User Perception of Fitness Recommender Systems

ANONYMOUS AUTHOR(S)

We consider fitness recommender systems as a new opportunity to promote a healthy lifestyle. Given that behavior change is an arduous and challenging journey, it is necessary to understand if the recommendations satisfy users with careful planning. However, our understanding of user perception after receiving fitness recommendations is limited. Offline cross-validation methods are the most common approaches to validate recommender systems. But the ultimate evaluation must come from a user study where the users receive the recommendations in real-life contexts. This is especially important for fitness recommender systems where users' health is at stake. In this paper, we describe a novel fitness recommender system, and design a mobile app to deploy the fitness recommendations. Next, we conduct a one-month user study where we aim at evaluating the fitness recommendations. Using semi-structured interviews, we present insights about recommendation acceptance and users' perceptions, attitudes, and pain points while receiving fitness recommendations.

CCS Concepts: • **Information systems** → **Recommender systems**; • **Human-centered computing** → *Empirical studies in HCI*.

Additional Key Words and Phrases: Recommender systems, behavior change, physical activity, sedentary lifestyle, personalization, persuasive technology.

ACM Reference Format:
Anonymous Author(s). 2021. User Perception of Fitness Recommender Systems. In *CHI '21: ACM CHI Conference on Human Factors in Computing Systems, May 8–13, 2021, Online Virtual Conference.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/xx.xx

1 INTRODUCTION

Obesity, diabetes, and heart diseases are some of the most serious health problems of modern societies. Many studies and reports identify inactivity as one of the primary factors [11, 14], and claim that being inactive is as dangerous as smoking [18]. To attain a more active lifestyle, individuals must make changes in daily habits and behaviors. The difficulty of such a behavior change lies not only in contemplating the change itself but often in figuring out how to adopt an efficient yet safe pattern over time. In particular, it is necessary to select the most suitable external intervention or the behavior change technique. As many previous studies show, people may set easy goals that make them soon lose motivation and therefore relapse [10], set goals too high that they risk injuring themselves or getting frustrated [6, 15], or take on a challenge that does not work for them.

The prevalence of wearable devices and smartphones enables us to track users' activities in an unprecedented way. This creates novel opportunities for recommender systems; for instance, machine learning (ML) based approaches to understand users' responses to external interventions and recommend personalized and effective interventions to facilitate behavior change. These interventions could vary for each user. For example, one user could be recommended to team up with an exercise buddy and another one to set a specific goal such as 12,000 steps per day. Despite the enormous studies on recommender systems [4], research in the area of *fitness recommender systems* is limited to a few studies [2, 3, 9, 13, 17]. These methods are limited as they rely on manually specified rules [17], use one-size-fits-all approaches modeling all users with a fixed daily goals [13], or abandon information from other users so they could not support inactive users with novel and effective recommendations [9]. Recent work has developed a data-driven approach recommending fitness strategies [3]. But this system generates recommendations for in-race marathon runners and not for normal fitness tracker users who tend to increase their level of activeness in the long-term.

On the other hand, the evaluation of fitness recommender systems demands more attention from researchers. Despite the potential impact of such recommender systems on users' behavior change, little is known about how users might perceive fitness recommendations. It is highly important to evaluate user experience and attitudes in real-life contexts where they are using fitness recommendations. One important notion to understand is 'recommendation acceptance' [16], whether users might accept or decline the recommended strategies. Also, it is important to understand users' perceptions, attitudes, and pain points while using fitness recommender systems.

In this paper, first, we introduce a novel fitness recommender system to increase the level of activeness (Section 2). We propose an intervention profiling pipeline to identify insightful information from past users to generate relevant recommendations for new users. These predictions are used to create a recommendation strategy where an optimal personalized intervention is recommended for each user. To deploy recommendations, we design and develop a prototype mobile app. Second, we report our user study where we conducted a one-month experiment with eleven participants (Section 3). Our study includes two phases: pre-recommendation and recommendation phases. In the pre-recommendation phase, we collected participants' baseline data for two weeks to learn which intervention would be the most suitable for them. During the follow-up recommendation phase, participants received actual recommendations. At the end of the one-month, we conduct interviews to collect feedback from participants. Last, we report our findings including qualitative insights and quantitative data including step numbers and engagement with the mobile app (Section 4). Our study contributes to the design of fitness recommender systems by shed lighting on contextual understanding of how users perceive fitness recommendations using a mobile app.

