Noname manuscript No.
(will be inserted by the editor)

Personalized Weight Loss Strategies by Mining Activity Tracker Data

Fabio Gasparetti · Luca Maria Aiello · Daniele Quercia

Received: date / Accepted: date

Abstract Wearable devices make self-monitoring easier by the users, who usually tend to increase physical activity and weight loss maintenance over time. But in terms of behavior adaptation to these goals, these devices do not provide specific features beyond monitoring the achievement of daily goals, such as a number of steps or miles walked, and caloric outtake. The purpose of this study is twofold. By analyzing a large dataset of signals collected by these devices, we identify significant clusters of similar behavior patterns related to user physical activities. We then examine specific patterns of step count in the context of recommendation of habits that more likely give rise to weight loss effects. The evaluation of the effectiveness of these personalized recommendations, based on a comparative study, proves how a recommender system based on the reinforcement learning paradigm is able to guarantee better performance for this task by balancing the trade-off between long-term and short-term rewards.

Keywords Health Recommender System · Human behaviour · Data Mining

1 Introduction

Wearable devices use proprietary algorithms that, by raw sensor signals and information input by the user, can usually estimate steps, distance, calorie burn, and hours of sleep (Majumder et al, 2017; Kamišalić et al, 2018). They are often considered as cost-effective solutions to support weight loss strategies (Bravata et al, 2007).

Indeed, in absence of specific diseases, such as hypothyroidism or dysthymia, or specific genetic factors, weight gain is most likely to occur due to the increasing intake of food, especially of high energy density (Rolls et al, 2005), or alterations of exercise regimens with respect to the one's energy expenditure (Melanson et al, 2013). While the former is primarily caused by unbalanced diets or high consumption of specific food (e.g., packaged snack foods), the latter is explained most of the times by little physical activity. At low levels of physical activity, a significant restriction in the food intake is needed to maintain a correct balance.

F. Gasparetti
Roma Tre University, Rome, Italy
E-mail: gaspare@dia.uniroma3.it

L.M. Aiello
Nokia Bell Labs, Cambridge, United Kingdom

D. Quercia
Nokia Bell Labs, Cambridge, United Kingdom

Fig. 1: Traditional apps associated with activity trackers tend to promote habits that increase daily step count goals (a), whereas health recommender systems may suggest different patterns to improve health status (b).

Over the long run, little physical activity negatively impacts the muscle mass and physical performance (Konopka and Harber, 2014). The muscle mass decline causes a reduction of the metabolic rate, which regulates the physical and chemical processes that happen naturally in our bodies in order to sustain life (Connolly et al, 1999). In other words, as we begin to lose weight, the body responds to the weight loss by reducing the metabolism, thereby slowing down further weight loss. So, an attempt to decrease the food intake will result in our body trying to return the body to its original weight. As the metabolism is slowed down, the caloric intake has to be even more reduced to maintain weight loss. And it often corresponds to situations which have less chances to be pursued for significant long periods, and more likely the lost weight is gained back once the regime is interrupted (Swift et al, 2014; Romieu et al, 2017).

Contrary to what one might think, excessive physical activity may generate *compensatory behaviors* (King et al, 2007), i.e., adjustments we may unconsciously make to regain the calories we just burned. Individuals tend to adapt metabolically to increased physical activity, with the result that the relationship with the energy expenditure is not linear (Pontzer et al, 2016). In other words, larger amount of physical exercise tends to reduce the resting metabolic rate. A deliberative compensatory response may also stimulate the appetite in persons who gets a lot of exercise, increasing their food intake and failing to lose as much weight as expected.

With respect to physical exercise, in order to prevent weight gain the American College of Sports Medicine states that 150-250 minutes per week of moderate intensity activity is effective (Jakicic et al, 2001). Similarly, a recent study suggests a duration of 45 minutes with a frequency of three to five times per week (Chekroud et al, 2018). The World Health Organization (World Health Organization, 2010) suggests to increment the time spent for physical activity to 300 minutes per week with a combination of aerobic and resistance training (e.g., walking and bicycle riding) for additional health benefits. However, a single recommendation for the optimum amount of physical activity for weight loss maintenance is still to be defined (Catenacci et al, 2008).

Human beings constantly adapt their behaviors to reach their goal settings. Activity trackers enable us to track, monitor and store quantifiable outcomes about our habits, especially in terms of physical activities, supporting this behavior adaptation (Oyibo et al, 2018). This kind of ubiquitous sensing is an unprecedented way to study human behaviors, both in the granularity of activities tracked and the length of observation period. The point is to become more aware of how much activity people are doing so that they can make positive changes. Self-monitoring weight and physical activity on a regular basis are two of the keys for long-term success at weight loss (Wing and Phelan, 2005).

This task of collecting and reporting laymen-friendly information, helping to better comprehend the health status as inferred by the tracked activities, is one of the purposes of the *health recommender systems* (Ricci et al, 2010; Wiesner and Pfeifer, 2014). But a simple reporting tool may not be enough to increase physical activity and improve health on the long-term (Leijdekkers and Gay, 2015; Lewis et al, 2015; Bravata et al, 2007; Jakicic et al, 2016a; Finkelstein et al, 2016). The recommendations should include effective and tailored strategies to specifically translate physical activity to health gains (Wiesner and Pfeifer, 2014; Martin et al, 2015; Thomson et al, 2016). In spite of the unquestioning value of these devices, a very few health recommenders make use of the large amount of data collected on daily basis to provide this sort of suggestions (Wiesner and Pfeifer, 2014; Schäfer et al, 2017).

The principal assumption of this study is that multiple recommendation strategies are feasible in the weight loss scenario, and they have to be continuously adapted to recent and future behaviors undertaken by the user. Consumer-grade fitness trackers are key players in the implementation of these strategies. But traditional recommendations provided by fitness apps that monitor the activity by way of these devices, promoting habits that increase daily step counts (see Fig. 1(a)) may not be optimal.

For instance, in the scenario of 150 minute per week recommendations, the physical activity goal can be reached in multiple short bouts, of 15 minutes each, spread throughout the week, or in just two 75-minute close-together sessions of vigorous aerobic activity a week, or any combination of the two. And specific sequences of activity patterns and habits may provide better outcomes than others in specific circumstances, in terms of noticeable weight loss.

Within this setting, we aim to address the following research question:

– **RQ**: Based on patterns of physical habits collected solely from activity trackers, can we provide personalized recommendations, which are able to improve the chances of weight loss success in comparison with traditional strategies that aim at maximizing the count of steps?

To answer this question, we conduct an exploratory and inferential data analysis on a real-life dataset of signals extracted from consumer-grade activity trackers over an extended period. The goal is to investigate the existence of repeated habits in the human behavior and their potential relationship with weight alterations. An instance of a health recommender system based on reinforcement learning evaluates the possibility to exploit these patterns in the context of weight loss. The system relies solely on the data collected from the trackers, without engaging the users in the self-report assessment of their habits.

