AI)

1. Introduction

Cardiovascular disease (CVD) is a global public health concern, and cardiac rehabilitation (CR) is a vital multidisciplinary treatment for reducing secondary cardiovascular risk by improving both optimal physical and psychosocial functioning [1] through exercise training, nutritional interventions, smoking abstinence, etc. [2]. Physical activity (PA) is a key component of CR, with a focus on PA counselling and training [3]. Regular PA can protect the heart and reduce the incidence of cardiovascular events [4]. During CR participation, 70-85% of patients report meeting the recommended PA guidelines

[5, 6]. However, maintaining PA over the long-term after CR can prove challenging, with only 38-56% of worldwide patients continuing to meet the guidelines a year later [7, 8].

Thus, it is crucial to encourage PA habit formation among CR and post-CR patients. Habits are believed to promote long-term activity and prevent loss of motivation [9]. To form a habit, a behavior must be performed repeatedly in a consistent environment [10]. The variables that influence the continuation of behavior after initiation must also be taken into account [11]. However, these variables are unique to each individual, making it challenging to develop a one-size-fits-all solution.

As AI is at the heart of healthcare innovation [12], it creates opportunities in applications including personalized treatment and behavior modification [13]. Research has shown promise in the development of personalized recommendation systems through the exploration of deep neural learning models [14] as well as by developing situation-aware recommendation systems utilising wearables to collect several vital inputs [15]. Furthermore, some have proposed that AI recommendation systems utilizing reinforcement learning can help increase user engagement and satisfaction beyond only predicting the users' preferences alone [16].

However, despite the potential benefits, personalized recommendation systems pose several challenges. Firstly, relying solely on reinforcement learning may lead to a so-called "cold start", where the model lacks information about new users which is a common challenge in recommendation systems. Secondly, barriers of trust between the users and the AI can prevent users from using the system due to the user's limited knowledge of AI and the complexity of AI itself [12]. This "black box" problem, which is inherent in some forms of AI [17], can be addressed through explainable AI (XAI). XAI can help the users know how the AI works [18] and thereby increase transparency between the AI system and users, track influencing factors, and improve the predictive models [19]. Lastly, when designing for habit formation habitual tendencies should be considered by incorporating behavior change techniques, such as rehearsal and repetition or self-monitoring of behavior [20]. As this can help users to develop and maintain healthy habits that are beneficial to their health habits such as being physically active [21].

Based on this reasoning motivated by the supporting research, we hypothesize that a smart AI solution that fulfills these habitual tendencies, can prevent a cold start and by including XAI principles would demonstrate the potential of a personalized recommendation system in a CR context. This paper features the design of a personalized recommendation system called "WeHeart" that makes possible testing of this hypothesis and the initial validation of it. The system uses measured and self-reported data from CR patients to provide personalized PA recommendations that vary in time and intensity. WeHeart, placed in the patient's home, uses an algorithm that takes into account personal circumstances when making recommendations. To ensure patient understanding, WeHeart employs Explainable AI (XAI) to show how recommendations are calculated, using Shapley Values to explain the contribution of various factors. This approach helps patients form PA habits, supports long-term activity, and prevents loss of motivation.

2. Methods & materials

The designed system consists of a recommendation model and a physical prototype called "WeHeart". For the recommendation model, the dataset needs to be selected and preparated for the integration in the overall system. To take the first steps in creating an AI system that creates personalized recommendations for cardiac rehabilitation is by avoiding a cold start. Therefore, this system can provide a recommendation reflecting the current state of the rehabillitee by including self-reported and sensor data and incorporating them in a supervised learning algorithm, specifically a classification model. This is especially important in the context of this paper as inaccurate health recommendations can have undesirable effects. This classification model will make use of PMData [22], a public dataset containing logging data of 16 healthy adults over a period of 5 months. The dataset contains data measured with a Fitbit Versa 2 smartwatch and subjective reporting. This dataset was chosen as it included both sensor and self-reported data and was openly available.