2 DESIGN

The architecture of the system is depicted in Figure 1. Our fitness recommender system includes three main parts: intervention profiling (Section 2.1), recommendation strategy (Section 2.2), and a mobile app (Section 2.3).

2.1 Intervention Profiling

2.1.1 Dataset. To build the system, we used a dataset collected in our previous lab study. The dataset includes eight weeks of fitness tracker data from 60 individuals. After three weeks of wearing fitness trackers, these individuals received one of two digital interventions. One group practiced "self-reflection" being able to see fitness information only about themselves. The other group practiced "peer-reflection" participating in physical activities together with a partner (i.e., seeing their own and their partners' step numbers). We used this dataset to predict how new users would respond to each intervention, given their pre-intervention data. This phase consists of three steps: data pre-processing, representation learning, and predictive modeling.

2.1.2 Data Pre-processing. After inspecting the dataset, we observed that in 17% of the pre-intervention days, people were not active at all. Given that the users likely forgot to wear their fitness trackers, we used the hot-deck method to impute missing data. More concretely, we replaced the missing daily time series data with a random valid daily time-series data generated by the same user on the same day of the week. Then, we performed data aggregation and

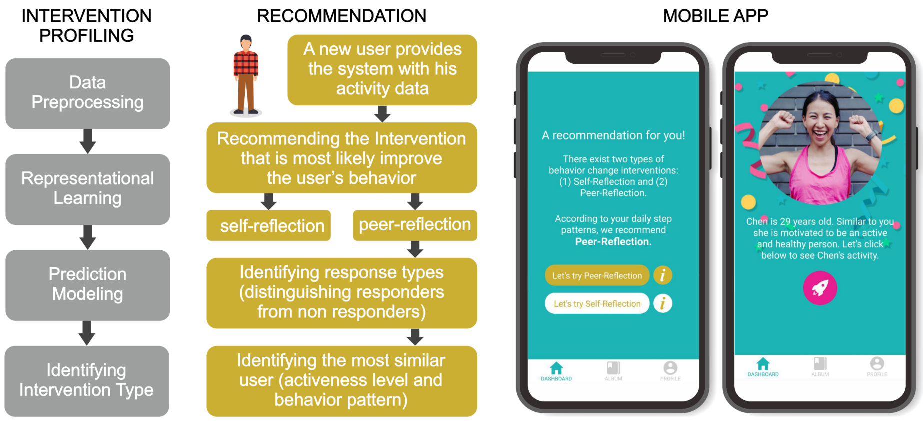

Fig. 1. Architecture of the fitness recommender system including behavior pattern extraction (left), recommendation strategy (middle), and the mobile app (right).

transformation. The aggregation step included segmenting each sequence into 10-minute non-overlapping sliding windows and summing up the minute-level step counts belonging to the same window. The transformation step included performing a box-cox transformation [1] on the minute-level step count observations to obtain transformed observations having a normal distribution.

2.1.3 Representation Learning. We used Recurrent Neural Network (RNN) autoencoder [19] to reduce the dimensionality of each daily time-series and generate embeddings (representing each time-series as a lower-dimensional data point) that preserves the low-level information as much as possible. We used these embeddings for the predictive model described in Section 2.1.4.

2.1.4 Predictive Modeling. To predict how a new person would respond to a given intervention our recommender system utilizes the pre-intervention sensor data. We defined response as the relative improvement of the person's post-intervention average daily step count compared to her/his pre-intervention average daily step count. We trained supervised ML models separately on the data from the people that received the self-reflection intervention and the people that received the peer-reflection intervention. Our models used the features extracted from the pre-intervention time-series data as predictors of user's relative improvement. We evaluated different ML models for our supervised learning task. Ridge regression model [12] performed the best in our evaluation, and this is why we used it in our recommender system. The input to this model is a set of embeddings (produced by the autoencoder) describing the user's behavior on each day of the week and a set of total daily step counts associated with each day of the week.