The paper is organized as follows. Section 2 reports relevant work on health recommender systems based on data extracted from consumer-grade wearable devices. Sections 3 and 4 explore the dataset of signals considered in this study, and provide insights on regular habits of user behavior. The proposed recommender system is introduced in Sect. 5, which is subjected of a comparative evaluation (Sect. 6). Finally, Sect. 8 draws final conclusions and outlines future work.

2 Related Work

In this section, we report on the research activity that covers both large-scale studies on physical habits, and different approaches to develop health recommender systems by exploiting consumer-grade activity trackers' functionality. The former are relevant both for better understanding the positive correlation between physical activities and health benefits when consumer-wearable activity trackers are employed. Additional references are discussed in the second part of the section, which reports the most recent research advancements in individualized exercise recommendations for weight loss.

2.1 Physical Activity and Wearable Devices

Wearable devices have been used to measure the impact of physical activity on relatively small groups of users in clinical studies (Pantelopoulos and Bourbakis, 2008; Jakicic et al, 2016b). Similar to our work to cluster people based on their activity, Fukuoka et al. analyzed accelerometer data of wearable devices worn by 215 women to identify clusters of people exhibiting different activity patterns, with special care about characterizing inactive individuals (Fukuoka et al, 2018).

Reference	Techniques	Input data	Testbed: users; timespan	Performance dimension	Goal
(Berkovsky et al, 2012)	GAME	ACC	180;-	PHY	Increase the physical activity during videogame playing
(Rabbi et al, 2015)	BAND	FOOD, PHY, EXP	16;14w	PHY,FOOD	Recommending physical activity and dietary behaviors
(McDaniel and Anwar, 2017)	-	STR, EXP	-	STR	Recommending stress-relieving activities
(Reimer et al, 2017)	NN	ACC, ECG, GPS	-	PHY	Recognition and prevention of stress conditions
(Yom-Tov et al, 2017)	RL	ACC	27;26w	PHY	Supporting adherence to physical activity goals for diabets patients
(Smyth and Cunningham, 2018)	CBR	PACE	12,968 races,N/A	PHY	Recommending pacing plans for marathon runners
(Vildjiounaite et al, 2018)	HMM	APP, GPS	54;100d	STR	Recognition of stress conditions
(Zhou et al, 2018)	PRE,RL	ACC	64;10w	PHY	Recommending personalized step goals
Proposed approach	RL	ACC, BMI	11,615;1y	W	Recommending weight loss strategies

Legend:
(ACC) Accelerometer data, (APP) Smartphone apps usage, (BAND) Multi-armed bandit online optimization process, (BMI) Body mass index, (CBR) Case-based Reasoning, (CHA) Changes in the user behavior, (ECG) Electrocardiogram data, (EXP) Data explicitly stated by users, (FOOD) Food intake, (GAME) Gamification (Virtual rewards gained for performing real physical activity), (GPS) Global Positioning System data, (HMM) Hidden Markov models, (NN) Neural Networks, (PACE) Sequences of average paces, (PE) Prediction error, (PHY) Amount or quality of physical activity performed (PRE) Predictive modeling, (RL) Reinforcement learning, (STR) Stress level, (W) Weight alteration.

Table 1: Comparison between surveyed health recommender systems.

Pattern analysis investigations of physical activity over larger populations are not easy to be carried out. Datasets consisting of significant amounts of accurate signals related to the user habits sampled by medical equipment, or inferred by self-reported studies, are not easy to build up and, in general, publicly-accessible. By contrast, consumer-level wearable devices, such as activity trackers, allow researchers to easily reveal basic patterns of movements, energy expenditure or vital health readings. Those outcomes may devise better ways to promote individual lifestyles and reduce activity inequality within populations.

Althoff *et al.* (Althoff et al, 2017) evaluated the variations of the activity, in terms of daily steps, between populations of different countries, measured through activity tracker apps installed on smartphones. The authors were also able to study the obesity-activity correlation, proving how a larger number of daily steps is associated with lower levels of obesity. Basically, these observational population-scale studies attempt to accurately infer causalities from longitudinal data, and understand the causes of health outcomes (Althoff, 2017).

Researchers have analyzed traces from personal mobile devices to quantify how physical activity is impacted by exogenous events. One such event was the introduction of a new location-based augmented-reality game “Pokemon Go” (Althoff et al, 2016; Graells-Garrido et al, 2016). The researchers found that Pokemon Go’s introduction resulted in increased activity levels. Another exogenous event that has been studied was that of “walking challenges.” After studying 2,500 walking competitions over a year, Shameli *et al.* concluded that these competitions were associated with an increase of activity, for the average user, by 23% (Shameli et al, 2017).

A recent study proves how consumer-grade wearable activity trackers are not only limited to provide insights into aggregated behavioral and demographic characteristics, but they might be useful in personalized health monitoring (Lim et al, 2018). In particular, the authors show how both resting heart rates and step counts are being associated with cardiovascular and metabolic disease markers.

2.2 Health Recommender Systems and Persuasive Technologies

Personalized, adaptive goal setting allows changing recommended habits over time based on prior individual behavior. It is a fundamental step toward weight loss strategies tailored to the specific users, which constantly adapt to potential alterations of the training program. Instead of traditional strategies that try to maximize the count of steps, personalized recommendations formulate fitness goals accordingly to perform physical activity more effectively.

Smyth and Cunningham propose a system that suggests tailored race-plans for marathon runners (Smyth and Cunningham, 2018). It is based on a Case-Based Reasoning paradigm, which attempts to find a suggestion by reusing and adapting a solution from similar past input instances. The input instances are represented by a sequence of average paces in the marathon, broken down into 5km-segments. The recommendations consist of a best achievable finish-time for the specific personal fitness level, and a suitable marathon race-plan to achieve it. A plan is considered as sequences of average paces for segments. In spite of the specific category of people and the type of recommended actions (i.e., marathon runners and average paces for segment, respectively), unrelated to the weight loss, an evaluation on a large dataset of runners proves the effectiveness in the training program, and the prediction accuracy of finishing times.