To choose the appropriate recommendation based on the provided data, a supervised classification model is used. Since at this stage, it is unknown which algorithm is best for the system, a total of three algorithms will be compared: decision tree classifier, decision tree regressor, and random forest classifier, to choose the best performing algorithm for the recommender system.

3. Results

3.1. Data preparation and classification

The variables used are step count, active minutes, sleep quality, and self-reported data. “PAI” is an existing single indicator of PA [23] which provides an easy-to-understand measure, however, it is processed with an inaccessible algorithm. Therefore a so-called PA score is calculated to be able to grasp the PAI of the potential user. The PA score represents the recommended activity norm of 150 minutes of moderate intensity by the World Health Organization [24]. And just like the PAI it is based on the heart rate response to PA [23]. The PA score is calculated with the moderately active minutes based on the aforementioned recommendation from World Health Organization, as seen in Equation 1, which is then added as a separate column called “score” in the final algorithm. With a PA score greater than 100, the user is considered to have performed the standard amount of PA during the week.

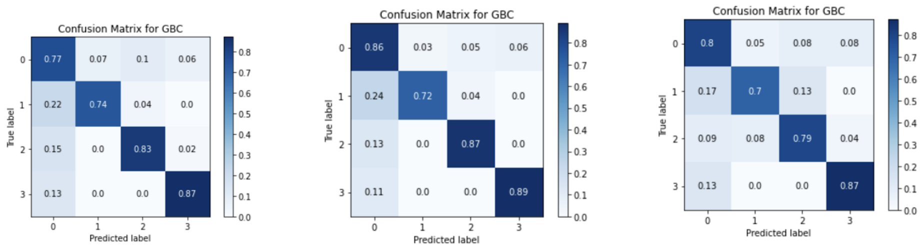
$$score = \frac{\text{sum of the moderately active minutes of the past 7 days}}{150} * 100\% \tag{1}$$

The data is classified with the score variable. Whilst the goal is to move towards a score of 100%, a user should not go from 0% to 100% at once. Therefore the mean score of the past 7 days is taken into account. If today’s score of the user is above the mean score of the past 7 days with an increment of 5% (as this only adds a small amount of 7,5 minutes of exercise), the user is not recommended to perform PA. However, if the score is below the described threshold the user is most likely recommended to exercise.

There are many factors that influence a user’s state. For example, sleep [25] or stress [26, 27] can influence how ready the user is for PA. WeHeart should take into account these factors. Besides sleep and stress, fatigue and mood [25, 28] are factors that can impact PA and are available to measure, either through wearable or self-reported data. Therefore, WeHeart will take into account these four features and will skew the recommendation based on this. If they are not already, each input will be translated to a score between 1 and 5.

3.2. Model selection and validation

The model is trained to estimate to which category a new observation belongs. The chosen classes for this model are Enough, Light exercise, Medium exercise and Intense exercise. The dataset is split into a training (80%) and a testing (20%) set. To select an appropriate algorithm for the classification model, three variations were considered: decision tree classifier, random forest classifier, and decision tree regressor. The results of the three different algorithms are compared by looking at the accuracy of the model and its confusion matrix, see Figure 1.



(a) Decision tree classifier; Accuracy: 0.80.

(b) Random forest classifier; Accuracy: 0.86.

(c) Decision tree regressor; Accuracy: 0.81.

Figure 1. Confusion matrices of the different classification models.

First, considering the accuracy, the Random forest classifier has the highest score of 0.86. Second, the confusion matrix shows the probability of classifying wrongly for all the different cases. The worst-case scenario is that someone who should be recommended to perform no exercise is recommended to do a high-intensity exercise. Which is the right top square of the confusion matrix. These cases are the lowest for both the Decision tree classifier and the Decision tree regressor. As the accuracy is the highest for the Random forest classifier, that algorithm will be used to train the model.

In the above cases, the model was trained on all participants together. However, as the goal is to produce personal recommendations a test was performed to see the accuracy of the model and if it would be trained on an individual level. However, this reduced the accuracy of the model to 0.57. Therefore, the model will not be trained on an individual level. Next to that, the classification learning curve showed that at least 400 cases are needed to reach the maximum efficiency of the model.