2.2 Recommendation Strategy

Before generating recommendations, the recommender system requires monitoring a user for a while. A new user provides the system with her/his activity data before the recommendations (i.e., pre-recommendation behavior). After identifying the behavior patterns of a specific user, the system generates recommendations in two stages:

2.2.1 Intervention Recommendation. Our recommender system can (i) predict the potential improvement under each different intervention and (ii) recommend the intervention that is more likely to improve the user behavior. For example,

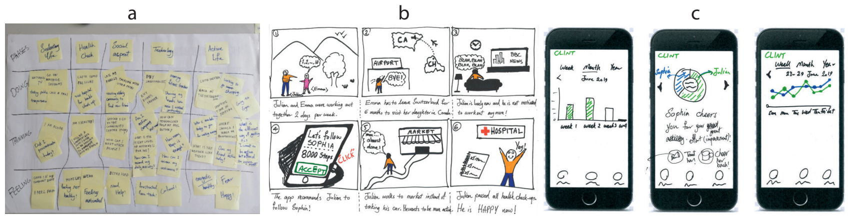

Fig. 2. (a) to-be scenario map, (b) storyboard, (c) low-Fidelity prototypes.

given our dataset, our recommender system identifies whether a new user should receive the "self-reflection" or the "peer-reflection" intervention, and then recommends the intervention with a higher possibility of improvement.

2.2.2 Peer Recommendation. Next, only for those users whose recommendations are "peer-reflection", the system goes further and provides more fine-grained recommendations. The recommender system identifies the best *role model* from the dataset that can persuade the user for higher physical activity and recommends the user to get paired with the role model. The user and the role model (together) should take a certain number of step numbers determined by the app. Next, we explain our rationale behind selecting the role model in details:

The system differentiates individuals (in the dataset) based on their responses to the peer-reflection recommendations they received. We considered all individuals that received the peer-reflection intervention and analyzed two consecutive days. For every two consecutive days, we determine whether the individuals responded or not to the intervention. We were interested in *responder individuals* those who increased their activity levels on the second day. Using data only from the responder individuals is likely to ensure the effectiveness of the recommendations, where a new user can increase her/his level of activeness by following active individuals. Among the responder individuals, the system seeks similarities in 24-hour activity patterns in a recent day and only choose the most similar responder individuals. The similarity includes both the level of activeness and the activeness pattern. In other words, for sedentary users, the system only selects individuals who were sedentary before the recommendation and they succeeded to change after the recommendation. Also, the system chooses individuals with similar activity patterns. For example, if a user is more active during night hours, the system omits those individuals who are active in the morning hours. This approach guarantees the feasibility of the recommendations for sustainable physical activity improvement. Thus, new users will not get frustrated or injured due to receiving unreachable goals.

2.3 Mobile App

Figure 2 demonstrates our design attempt. We employed a systematic interaction design approach following Cooper et al. [7]. We started with drafting a *to-be scenario map* brainstorming different issues users might face when using a fitness recommender app. Next, we created a *persona* and drew a *storyboard* as a part of our visual thinking. We also constructed *context scenarios* narrating everyday life stories about ideal user experiences. Next, we created low and high fidelity *prototypes* and a *journey map* before developing the actual mobile app. We implemented the mobile app using `Expo` [8], where we were able to build and deploy the app on both `Android` and `iOS` platforms. The recommender system was fully implemented at the back-end and accessed by the mobile app client through an API developed in `Python`. The mobile app could communicate with the back-end server to get the data including current step numbers and recommendations for the logged user, and could graphically display it in the front end. The recommender system notifies users with novel recommendations, once they opens and refreshes the app.

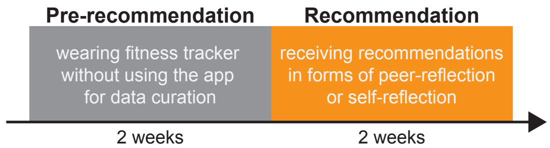

Fig. 3. Timeline of the User Study.

3 EXPERIMENT

We conducted a 4-week experiment which consists of 2-week pre-recommendation and 2-week recommendation phases (Figure 3). We recruited eleven participants (4 females and 7 males) with an age range between 20 and 35. Except for one, all were university students (7 Ph.D., 3 M.Sc.). Most of the participants had experienced the use of fitness trackers in their everyday life. But none of them had used a fitness recommender app. After collecting demographic information, we distributed `Fitbit Charge 3` fitness trackers among our participants. We started the two-week pre-recommendation phase, collecting users' fitness data without using the recommender app. Participants used `Fitbit` original app to sync and track their step numbers. During the pre-recommendation phase, participants did not receive any recommendations. We also asked participants not to use any other fitness tracker or fitness app during the experiment. We fed the recommender system with the collected data to generate the recommendations in the second phase.