With the pervasive purpose of providing low-effort suggestions that should enable actual adoption by a large number of users, Berkovsky *et al.* explored adaptive persuasive technologies in the form of activity-motivating computer games to engage players in physical activity (Berkovsky et al, 2012). The authors propose to associate the rewards and incentives that inspire intrinsic motivation in the player, with physical activity to be performed in the game. A framework for behavioral change support based on a goal hierarchical catalog has been proposed in (Reimer and Maier, 2016). The framework is based on the *nudging* theory (Baron, 2010), which is believed to help people move towards healthier habits. The goal catalog is being personalized to the user behavior. The hierarchical catalog subdivides higher-level long-term goals, such as reducing the Body mass index (BMI), into short-term subgoals, such as "walking 10,000 steps" or "30-minute swimming". Multiple potential subgoals allow the users to achieve the high-level goal by selecting among different strategies. The nudges are used both to praise the good performances, and to make suggestions or reminders otherwise. The authors constantly monitor the user activities that better lead the users towards their goals. A Collaborative filtering approach generates recommendation in terms of sub-goals to pursue first to reach the goal. To the best of our knowledge, recommender systems based on this framework are yet to be developed and evaluated.

MyBehavior app (Rabbi et al, 2015) is one of the first recommendation systems that automatically generates health feedback from physical activity and food log data. It provides personalized suggestions by learning the user's physical activity and dietary behavior. It is based on the Multi-armed bandit online optimization process, which aims at influencing user behaviors by suggesting actions that maximize the calorie loss. The app monitors frequent and infrequent behaviors, considering the former as lower-effort to adopt and, therefore, yielding more chances of calorie loss compared to healthy behaviors assumed in very infrequent routines. The pilot study involved 16 people over a 14-week period shows a significant alterations of behavior habits after few weeks. Similar to the Rimer and Maier's approach, the app requires to know the current user activity mapped to a high-level categorization, with information about the associated metabolic rates (Ainsworth et al, 2011). But consumer-grade activity trackers do not include any inference engine for this task. If few of these activities can be partially inferred by an additional recognition component on raw data from motion sensors (Bulling et al, 2014), the user is asked to self-report them in all other cases. Conversely, the food logging process is to be borne entirely by the user. Besides the burden on the users for explicitly stating their physical and food habits, additional investigations on the accuracy of the self-report measures of physical activity are required (Baranowski, 1988; Sallis and Saelens, 2000).

In the context of prevention of chronic stress, Maier *et al.* (Maier et al, 2014; Reimer et al, 2017) developed a smartphone app for continuous monitoring the user's stress level, giving warnings when it raises up to a certain threshold. The estimation of the stress level is based on samples from an electrocardiography sensor. McDaniel and Anwar's smartphone app (McDaniel and Anwar, 2017), and Vildjiounaite *et al.* algorithm (Vildjiounaite et al, 2018) share similar goals by exploiting the smartphone's sensors. By constantly tracking user activities, these apps can alert people who may be overstressed and trigger interventions. In spite of correlations between stress levels and weight alterations, these apps limit their scope to the identification and notification of high stress levels, without personalized and positively reinforcing interventions tailored to the user previous and current habits.

Reinforcement learning is a very effective tool to provide recommendations that are informed by the past user behaviour, thus effectively triggering behavioural changes. It has proven effective in recommending digital items such as news (Zheng et al, 2018) and music playlists (Liebman et al, 2015) but also in the medical domain in identifying optimal clinical strategies from patients records (Zhao et al, 2009). However, reinforcement learning has been used seldom in the context of recommending activity routines to achieve a target health outcome.

Yom-Tov et al. (Yom-Tov et al, 2017) experimented with a reinforcement learning approach on a small group of patients with diabetes to encourage them to increase their physical activity. Unlike our approach, their system focused on the one dimension—physical activity—with the only goal of increasing it by sending motivational messages. Their reinforcement learning approach was aimed at learning the most effective sequencing of those messages to trigger a patient's positive reaction.

CalFit (Zhou et al, 2018) is a fitness app that automatically sets and notifies users with tailored daily step goals by exploiting a reinforcement learning approach adapted to the context of physical activity interventions. It explicitly models a measure of user's self-efficacy, which represents the capabilities of each person to successfully execute courses of actions. Achieving a goal increases the user's self-efficacy, with a positive feedback to future behaviors and ambitious goals. But too far-reaching goals might negatively affect the willingness to keep following the suggested actions. The evaluation consists of 64 academic staff employees monitored over a 10-week period. The authors prove how tailored goals increase the chance to see persons adhere to the suggested activity levels during the last four weeks. The short-term timespan of the evaluation and the lack of measurements of weight alterations of the students call for additional investigations of the proposed app.

Table 1 summarizes the reviewed approaches. They faced with the task of engaging users in more physical activities, (Berkovsky et al, 2012; Reimer and Maier, 2016), healthier lifestyles (Rabbi et al, 2015), or improving physical performances (Smyth and Cunningham, 2018). However, none of them explicitly consider the weight dimension during the recommendation process. In terms of algorithms, (Smyth and Cunningham, 2018) effectively implement a case-based approach for marathon runners which explicitly represents state-action pairs, which is common to the proposed RL-based approach. MyBehavior uses phone data to provide food and activity suggestions that maximize the chances of achieving calorie loss goals by casting the problem to an optimization process (Rabbi et al, 2015). It relays on logs of food intake and exercises that cannot always be automatically inferred. But the concept of reward is taken into consideration for learning the best strategies, similarly to our approach. Reimer and Maier (2016) combines a hierarchical goal catalog and Collaborative filtering to generate recommendation in terms of goals and sub-goals to pursue. Human activity recognition based on signals extracted from activity trackers or smartphones is usually able to infer very basic activities, such as, walking, running and standing. Since we do not deem this level of activities detailed enough for a exhaustive goal catalog, it has not included in the state and action representation of the proposed RL-based approach. Reinforcement learning has already been considered for the behavior change task by CalFit (Zhou et al, 2018). The representation of the states resembles the one proposed in our approach (that is, number of steps taken in a day), but the goal of the app is to maximize the chances to see the users adhere to the suggested level of physical activity. Hence, a direct correlation to the weight loss cannot be estimated.

In addition to that, the current findings need to be interpreted with caution due to the relatively small number of real-world experiments, limited timespans and number of participants.

3 Dataset of Activities

In order to understand and analyze specific traits of human behavior and significant and repeated habits, a large amount of digital traces is required, unless a laborious and long-lasting exploratory data analysis on many surveyed participants is undertaken.

Our investigation is based on a large real-world dataset of 11,615 users collected by consumer-grade health-monitoring devices manufactured by Nokia over a 1-year time span, namely from Thu 31 March 2016 to Fri 31 March 2017. As for the demographic statistics, most of the users are in the 20-59 range of years, with an almost equal sex ratio of females and males, (43.6% vs 56.4%). The specific age distribution is summarized in Figure 2. The dataset includes samples from two

	ST	SL	BMI
day 1	4510	7.10	35.16
day 2	3784	8.43	-
day 3	4554	6.92	-
day 4	3809	8.65	-
day 5	5378	8.05	36.23
day 6	3910	7.88	-
day 7	4115	9.52	-
...			