Graphical visualization of the WeHeart system can be found in Figure 2, which shows the input and output of the different data types into the classification model.

3.3. The physical prototype "WeHeart"

Whilst it is important that the model provides suiting and reliable recommendations, it is just as important to consider how this information is presented to its users. As the target group of this study mostly consists of older adults with limited experience with digital systems, it was decided to design a physical device. This device is placed in the living

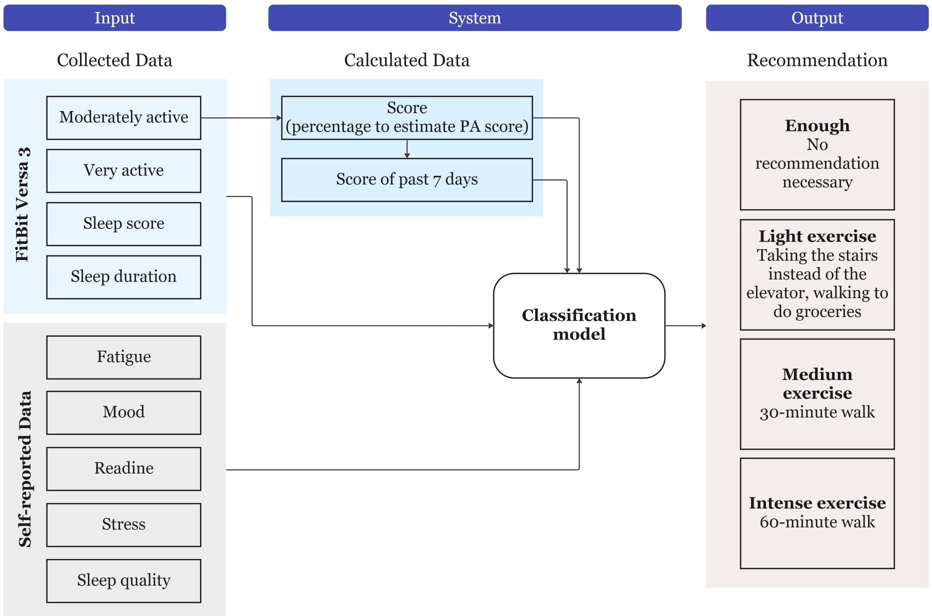


Figure 2. Input and output data of the classification model.

environment of the user and provides them to interact with the recommendation model by answering questions to provide self-reported data (such as stress and mood), see Figure 3a, evaluating their current PA score in relation to previous days and requesting a PA recommendation 3b. Once the user decides to accept the recommendation, they can activate the printing function by pressing the button at the top of WeHeart (Figure 3c). The recommendation includes the type of exercise, the intensity, and the duration. SHAP is used to visualize the contribution of each feature to the recommendation to make the algorithm more understandable.

(a) Self-reported data input interface.

(b) PA score overview.

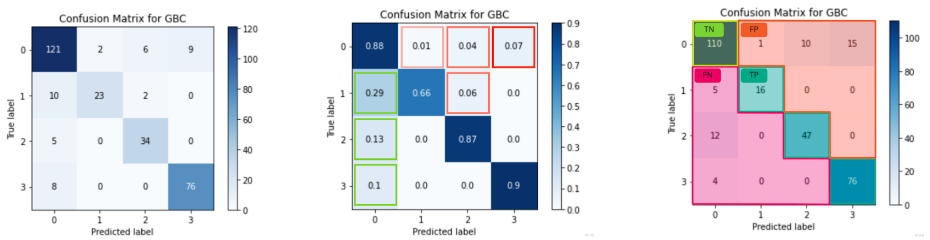
(c) Physical device.

Figure 3. Pictures of the physical prototype.

The accuracy of the model is 0.86 which says that 86% of the predictions were correct. This is insufficient information to determine how well the model is performing. Another metric to consider is null accuracy, which is the accuracy that could be obtained by consistently predicting the most frequent class (in this case the target 0). The null accuracy is 0.46. However, the confusion matrix should be examined to learn more about the classification model's performance (see Figure 4a).