After two weeks, we installed the app on participants' smartphones. Right after logging into the app, participants received a recommendation. The app suggested the most appropriate intervention for a given user. Based on participants' physical activity patterns on the week before, they either received self-reflection or peer-reflection. Participants could comply or select the opposite recommendation. After one week of use, participants received the opposite recommendation.[1] We switched the recommendations in the second week since we were interested in collecting feedback about both types of interventions. At the end of the recommendation phase, we conducted semi-structured interviews, where we examined users' requirements and concerns regarding two weeks of experience of using the recommender app. We asked how did they interact with the recommender app, what was their motivation for physical activity, which intervention did they prefer, and how did they find the recommendations. After the interview, we allowed participants to keep the fitness trackers with themselves. We transcribed all audio recordings of the interviews and applied thematic analysis [5] to analyze the interview content. We created an initial list of codes, including more than 200 meaningful codes, and generated respective themes. Last, we found relevant quotes from participants to report in the paper. Participants were anonymized and identified using numbers.

4 MAIN FINDINGS

In the first week of the recommendation phase, three participants (P1-P3) received the self-reflection, and the remaining eight participants (P4-P11) received the peer-reflection. Three participants (P1, P7, P8) in both the first and second weeks received the personalized intervention, but the other eight participants only in the first week received a personalized intervention. None of the participants declined the recommended interventions.

4.1 Qualitative Findings

4.1.1 Satisfaction. The majority of participants mentioned, overall, they ***liked*** the concept. Six participants found the concept novel and interesting. Five participants found the app effective for increasing their fitness level. [P11]: "*The*

[1]Those who received self-reflection in the first week of the recommendation phase received peer-reflection in the second week and vice versa.

app has some extra elements which I've never seen before on other apps. For example, the idea of recommendation is quite novel for me. Especially the peer reflection and receiving every day a new peer is really nice." [P5]: "*It was motivating. I remember when I had to move somewhere, I tried to walk and not taking the bus or metro.*"

4.1.2 When peer recommendation meets the expectations. Participants who were recommended with peer-reflection reported being ***interested in the social aspects*** of the app, where they were always curious to know how their peer was doing. [P7]: "*I liked more working with other guys. If I do it alone, it is less likely to walk more.*" [P11]: "*I am the kind of person who likes to do everything in a group while interacting with other people. I think the idea of interacting with people through my step numbers is exciting.*" They also reported checking the app more in the peer reflection since they were ***curious about their partners***. [P9]: "*During the peer-reflection, I opened the app much more because I wanted to see how many steps my partners had. That I don't do it usually.*"

One of the participants described his motivation to walk more with the peer-reflection. [P11]: "*During the recommendation, several times I walked to school instead of taking the bus. But it did not happen for me before.*" Three participants mentioned the ***novelty*** of the peer-reflection was the reason to prefer it. [P6]: "*All these apps in the market have typical 10,000 steps goal. I also can see in my tracker itself when I reach 10,000 steps as it usually congrats me. But this [peer-reflection] is something new. So, there is a reason for me to use it.*" In addition, two participants mentioned preferring peer-reflection because the recommended goals were ***adaptive*** and thus ***feasible*** to achieve. [P4]: "*Daily goals were feasible for me. Sometimes I received 13,000 steps in total with my peer. Sometimes it was a total of 19,000 that I could not achieve and the next day it reduced again. But the typical* `Fitbit` *goal (10,000 steps) is not achievable for me. So, I've had more achievements with the peer-reflection recommendations.*"

Two participants reported feeling ***supported*** to achieve their goals using the peer-reflection recommendation. [P5]: "*Some day I got recommended a very high total goal. I was busy and I realized that I cannot be super active on that day. But when I checked the app in the evening, I was so surprised that my peer accomplished much more than needed like more than three quarters. It was nice!*" Last, one of the participants reported he liked the ***ambiguity*** in the app when he needs to speculate about peers' step numbers. [P6]: "*You see that somebody else is working for the same target and you can cover him, and it can motivate you. The interesting thing is that there is no certain goal and you don't really know how much you should cover as it depends on your peer. Even you may think 'Okay probably he will fail!' so you should walk more to compensate your peer. This ambiguity during the day is interesting.*"