Table 2: Few samples of the three considered signals for a random user, where ST stands for number of steps, SL for hours of sleep and BMI for Body mass index.

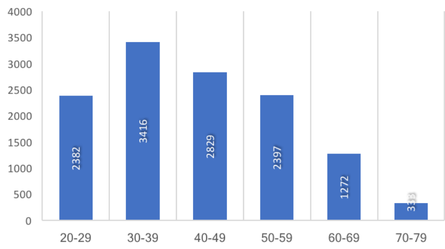

Fig. 2: Age distribution of the users.

categories of devices: wristband activity trackers and digital scales. Few samples of the dataset are reported in Table 2. Since each sample is generated by a specific sensor at a certain timestamp, we can represent the sequence of samples from a sensor as a *signal*.

In the temporal investigation on human behaviors and weight alterations, we require that a significant number of samples per signal occur sequentially. A typical signal that is considered strictly correlated to weight, which also shows less sparsity characteristics (as shown in Figure 3), is the number of steps s_{st}, which is derived by the user's motion adjusted for the height of the person. We suppose that an alteration of the weight is feasible by affecting the behavior patterns related to the physical activity, which has a proxy in the s_{st}-signal sampled by the activity trackers.

Based on the bioelectrical impedance analysis, digital scales send a low electrical current through one foot and reading the current with a sensor under the other foot, and monitor the current flowing through the lean mass, which is the most conductive in the human body. By that measure, digital scales can calculate the BMI metric (also named Quetelet index), the measurement of body fat based on height and weight. It is defined as the ratio between the body mass and the square of the body height and, therefore, is expressed in kg/m^2-unit. It basically quantifies the amount of tissue mass (muscle, fat, and bone) in an individual. It is often considered an inexpensive and easy-to-perform method of screening for weight category. According to the National Heart, Lung, and Blood Institute (NHLBI) guidelines (Lung, and Institute, 2018), the following categories are considered:

- less than 18.5: Underweight range.
- 18.5 to <25: Normal or healthy weight.
- 25.0 to <30: Overweight range.
- 30.0 or higher: Obese range.

The BMI is sampled by the s_{bmi} signal. Digital scales are rarely used at daily basis, therefore the frequency of the samples is lower compared with the other signals.

Finally, the devices included in the dataset are also able to continuously monitor the movements during sleep — also known as actigraphy — and assess sleep-wake cycles. By a combination of accelerometers and heart rate data, they can give user an estimate on how much hours are spent asleep s_{sl}, sampled each day. Accumulated literature proves how short sleep duration is associated with concurrent and future obesity (Patel and Hu, 2008). The American Academy of Sleep Medicine (AASM) and Sleep Research Society (SRS) suggest seven or more hours per night on a regular basis to promote optimal health (Watson et al, 2015). But there is no evidence that specific patterns of hours spent asleep during the week promote weight loss. For this reason, the proposed recommender system limits its scope to the physical activity represented by the number of steps. For the sake of completeness, an investigation of weekly sleep patterns is reported in the following sections.

The accuracy of activity and fitness trackers has been frequently evaluated in the past (Evenson et al, 2015; Ferguson et al, 2015; Straiton et al, 2018). Generally, the studies indicate higher validity for the relative measurement of physical activity measured by the number of steps, with

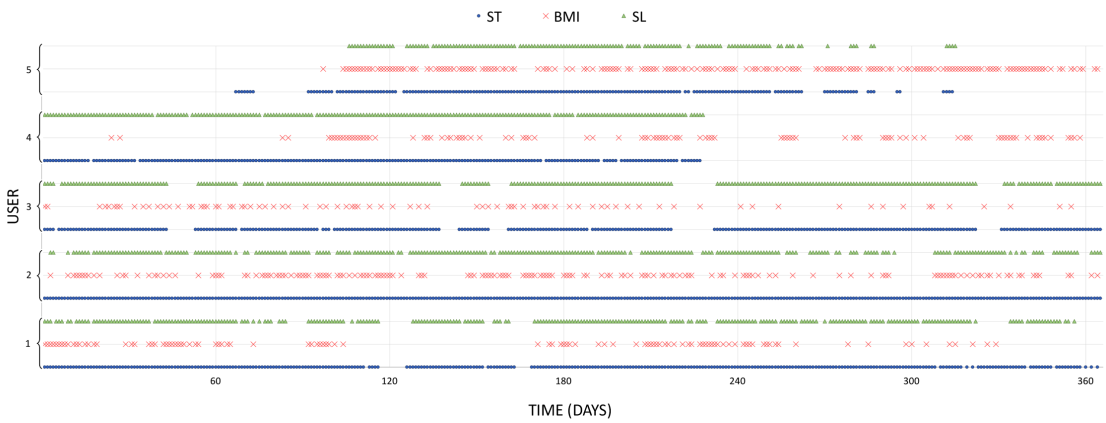

Fig. 3: The sparsity of signals captured by activity trackers for a few random users.

potential risks of steadily undercounting them in specific circumstances, such as in slow, short or non-stereotypical gait patterns (Brodie et al, 2018). As for BMI, many variables might affect the results, including hydration levels, recent exercise activity and underlying medical conditions. But different studies have shown that bioelectrical impedance analysis is a fairly accurate method for estimating body fat (Dehghan and Merchant, 2008; Thibault and Genton, 2014; Demura and Sato, 2015). This study is focused on weight alterations instead of absolute measures. So, we make the hypothesis that, on average, significant long-term alterations of weight identified by tracking body fat changes over time can be recognized by sequences of BMI samples, ignoring minor or single fluctuations of the measure.

As for basic descriptive statistics on the dataset, we have: $\mu_{st} = 6855.20$ ($\sigma_{st} = 3602.22$), $\mu_{sl} = 7.13$ ($\sigma_{sl} = 1.47$), $\mu_{bmi} = 27.07$ ($\sigma_{bmi} = 0.42$). The total number of samples for the ST, BMI, SL signals in the dataset are 2,108,420, 445,444 and 1,353,261, respectively.

Appropriate steps, such as anonymization procedures, have been taken to ensure this research followed fundamental ethical principles, e.g., the subjects cannot be identified or exposed to risks or liabilities throughout the analysis and the evaluation. User locations and further user data were also omitted by the dataset, whenever available, limiting the input data to the above-mentioned signals and user age.

4 Data Analysis

In order to devise recommender systems based on digital records of human behavior obtained by activity trackers, the assumption that these behaviors can be easily observed and matched one another has to be investigated.

The methodology for investigating patterns of physical activity is suitable for signals that consist of samples of daily activities, which last for extended periods (e.g., weeks). The data have therefore identical sampling rate throughout the whole set of users, which corresponds to the amount of physical activity performed during one day.