As previously discussed in the preceding section, there exist various categories of errors, each with its own level of severity. These errors are represented by colored lines in Figure 4b. The green boxes are mistakes that do not cause any direct problems, as in these cases the algorithm proposes a lighter exercise than its original label (False Negatives). However, the red boxes can cause more severe problems, as a more intense exercise is proposed or the worst case scenario in which the user should be recommended to perform no exercise whilst the algorithm proposes a high-intense exercise (False Positives). The false positive rate was computed to be 0.19, which is sub-optimal as it ideally should be 0. The confusion matrix has been analyzed by calculating the sensitivity, specificity, False Positive Rate, and Precision. To calculate this, classes 1-3 are seen as positives, and 0 is seen as negative as it is especially important that the user is correctly recommended to either perform an exercise or not (see Figure 4c).

Lastly, the model is validated by a 5-fold validation that has been performed and showed that the mean Accuracy is 0.83 with a standard deviation of 0.01. This shows that the results of the model were not biased by the testing and training set used.



(a) With cases.

(b) With normalized values.

(c) Annotated with FP, TN, FN and TP.

Figure 4. Confusion matrices

3.4. Explainable AI

Considering the personalized recommendations, it is crucial that users can trust the system that they are working with. To achieve trust, we have adopted the Shapley Values as an XAI technique [29] as a visualisation of explaining the models' predictions by calculating feature importance, which provides a framework to interpret decision trees and ensemble tree models. WeHeart uses the global interpretability of the Shapely Additive exPlanations (SHAP) model in a Figma concept to demonstrate the extent to which each input variable contributes positively or negatively to the target variable. This greatly increases the transparency of the model used in this paper by explaining to users why they received this recommendation and the extent to which factors influence the results. For example, if a recommendation is different than expected, the user can see it is caused by the high level of stress that the user reported. To sum up, it allows the user to identify and compare the effects of the different factors and make changes accordingly.

4. Discussion

The objective of this research is to develop an AI system that could assist individuals undergoing cardiac rehabilitation in adopting physical activity habits, thereby reducing the likelihood of a future cardiac event [4]. As highlighted in the related works, ensuring that users can trust the recommendations is crucial in a healthcare setting, as offering inappropriate suggestions can have unfavorable consequences. Based on the performed tests and analysis a random forest classifier would be the most suitable model to use with existing data in order to reduce erroneous recommendations and thereby avoid a "cold start". However, at this point, the confusion matrix shows that there are errors where the algorithm proposes a more intense exercise than the user should do. This error is dangerous because it could lead to health risks for the rehabilitees and mistrust of the system. With the random forest classifier, the chance of such an error is the lowest but nonetheless, in future designs, these errors should be avoided. This could be done by applying certain thresholds. For example, by using ROC-Curve as a diagnostic map to find a threshold to minimize the False Positive Rate.

4.1. Limitations

While this paper presents promising results for a personalized recommendation system for cardiac rehabilitees, there are several limitations that must be considered before the system is used in practice. Some of these limitations were encountered during model development, while others require further consideration.

Firstly, the data set for this model contains data from healthy adults. To further evaluate the model's effectiveness, data should be collected from cardiac rehabilitation patients. It is important to acknowledge that the model's results may vary regarding gradual improvement, as healthy adults already have a baseline level of physical activity. Therefore, it is possible that utilizing data from the target population will lead to necessary changes in the model.

Secondly, the classification of the training and testing data can be further improved. Currently, existing data has been classified by looking at whether someone has reached the minimum amount of PA (score) and at self-reported data (e.g. stress level or mood). Literature showed that these factors have an influence on the performance of PA [25, 26, 27]. However, this classification could be improved to research in more detail the weight of different features.