4.1.3 Self-reflection satisfies self-reliant users. Three participants who received self-reflection mentioned they would like to be ***self-dependent*** and do not rely on someone else for achieving goals. [P2]: "*I don't like the idea of depending on someone. I would like to set my own goal. In my everyday life, I aim to be active during the day. That is my routine. I know that I should either walk or go to the gym. I do not really care if others are going to walk or not to walk. It doesn't bother me. It's my activity level that bothers me. So, I don't really want to change my routine just because some other guy walks or doesn't walk enough.*" [P3]: "*The self-reflection is enough for me to get motivated. Because I always have to walk more than 10,000 steps no matter how many steps the other person needs to walk.*"

These participants mentioned they found goals in the self-reflection somehow ***easier to plan***. [P2]: "*I think, in the peer-reflection, my peer was not active. So, I was not motivated to check the app. You know, the goal in the peer scenario appears from somewhere in blue and I didn't really understand it. You received some target and you cannot estimate how to achieve it because you really don't know the peer. But in the self-reflection, I can easily plan for 10,000 steps and I know how to achieve it.*" But, four participants mentioned they would ***prefer to see adaptive goals*** instead of 10,000 steps per day. [P2]: "*I wish my goal could be adjustable like a personal coach that adjusts the training plan depending on my* *well-being or my situation. I think if I use this app for a sufficiently long time the app could learn some patterns and then the recommendation might fluctuate from day to day, for example, saying that someday I should be more active and the other day I can be less active.*" Two participants believed having adaptive goals in the self-reflection is a better idea. But they think the goals should be set by the recommender system, and not by users.

4.1.4 Recommendations should be transparent. Four participants suggested that the app should provide ***transparent*** recommendations, for example, by giving a more descriptive introduction about their peers, where such a description will help them to do better daily planning. [P1]: "*I think the description of peers is short and generic. Maybe it should be more concrete. Something like usually this person is more active in the evening. For example, I know that if he doesn't walk 3,000 more steps until the rest of the day, I need to do it by myself.*"; [P6]: "*The introduction should be more elaborative. Maybe interacting with this person can be another option. Even though I like the ambiguity in the app, I prefer also to communicate with this person to be able to plan for my day. Like if the peer could say: 'Hey! today I am hiking and I will walk the entire day' then I can plan my activity accordingly.*"

As mentioned by P6, participants were curious to interact with their peers to better plan for their daily steps in the peer-reflection mode. Eight participants mentioned they preferred to ***know their peers in-person***. [P9]: "*It would be better if I know my partner, then I can contact her or him saying: 'Hey! go and move a little bit!' But it is difficult to ask strangers to walk for me.*" One of the participants guessed that her peers were virtual, and she thought it would be interested if recommendations are generated based on real partners instead. [P4]: "*I know these are virtual peers but maybe this can impact you if you know that these are real-time people who also see the same goal at the same time as me. If you get paired up with a peer in real-time might be more beneficial.*"

4.1.5 Active users need calorie-based recommendations. Among participants, four of them had an active lifestyle. P2, P6, and P10 were regularly exercising in fitness centers. P6 was the most active user who was additionally playing tennis every week. In addition, P5 mentioned she was doing regular exercise in her house before starting the day. These participants reported ***the step count is not the best way to represent the recommendations***. [P6]: "*Steps are not even half of my energy burn during the day. Because if I go to the gym, I burn most of my calories not by steps. Also, when I play tennis maybe I cover 5,000 steps in an hour, but the energy consumption will be much higher because I spend a lot of energy while hitting the ball, during accelerations, or jumps that they consume energy much more than usual walking or running. The tracker can measure the burned calories, and it is kind of reliable because it reflects my day more compared with my step numbers.*" [P5]: "*I would honestly prefer to see everything based on calories then I would be definitely motivated.*" To conclude, it is suggested that for users with higher fitness levels, the recommender systems should provide recommendations based on calorie expenditure and not the step count.

4.2 Quantitative Findings

Here we provide the quantitative findings.[2] We visualize the amount of the recommended goals in each intervention and the actual step numbers. Figure 4a shows the recommended step counts for participants who received the peer-reflection in the first week was higher compared to their actual step numbers. Thus, the recommended amounts might be effective to encourage users to increase their activeness. Figure 4b shows the recommended 10,000 step count for participants who got recommended with self-reflection was lower than their actual step numbers. Although the goal in the self-reflection

[2]We illustrate the data only for the sake of visualization. Given the number of participants and the duration of the study, we do not run any statistical test and do not make any claim for the improvement of fitness.