The sequence of samples indexed in time order is a *discrete signal*. Hence, the dataset consists of discrete-time series $D = \{f^{(s_i)}_{u_j}(t)\}$, where s_i is the i-signal, t is the day of sampling and u_j is the monitored user.

An aspect to be taken into consideration for interpreting physical activity data is that it is typically incomplete (Tang et al, 2018). Wrist-based devices are not steadily worn. Some users are keener to use them when they go to the gym or for a run. Conventional batteries may be fine for

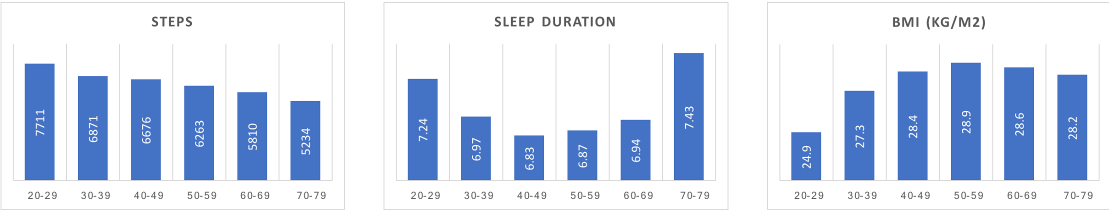

Fig. 4: Macro averages of the considered signals according to age demographic.

sensors and other very low power wearable devices, but they struggle to keep up with the demands of more capable wearables such as fitness bands and smartwatches. For these devices battery life is measured in days. As a result, samples of the signals are often fragmented. Long sequences of uninterrupted samples are less likely to be found. This also means that many temporal and frequency domain analysis techniques that require a significant number of consecutive samples are not suitable in our context. On the other hand, activity trackers may collect signals whose samples may cover also several weeks without interruption. But clustering on the raw time series is often unsatisfactory. Indeed, signals generated from two different users have low chance to show any sort of similarity.

One way to address these issues is to investigate significant weekly patterns, which is a reasonable short time frame if the signals about steps and sleep hours are sampled per day. Indeed, social constraints on daily life, such as working hours or university class schedules, usually impose a frequency and repetition of many of the activities carried out during the week. In the scenarios where the signals are not regularly sampled throughout extended periods, it is still possible to extract partial short-term consecutive values. We should also expect a certain level of regularity in the occurrences of patterns, especially if they all share the same starting day of the week. Formally, we define a *segment* $\hat{f}_{u_j}^{(s_i)}(t, \Delta)$ as a single sequence of samples as follows:

$$\hat{f}_{u_j}^{(s_i)}(t, \Delta) = \left\{ f_{u_j}^{(s_i)}(t') \in D | t \leq t' < t + \Delta \right\} \tag{1}$$

where Δ is a time interval expressed in days. The concatenation of the segments forms the original signal.

Signals monitored by activity trackers are influenced by the age of the user. For instance, physical activity is known to decline with age (Doherty et al, 2017), BMI tends to increase and sleep range shrinks, which is consistent with the demographic statistics from the dataset, as reported in Figure 4. In order to reduce the complexity of the analysis, the identification of the patterns and trends of user activities is performed with no respect to the age of each user. For that reason, the signals are standardized by means of the z-score function:

$$z_{u_j}^{(s_i)}(t) = \frac{f_{u_j}^{(s_i)}(t) - \mu_{u_j}^{(s_i)}}{\sigma_{u_j}^{(s_i)}} \tag{2}$$

where $\mu_{u_j}^{(s_i)}$ and $\sigma_{u_j}^{(s_i)}$ are the sample mean and standard deviation of the signal s_i for the u_j-user, respectively. Similar to Eq. 1, we can derive the standardized segment $\hat{z}_{u_j}^{(s_i)}(t, \Delta)$.

4.1 Clustering

In order to combine multiple similar segments and reduce the total number of different patterns to analyze due to their high variability, a classification is performed on each of the two input signals

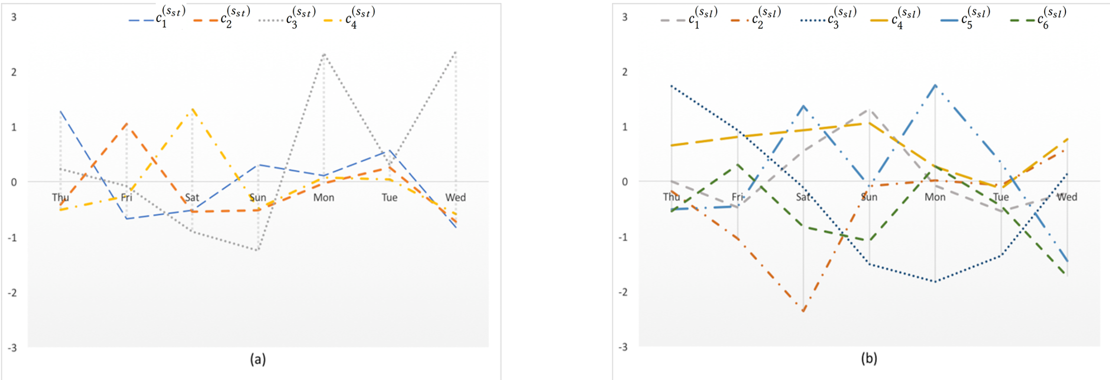

Fig. 5: Seven-day clusters' medoids for signal segments related to step (a) and sleep (b) behaviors. On y-axis the standardized measure (z-score) of the samples.

s_{st} and s_{sl}, that is, directly on their z-score segments. Since we expect slow alterations of BMI, the s_{bmi} samples will be subjected to a more traditional trend analysis.

An unsupervised procedure uses the unlabeled input data to estimate the classification parameters and group the data in predefined K clusters. This kind of partitioning-based clustering requires to specify the number of clusters to be generated in advance. A traditional analysis of the silhouette widths provides us an optimal configuration of numbers of clusters (Mirkin, 2011), namely, four clusters for s_{st} and six clusters for s_{sl}, respectively. The clustering output is a partition of the data so that each segment $\hat{z}_{u_j}^{(s_i)}(t, \Delta)$ belongs to exactly one cluster in:

$$C^{(s_i)} = \{c_1^{(s_i)}, c_2^{(s_i)}, \cdots, c_K^{(s_i)}\} \tag{3}$$

The PAM *Partitioning Around Medoids* (Kaufman and Rousseeuw, 1987; Greenlaw and Kantabutra, 2013) is an adaption of the k-means which shows robust characteristics in the presence of noise. This aspect is very relevant in our scenario, where devices are not steadily worn by the users and daily samples may occasionally show partial representations of the human activity. One more advantage is that the algorithm outputs one *medoid* $\hat{c}_k^{(s_i)}$ per cluster, which is the segment whose average dissimilarity to all other segments in the cluster is minimal. This instance-based representation allows us to better understand and visualize the patterns that each group of segments represents.