Lastly, the current model also focuses insufficiently on gradual improvement. Specifically, it takes into account if one's current score has increased by 5% from the mean of the past 7 days. Although this prevents someone from going from 0 to 100 at once it is only classified if enough physical exercise was done. Instead of classifying if someone did or did not do enough exercise, other models such as a regression model could be used to actually predict someone's progress and improvements. This could make the recommendations suitable to the needs of the rehabilitee. There are also some limitations to the actual implementation and explainability of the system. While the tangible prototype of the system is designed to provide users with an explanation of the recommendations based on the Shapley Values visualization, its efficacy in promoting user understanding and trustworthiness of the results remains unclear as it has not been evaluated with rehabilitees. It may lead to users disregarding the results, as exemplified by the scenario where, due to bad sleep quality the recommendation is for lower PA than usual, the user might still go for a more intense workout. It might be important to prevent these situations and to also inform someone of the dangers of doing too much physical exercise immediately after a cardiac event [30].

4.2. Future Work

The current paper shows the development of a part of AI system for supporting PA of cardiac rehabilitees. The continuation of this research should include a reinforcement learning (RL) algorithm to further enhance the personalization of the system. RL can also address limitations such as gradual improvement, enhance the long-term effects and capture dynamic sentiments of users, which traditional recommendation systems fail to consider. Previous studies have employed RL in health-related recommendation systems, such as the "Actor-Critic" model proposed by Mulani et al. [31], which continuously updates information-seeking strategies based on user feedback. Another study by Zheng et al. [32], showed that RL methods can lead to recommendations with high concordance with clinicians' prescriptions and improve clinical outcomes.

Another potential avenue for further research in this field could explore the practical implementation of such a recommendation model in the homes of cardiac rehabilitees. By doing a study like this the model can be tested with actual patients and in a relevant context, in their home environment.

5. Conclusion

This paper reports the first phase of the design of a personalised embodied XAI system that will persuade CA patients to have the adequate amount of physical activity. The findings of this part of the study demonstrate that a classification model, incorporated into a physical system WeHeart, can be used to make personalised suggestions regarding PA, based on wearable sensor data and self-reported data. The use of this recommendation system has the potential to assist cardiac rehabilitees in incorporating PA into their daily lives and reducing the risk of future events and forming habit of doing PA at home. The classification algorithm of WeHeart could be used to start with personal recommendations. Further exploration of reinforcement learning algorithms could contribute to the quality and personalization of the model. Future work could focus on the actual implementation of such a recommendation model with cardiac rehabilitees and on the development of a reinforcement algorithm to make the PA suggestion more aligned with the preferences of the rehabilitees. Overall, the results of this research demonstrate the potential benefits of using data that is easily accessible from wearable sensor data and self-reported data to generate initial recommendations. In addition to the promising results obtained from the use of a classification model to generate personalized PA suggestions, the WeHeart system incorporates explainable artificial intelligence (XAI) in the form of Shapley values. This feature enables the system to provide clear explanations of how and where the recommendations are generated, thus increasing users' trust in the system. By incorporating XAI, WeHeart represents an important step towards developing more personalized, and transparent basis for recommendations as a critical feature for building user trust in AI-powered systems.

References

[1] Kariann R. Drwal, Bonnie J. Wakefield, Daniel E. Forman, Wen Chih Wu, Bjarni Haraldsson, and Ramzi N. El Accaoui. Home-based cardiac rehabilitation: Experience from the veterans affairs. *Journal of Cardiopulmonary Rehabilitation and Prevention*, 41:93–99, 3 2021.

[2] Sophie K. Kieffer, Nina Zisko, Jeff S. Coombes, Javaid Nauman, and Ulrik Wisløff. Personal activity intelligence and mortality in patients with cardiovascular disease: The hunt study. *Mayo Clinic Proceedings*, 93:1191–1201, 9 2018.

[3] Novella Martinello, Stephanie Saunders, and Robert Reid. The effectiveness of interventions to maintain exercise and physical activity in post-cardiac rehabilitation populations: A systematic review and meta-analysis of randomized controlled trials. *Journal of Cardiopulmonary Rehabilitation and Prevention*, 39:161–167, 5 2019.