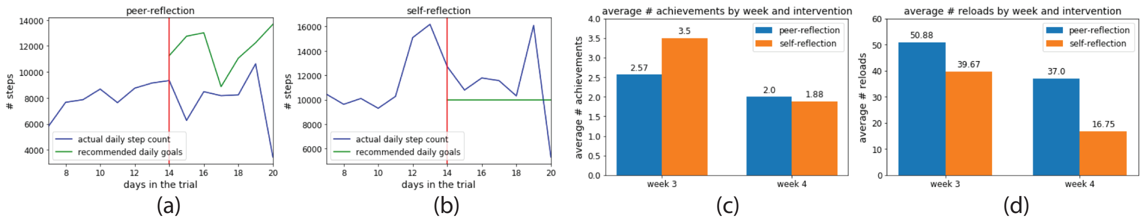

Fig. 4. Actual step counts versus recommended step numbers using peer-reflection (a) and self-reflection (b). User achievements (c) and engagement with the app (d).

group was feasible for these participants it could be changed and set for higher numbers to motivate users to increase their activeness.

Figure 4c demonstrates the average number of days that participants achieved the recommended numbers. We observed higher achievements with the self-reflection than the peer-reflection during the first week. These results might be due to the reason that participants who received self-reflection, in general, were more active users that can easily achieve 10,000 steps per day. Also, during the first week of the recommendation phase participants had higher achievements compared to the second week. Last, we illustrate user engagement using the number of data synchronization between the `Fitbit` server and our host server. The synchronization happened when participants opened the app to see their step numbers. Figure 4d shows that during both weeks participants engaged with peer-reflection much more than self-reflection. This is in line with interviews where participants reported that they often checked the app to interact with their virtual peers. We found that participants engaged with the app more during the first week than the second week which could be due to the novelty effect. Further longitudinal research is required to study the recommendation acceptance and the app abandonment rates over time.

5 CONCLUSION AND FUTURE WORK

We aimed at understanding users' perceptions when receiving fitness recommendations. To this end, we developed a fitness recommender system in the form of a mobile app. We used an ML approach based on probabilistic profiles to build a two-step recommendation system that first recommends the most suitable intervention and then provides users with actionable goals. Our user study showed that participants highly engaged with the app and used it frequently on different occasions. We found a high acceptance rate for recommendations, where participants accepted and were satisfied with the type of intervention they received. Interest in the social aspect, novelty of the concept, and perceived support from their peer were the main reasons motivating users to check their app more often and get motivated to walk in the peer-reflection mode. On the other hand, being self-dependent and physically capable were the two main reasons for preferring self-reflection. Participants expressed their desire to receive more transparent and adaptive recommendations. Besides, those who were physically more active asked for the calorie-based recommendations.

We plan to run a future study to understand the user perception of fitness recommender systems in-the-wild setups and longer durations. This study will seek to understand how such systems support users with behavior change to become more active individuals.

REFERENCES

[1] Rebecca R Andridge and Roderick JA Little. 2010. A review of hot deck imputation for survey non-response. *International statistical review* 78, 1 (2010), 40–64. https://doi.org/10.1111/j.1751-5823.2010.00103.x

[2] Shlomo Berkovsky, Jill Freyne, Mac Coombe, and Dipak Bhandari. 2010. Recommender Algorithms in Activity Motivating Games. In *Proceedings of the Fourth ACM Conference on Recommender Systems* (Barcelona, Spain) *(RecSys '10)*. Association for Computing Machinery, New York, NY, USA, 175–182. https://doi.org/10.1145/1864708.1864742

[3] Jakim Berndsen, Barry Smyth, and Aonghus Lawlor. 2019. Pace My Race: Recommendations for Marathon Running. In *Proceedings of the 13th ACM Conference on Recommender Systems* (Copenhagen, Denmark) *(RecSys '19)*. Association for Computing Machinery, New York, NY, USA, 246–250. https://doi.org/10.1145/3298689.3346991

[4] Jesús Bobadilla, Fernando Ortega, Antonio Hernando, and Abraham Gutiérrez. 2013. Recommender systems survey. *Knowledge-based systems* 46 (2013), 109–132. https://doi.org/10.1016/j.knosys.2013.03.012

[5] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (Jan 2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[6] Priscilla M Clarkson and Monica J Hubal. 2002. Exercise-induced muscle damage in humans. *American journal of physical medicine & rehabilitation* 81, 11 (2002), S52–S69. https://doi.org/10.1097/00002060-200211001-00007

[7] Alan Cooper, Robert Reimann, David Cronin, and Christopher Noessel. 2014. *About face: the essentials of interaction design*. Wiley, Hoboken, NJ, US. https://www.wiley.com/en-es/About+Face:+The+Essentials+of+Interaction+Design,+4th+Edition-p-9781118766576

[8] Expo. 2020. Make any app. Run it everywhere. https://expo.io/. Accessed: 2020-11-30.