The within-class and between-class variabilities, subjected to minimization and maximization, respectively, are formulated as follows:

$$\begin{aligned} I_w &= \sum_{k=1}^{\left|C^{(s_i)}\right|} \sum_{j=1}^{N_u} d^2\left(\hat{z}_{u_j}^{(s_i)}(t, \Delta), \hat{c}_k^{(s_i)}\right), \text{if } \hat{z}_{u_j}^{(s_i)}(t, \Delta) \in c_k^{(s_i)} \\ I_b &= \sum_{k=1}^{\left|C^{(s_i)}\right|} \left|c_k^{(s_i)}\right| d^2\left(\hat{c}_k^{(s_i)}, \bar{c}^{(s_i)}\right) \end{aligned} \tag{4}$$

where d is the distance between one segment and its cluster's medoid, $\bar{c}^{(s_i)}$ is the weighted average of the cluster centers and N_u is the number of users in the dataset. Since the input signals are normalized by the z-score function, we consider a traditional Euclidean distance for the d implementation.

Table 3: Average percentage difference of the medoids with respect to the age of the population.

	20-29	**30-39**	**40-49**	**50-59**	**60-69**	**70-79**	**80-90**
$c_1^{(s_{st})}$	-6.55	-0.88	-0.08	2.63	3.16	8.42	-2.99
$c_2^{(s_{st})}$	4.97	1.07	-1.97	-1.85	-3.12	5.52	2.03
$c_3^{(s_{st})}$	-10.38	-2.27	5.80	-0.97	6.10	-5.36	19.38
$c_4^{(s_{st})}$	11.96	2.08	-3.74	0.19	-6.14	-8.57	-18.42
$c_1^{(s_{sl})}$	-0.85	0.34	12.83	6.62	-25.17	-36.92	-47.90
$c_2^{(s_{sl})}$	14.10	1.64	-2.54	-9.71	0.59	-0.55	6.08
$c_3^{(s_{sl})}$	-3.93	-5.23	-13.12	4.96	25.44	15.58	-12.23
$c_4^{(s_{sl})}$	-3.21	3.28	-0.26	-2.01	-1.65	10.63	21.16
$c_5^{(s_{sl})}$	-3.40	-0.16	4.62	5.46	-7.89	-10.49	-31.27
$c_6^{(s_{sl})}$	-2.71	0.13	-1.52	-5.31	8.67	21.76	64.16

4.1.1 Characterization of the clusters

The medoids for signals s_{st} and s_{sl} obtained by clustering are depicted in Figures 5(a) and 5(b), respectively.

As for the step behaviors, three of the four patterns, namely, $c_1^{(s_{st})}$, $c_2^{(s_{st})}$ and $c_4^{(s_{st})}$, share similar but shifted attitudes during the workweek. In these patterns, the increment of the number of steps happens one day of the week. On the remain days, the steps do not deviate significantly from the average. The segments which show increments on Friday and Saturday are more likely associated with (20-39) age range, while Thursday is preferred by adults in the (40-79) range, as reported in Table 3.

One medoid represents an opposite scenario, when the users do relatively more steps on the beginning of the week, and less on weekends ($c_3^{(s_{st})}$). This last cluster is being characterized by adults > 40 years old. The identified sleep patterns are more diversified. Two patterns, namely $c_2^{(s_{sl})}$ and $c_6^{(s_{sl})}$, represent people that spend considerable less time sleeping during the weekends. If we look at the age distribution, these two clusters better represent two categories of adults, (20-29) and (60-69), respectively (Table 3).

Quite the opposite, medoids $c_1^{(s_{sl})}$ and $c_5^{(s_{sl})}$ are associate to attitudes toward more resting on weekends, which are usually related to users > 40 years.

The cluster analysis allows us to investigate collective behaviors by grouping similar weekly patterns of daily sampled physical activity. Indeed, this exploratory mining proves how significant patterns for the step dimensions can be obtained. But for delivering personalized recommendations of physical activity that take into consideration unique patterns of user behaviors, grouping together several different signals into one single approximate representation (or centroid) may not be the right choice in terms of accuracy of the user profiles. Furthermore, unlabeled instances of user behaviors might be available throughout the monitored period, some with very low frequencies or with different offsets with respect to the first day of the week chosen in the analysis. Section 5.2.1 introduces the specific representation of the user physical activity that is considered during the recommendation process.

4.2 Trend Identification

One additional investigation of the dataset aims at identifying significant upward (downward) trends, i.e., when a sequence of samples of one signal consistently increases (decreases) over a

Table 4: Four-week segment trend analysis: ↑ upward trends, ↓ trends and the rest of the segments (↔) that have not been classified with a trend.

sig	↓-trend	↑-trend	↔-trend
s_{st}	10,493 (6.32%)	10,258 (6.18%)	145,274(87.5%)
s_{sl}	2,674 (4.04%)	2,138 (3.23%)	61,316 (92.64%)
s_{bmi}	1,873 (35.83%)	1,028 (19.66%)	2,326 (44.51%)

period of time in statistical terms, regardless if the trend is linear or not. This investigation does not take into consideration potential correlations between signals.

Mann-Kendall test (Gilbert, 1987) is a non-parametric test for trend analysis which makes no restriction on the distribution of the values. For evaluating the validity of the null hypothesis (i.e., no trend), the statistics are based on the sign of differences, not directly on the values of the samples. For this reason, it is less affected by potential outliers in the raw data. These trends are classified upward or downward for time frames of four weeks (4Δ). For instance, a downward trend for a four-week interval extracted from the user's BMI signal is interpreted as a statistically significance decrease in her BMI in that timespan.

Table 4 reports the number of segments that have been attributed to a upward or downward trend with respect to the overall four-week segments extracted from the dataset. A relatively higher number of downward trends (35.83% vs 19.55%) are discovered on BMI signals, which is expected by users who decide to buy and extensively use activity trackers.

But there is no significant trend difference between the number of four-week segments characterized by a consistent increment of the number of steps or hours spent asleep with respect to segments showing a reduction. In addition, a high percentage of segments are not characterized by any tendency, namely, 87.5% and 92.64% , for s_{st} and s_{sl} respectively. In other words, while there exists a significance variance in these patterns by considering periods of seven days or less, it is less likely that people assume steady upward or downward trends which last one month for the s_{st} and s_{sl} signals.

5 A Health Recommender System

One of the principal roles of health recommender systems is encouraging users to perform behaviors that more likely bear positive effects and conversely discourage them from negative actions. In the weight loss scenario, users tend to select effective exercise strategies that lead to states with greatest value. Typically, this happens by following regular physical routines, which are often associated with the positive effects in terms of BMI reduction. Fitness smartphone apps support users by monitoring their physical activity in terms of steps taken or runs made. The user is able to explore past achievements and achievements still to come to reach personal goals.