[4] Shari S. Bassuk and Jo Ann E. Manson. Epidemiological evidence for the role of physical activity in reducing risk of type 2 diabetes and cardiovascular disease. *Journal of Applied Physiology*, 99:1193–1204, 9 2005.

[5] Chris M. Blanchard, Robert D. Reid, Louise I. Morrin, Lisa McDonnell, Kerry McGannon, Ryan E. Rhodes, John C. Spence, and Nancy Edwards. Demographic and clinical determinants of moderate to vigorous physical activity during home-based cardiac rehabilitation: The home-based determinants of exercise (home) study. *Journal of Cardiopulmonary Rehabilitation and Prevention*, 30:240–245, 2010.

[6] Dominique Hansen, Paul Dendale, Anita Raskin, Annick Schoonis, Jan Berger, Irmien Vlassak, and Romain Meeusen. Long-term effect of rehabilitation in coronary artery disease patients: randomized clinical trial of the impact of exercise volume. *http://dx.doi.org/10.1177/0269215509353262*, 24:319–327, 2 2010.

[7] Beth C. Bock, Rosa E. Carmona-Barros, Jeanne L. Esler, and Peter L. Tilkemeier. Program participation and physical activity maintenance after cardiac rehabilitation. *http://dx.doi.org/10.1177/0145445502238692*, 27:37–53, 7 2016.

[8] Kazuhiro P. Izawa, Sumio Yamada, Koichiro Oka, Satoshi Watanabe, Kazuto Omiya, Setsu Iijima, Yasuyuki Hirano, Toru Kobayashi, Yusuke Kasahara, Hisanori Samejima, and Naohiko Osada. Long-term exercise maintenance, physical activity, and health-related quality of life after cardiac rehabilitation. *American Journal of Physical Medicine and Rehabilitation*, 83:884–892, 12 2004.

[9] Benjamin Gardner and Amanda L. Rebar. Habit formation and behavior change. *Oxford Research Encyclopedia of Psychology*, 4 2019.

[10] Phillippa Lally, Cornelia H.M. Van Jaarsveld, Henry W.W. Potts, and Jane Wardle. How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology*, 40:998–1009, 10 2010.

[11] Ron Borland. Habits and temporality: a commentary on hall and fong's temporal self-regulation theory. *http://dx.doi.org/10.1080/17437191003700816*, 4:66–69, 2010.

[12] Felix Gille, Anna Jobin, and Marcello Ienca. What we talk about when we talk about trust: Theory of trust for ai in healthcare. *Intelligence-Based Medicine*, 1-2:100001, 11 2020.

[13] Arjun Panesar. *Machine Learning and AI for Healthcare Big Data for Improved Health Outcomes*. Apress, second edition edition, 2021.

[14] X Zhou, W Liang, Kevin I, Wang K, and S Shimizu. Multi-modality behavioral influence analysis for personalized recommendations in health social media environment. *IEEE Transactions on Computational Social Systems*, 6:888–897, 2019.

[15] Aldosary Saad, Hassan Fouad, and Abdallah A. Mohamed. Situation-aware recommendation system for personalized healthcare applications. *Journal of Ambient Intelligence and Humanized Computing*, 1:1–15, 2 2021.

[16] Qian Zhang, Jie Lu, and Yaochu Jin. Artificial intelligence in recommender systems. *Complex and Intelligent Systems*, 7:439–457, 2 2021.

[17] A. Ferretti, M. Schneider, and A. Blasimme. Machine learning in medicine:. *European Data Protection Law Review*, 4:320–332, 9 2018.

[18] Hui Wen Loh, Chui Ping Ooi, Silvia Seoni, Prabal Datta Barua, Filippo Molinari, and U. Rajendra Acharya. Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022). *Computer Methods and Programs in Biomedicine*, 226:107161, 11 2022.

[19] Urja Pawar, Donna O'Shea, Susan Rea, and Ruairi O'Reilly. Explainable ai in healthcare. *2020 International Conference on Cyber Situational Awareness, Data Analytics and Assessment, Cyber SA 2020*, 6 2020.