[9] Robert G Farrell, Catalina M Danis, Sreeram Ramakrishnan, and Wendy A Kellogg. 2012. Intrapersonal retrospective recommendation: lifestyle change recommendations using stable patterns of personal behavior. In *Proceedings of the First International Workshop on Recommendation Technologies for Lifestyle Change (LIFESTYLE 2012), Dublin, Ireland (in RecSys'12)*. Association for Computing Machinery, New York, NY, USA, 24. https://www.researchgate.net/profile/Yu_Chen66/publication/287905340_CoFeel_Using_emotions_for_social_interaction_in_group_recommender_systems/links/58d4c2c3a6fdcc1bae4d4419/CoFeel-Using-emotions-for-social-interaction-in-group-recommender-systems.pdf#page=24

[10] Hyunho Lee and Youngseok Lee. 2017. A Look at Wearable Abandonment. In *2017 18th IEEE International Conference on Mobile Data Management (MDM)*. IEEE, Piscataway, NJ, USA, 392–393. https://doi.org/10.1109/MDM.2017.70

[11] Yanping Li, An Pan, Dong D Wang, Xiaoran Liu, Klodian Dhana, Oscar H Franco, Stephen Kaptoge, Emanuele Di Angelantonio, Meir Stampfer, Walter C Willett, et al. 2018. Impact of healthy lifestyle factors on life expectancies in the US population. *Circulation* 138, 4 (2018), 345–355. https://doi.org/10.1161/CIRCULATIONAHA.117.032047

[12] Donald W Marquardt and Ronald D Snee. 1975. Ridge regression in practice. *The American Statistician* 29, 1 (1975), 3–20. https://doi.org/10.2307/2683673

[13] Ryo Otsuki, Osamu Sugiyama, Kohei Matsumura, Masahiro Tada, Haruo Noma, and Tomohiro Kuroda. 2017. When Should the Health Promotion Agent Recommend the User to Walk? Approach for Searching an Appropriate Walking Pattern for the User. In *Proceedings of the 5th International Conference on Human Agent Interaction* (Bielefeld, Germany) *(HAI '17)*. Association for Computing Machinery, New York, NY, USA, 469–472. https://doi.org/10.1145/3125739.3132589

[14] Neville Owen, Phillip B Sparling, Geneviève N Healy, David W Dunstan, and Charles E Matthews. 2010. Sedentary behavior: emerging evidence for a new health risk. *Mayo Clin Proc.* 85, 12 (2010), 1138–1141. https://doi.org/10.4065/mcp.2010.0444

[15] Scott A Paluska and Thomas L Schwenk. 2000. Physical activity and mental health. *Sports medicine* 29, 3 (2000), 167–180. https://doi.org/10.2165/00007256-200029030-00003

[16] Pearl Pu, Li Chen, and Rong Hu. 2011. A User-Centric Evaluation Framework for Recommender Systems. In *Proceedings of the Fifth ACM Conference on Recommender Systems* (Chicago, Illinois, USA) *(RecSys '11)*. Association for Computing Machinery, New York, NY, USA, 157–164. https://doi.org/10.1145/2043932.2043962

[17] Akane Sano, Paul Johns, and Mary Czerwinski. 2015. HealthAware: An Advice System for Stress, Sleep, Diet and Exercise. In *Proceedings of the 2015 International Conference on Affective Computing and Intelligent Interaction (ACII) (ACII '15)*. IEEE Computer Society, USA, 546–552. https://doi.org/10.1109/ACII.2015.7344623

[18] WebMD. 2012. Physical Inactivity May Be as Deadly as Smoking. https://www.webmd.com/fitness-exercise/news/20120717/physical-inactivity-may-be-deadly-as-smoking. Accessed: 2020-11-21.

[19] Wennian Yu, Il Yong Kim, and Chris Mechefske. 2021. Analysis of Different RNN Autoencoder Variants for Time Series Classification and Machine Prognostics. *Mechanical Systems and Signal Processing* 149 (2021), 107322. https://doi.org/10.1016/j.ymssp.2020.107322