Reinforcement Learning (RL) is an area of machine learning whose goal is to develop algorithms that adjust the actions of agents in such a way that positive rewards are being maximized over a long term (Sutton and Barto, 1998). The actions are determined by a policy learned by exploring a state-action space. Therefore, it is reasonable to assume an analogy between an action and the amount of physical activity the user decides to perform in order to maximize the long-term reward measured by BMI reductions.

The proposed health recommender system monitors the recent user physical activity in terms of number of daily steps over a number of days, and suggests a range of steps to be taken in the near future in order to maximize the chances to lose weight. The RL approach is instantiated on the dataset of activities (Sect. 3), and learns the best strategies in terms of steps to be suggested given the recent activity of the user.

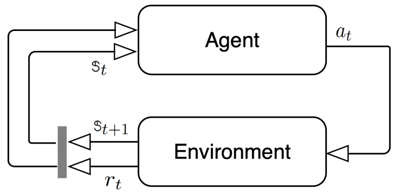

Fig. 6: In Reinforcement Learning, the agent takes actions in the environment and receives an observation, which is typically represented by a reward and a state, that are fed back into the agent.

5.1 Reinforcement Learning

Reinforcement learning is often considered as an effective approach to solve real-world decision-making problems. Figure 6 shows the standard RL setting where the agent interacts within the environment over a number of discrete time steps. At each time step, the agent examines the current state and selects an action from some set of possible actions according to its search strategy, named *policy*. Since in most of the cases training sets consisting of state-action data are not available, the RL makes use of a reward, that is, an estimate return for good and bad behavior as a result of the last taken actions.

A naive approach for ensuring the optimal action is taken at any given time is to simply choose the action which the agent expects to provide the greatest reward, which corresponds to a *greedy* policy. The problem with greedy strategies is related to suboptimal solutions. The benefits of a specific behavior adopted for a short period may show different benefits if it is assumed for long periods. A cleverer recommendation, which is usually implemented in RL strategies, would take into consideration the chain of actions eventually leading to reward, and the potential reward expected in the future. The RL's task is to maximize the long-term cumulative reward.

The interaction of the agent with the environment affects the learning, and vice versa. Specific actions determined by the agent's policy select the future states. The reward obtained from these states determines the future strategy. For this reason, many strategies include a trade-off between exploration, where the agents gather more information that might lead them to better decisions in the future, and exploitation, where they make the best decision given current information. Section 5.2.3 details the approach used in the proposed recommender to balance between the two paradigms.

The RL process is parameterized by a tuple $(\mathcal{S}, \mathcal{A}, \mathcal{T}, r, \gamma)$, where:

- $\mathbb{S}$ is the set of environment states.
- $\mathcal{A}(\mathbb{S})$ is the function that returns the available actions in the $\mathbb{s}$ state.
- The probabilities of transition from one state to another $(\mathcal{T})$.
- The reward function $r(\mathbb{S}) \rightarrow \mathbb{R}$ that estimates the immediate utility of being in a certain state.
- γ is the discount factor, which takes on values in $(0, 1)$ and has the effect of valuing the influence of short-term and long-term rewards. Hence, the rewards received t_n steps in the future are worth less than rewards received now, by a factor of γ^{t_n}.

The goal of RL approaches is to derive a policy π that describes how to act in each state as a result of the acquired experience. The policy takes the form of a function $\pi(\mathbb{S} \times \mathcal{A}) \rightarrow \mathbb{R}$. An optimal policy maximizes the total discounted expected reward, formally $\sum_{t=0}^{\infty} \gamma r(\mathbb{s}_t)$, by determining the probability of taking the a action in the $\mathbb{s}$ state. One popular approach to determine the policy π is Q-learning (Watkins and Dayan, 1992). Intuitively, the Q value is referred to as the state-action value that is iteratively estimated by exploring the state space or, more formally, the expected

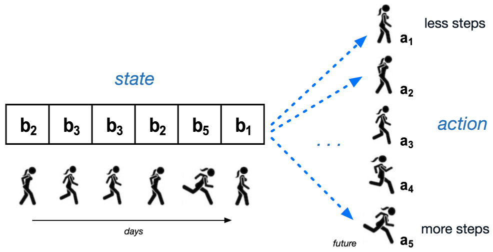

Fig. 7: Given the current state which represents the recent user physical activity in terms of number of steps, we aim at recommending a range of steps to take in the near future.

discounted reward for executing action a at state $\mathbb{s}$ by following the policy π. The goal in Q-learning is to estimate the Q values for an optimal policy.

At the beginning, the Q values are being initialized to an arbitrary constant. Assuming fixed intervals of real time, at each step, given the current state, the algorithm chooses one action and evaluates the state that it has led to. If it has led to an undesirable outcome, the Q value of that action from that state is reduced, so that other actions will have greater chances next time the state is re-evaluated. Similarly, if the reward is positive, it is more likely to choose the same action on that state in the future.

The Bellman equation (Watkins and Dayan, 1992) expresses the relationship between the action value of a state $\mathbb{s}_t$ and the action values of its successor states $\mathbb{s}_{t+1}$, $\mathbb{s}_{t+2}$, etc. The equation approximates the expected future discounted reward. The state-value function can be decomposed into immediate reward, and discounted value of successor state given a certain policy. After the selection of the a_t-action, the environment responds with the reward feedback r_t, by which the mean reward of action a can be estimated as follows:

$$Q(\mathbb{s}_{t+1}, a_t) \leftarrow Q(\mathbb{s}_t, a_t) + \alpha \left[r_t(\mathbb{s}) + \gamma \max_{a' \in \mathcal{A}(\mathbb{s}_{t+1})} Q(\mathbb{s}_{t+1}, a') - Q(\mathbb{s}_t, a_t) \right] \quad (5)$$

where $\alpha \in [0, 1)$ is the learning rate, which determines to what extent newly acquired evidence overrides previously acquired knowledge.

It is important to note that, even for identical environments, the value function changes depending on the policy. This is because the value of the state changes depending on how one acts, since the way she acts in that particular state affects how much reward we expect to see.

5.2 RL-based Health Recommendation

For implementing the RL paradigm to our context, the following definitions are required:

- A finite set of states $\mathbb{S}$ and available actions $\mathcal{A}$ related to the user behaviors tracked by the wearable devices.
- An estimation of the utility of taking specific actions in a certain state.