[20] Susan Michie, Michelle Richardson, Marie Johnston, Charles Abraham, Jill Francis, Wendy Hardeman, Martin P. Eccles, James Cane, and Caroline E. Wood. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine*, 46:81–95, 8 2013.

[21] Robert A. Carels, Jacob M. Burmeister, Afton M. Koball, Marissa W. Oehlhof, Nova Hinman, Michelle Leroy, Erin Bannon, Lee Ashrafioun, Amy Storfer-Isser, Lynn A. Darby, and Amanda Gumble. A randomized trial comparing two approaches to weight loss: Differences in weight loss maintenance. *Journal of Health Psychology*, 19:296–311, 2 2014.

[22] Vajira Thambawita, Steven Alexander Hicks, Hanna Borgli, Håkon Kvale Stensland, Debesh Jha, Martin Kristoffer Svensen, Svein Arne Pettersen, Dag Johansen, Håvard Dagenborg Johansen, Susann Dahl Pettersen, Simon Nordvang, Sigurd Pedersen, Anders Gjerdrum, Tor Morten Grønli, Per Morten Fredriksen, Ragnhild Eg, Kjeld Hansen, Siri Fagernes, Christine Claudi, Andreas Biørn-Hansen, Duc Tien Dang Nguyen, Tomas Kupka, Hugo Lewi Hammer, Ramesh Jain, Michael Alexander Riegler, and Pål Halvorsen. Pmdata: A sports logging dataset. *MMSys 2020 - Proceedings of the 2020 Multimedia Systems Conference*, pages 231–236, 5 2020.

[23] Amanda L. Hannan, Wayne Hing, Jeff S. Coombes, Suzanne Gough, Mike Climstein, Geoff Adsett, Rohan Jayasinghe, and James Furness. Effect of personal activity intelligence (pai) monitoring in the maintenance phase of cardiac rehabilitation: a mixed methods evaluation. *BMC Sports Science, Medicine and Rehabilitation*, 13:1–18, 12 2021.

[24] World Health Organization. Physical activity, 2022.

[25] Hugh H.K. Fullagar, Sabrina Skorski, Rob Duffield, Daniel Hammes, Aaron J. Coutts, and Tim Meyer. Sleep and athletic performance: The effects of sleep loss on exercise performance, and physiological and cognitive responses to exercise. *Sports Medicine*, 45:161–186, 2 2015.

[26] Peter Salmon. Effects of physical exercise on anxiety, depression, and sensitivity to stress: A unifying theory. *Clinical Psychology Review*, 21:33–61, 2 2001.

[27] Matthew A. Stults-Kolehmainen and Rajita Sinha. The effects of stress on physical activity and exercise. *Sports Medicine*, 44:81–121, 1 2014.

[28] Mark Hamer, Romano Endrighi, and Lydia Poole. Physical activity, stress reduction, and mood: Insight into immunological mechanisms. *Methods in Molecular Biology*, 934:89–102, 2012.

[29] Scott M Lundberg, Paul G Allen, and Su-In Lee. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 2017.

[30] Ray W. Squires, Leonard A. Kaminsky, John P. Porcari, Jeanne E. Ruff, Patrick D. Savage, and Mark A. Williams. Progression of exercise training in early outpatient cardiac rehabilitation: An official statement from the american association of cardiovascular and pulmonary rehabilitation. *Journal of Cardiopulmonary Rehabilitation and Prevention*, 38:139–146, 2018.

[31] Jayraj Mulani, Sachin Heda, Kalpan Tumdi, Jitali Patel, Hitesh Chhinkaniwala, and Jigna Patel. Deep reinforcement learning based personalized health recommendations. *Studies in Big Data*, 68:231–255, 2020.

[32] Hua Zheng, Ilya O. Ryzhov, Wei Xie, and Judy Zhong. Personalized multimorbidity management for patients with type 2 diabetes using reinforcement learning of electronic health records. *Drugs*, 81:471–482, 3 2021.