In our scenario, we are faced repeatedly with a choice among various ranges of volumes of physical activity to perform. The result of this activity provides the user with a reward in terms of health benefits, i.e., weight loss. In the RL paradigm, the amount of physical activity performed by the user corresponds to the action, as depicted in Figure 7. Given the current state which represents the recent user physical activity in terms of number of steps, we aim at recommending actions to maximize the expected total reward over some time period. In the rest of the section states, actions and rewards are formally introduced.

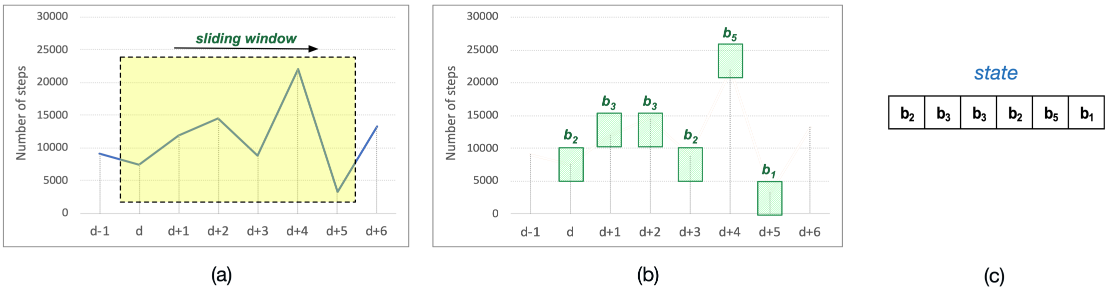

Fig. 8: Binning (b) of the daily sampled signal related to the number of steps (a). The sequence of bins constitutes the RL state (c). The (st) superscript has been omitted for brevity.

Table 5: Binning intervals for one user extracted from the dataset.

$b_1^{(st)}$ and a_1	$b_2^{(st)}$ and a_2	$b_3^{(st)}$ and a_3	$b_4^{(st)}$ and a_4	$b_5^{(st)}$ and a_5
0-3904	3905-5706	5707-7353	7354-9518	9519-20275

$b_1^{(bmi)}$	$b_2^{(bmi)}$	$b_3^{(bmi)}$	$b_4^{(bmi)}$	$b_5^{(bmi)}$
0-26.90	26.91-27.20	27.21-27.48	27.49-27.81	27.82-29.19

5.2.1 Representations of States and Actions

The state at step t refers to whatever information is available to the algorithm at step t about its environment. In order to collect the largest number of behaviors, the two categories of signals s_{st}, and s_{bmi} are subjected of equal-frequency data binning which groups the samples with similar values into one of the five bins $B^{(s_i)} = \{b_1^{(s_i)}, b_2^{(s_i)}, \cdots, b_5^{(s_i)}\}$, where every bin has the same number of samples. The number of bins is determined by a hyperparameter optimization described in Sect. 6. The binning is calculated on a per-user basis and on the z-score representation of the signals (see Sect. 4). Table 5 reports the bins of a specific user extracted from the dataset.

A fixed-sized sliding window of $\Delta^{(s_{st})}$ consecutive days is considered for extracting the available set of states. The window slides across the time series, one day at a time. The states are therefore represented by a bin-representation of the segment $\hat{f}_{u_j}^{(s_{st})}(t, \Delta^{(s_{st})})$. In other words, the original data values are replaced by a fixed value representative of the bin interval they fall in. Figure 8 shows the result of the binning on the original daily sampled signal (b), and the state consisting of the sequence of bin *ids* associated to a certain range of taken steps.

More formally, the RL state is defined as a sequence of $\Delta^{(s_{st})}$ samples, as follows:

$$\mathbb{S} \equiv B^{(st)} \times B^{(st)} \times \cdots \times B^{(st)} \tag{6}$$

The actions that the user can take consist on the number of steps taken the day just after the selected window. Since the step signal of the user is subjected to data binning, it is reasonable to assume that a recommended actions are represented by one of the available bins in $B^{(st)}$. So, each action corresponds to a range of steps that the user may take, that is:

$$\mathcal{A} \equiv B^{(st)} \tag{7}$$

Since we have five bins, A consists of five actions: $\{a_1, a_2, a_3, a_4, a_5\}$, where a_1-action represents the number of steps in the range identified by the b_1 bin, and so on (Table 5 shows this equivalence).

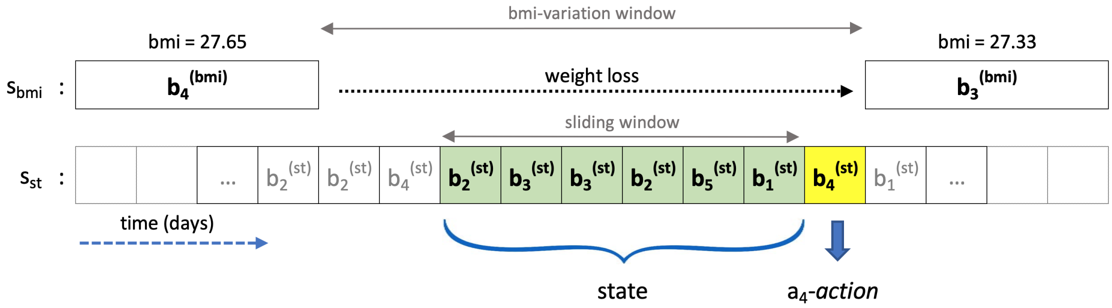

Fig. 9: How states, actions and rewards are determined.

Every day the user is provided with recommendations for the upcoming period which consist of a numerical range of steps to take.

By monitoring the user activity, it is possible to build up a state-action knowledge base, which represents the range of steps taken after each single period identified by a state. Figure 9 depicts the binning representation of the step signal of the user, that is, the state $< b_2^{(st)}, b_3^{(st)}, b_3^{(st)}, b_2^{(st)}, b_5^{(st)}, b_1^{(st)} >$ and the a_4 action determined by the $b_4^{(st)}$-range of steps taken by the user.

5.2.2 Definition of Reward

Rewards are determined by positive or negative alterations of the weight when the user takes specific actions in certain states. BMI signals are usually available occasionally, every time the user decides to use a smart weight scale. Hence, timely rewards following each actions cannot be sampled, and state-action pairs with an associated BMI variation sampled right before and after the state are limited. For this reason, for each pair of consecutive BMI samples, all the state-action pairs between the samples are extracted and associated to the related positive or negative BMI alteration.

In the example depicted in Figure 9, the user generates two BMI samples, 27.65 and 27.33, over a period of 10 days. The sliding window generates several state-action pairs. The two BMI samples are subjected to data binning, so $b_4^{(bmi)}$ and $b_3^{(bmi)}$ are the two corresponding intervals according with Table 5. Since we have a one-bin BMI drop, we suppose that the physical activity performed by the user in that period may have contribute to the reduction. The three state-action pairs falling within the period are extracted and associated to a positive reward.

Hence, each state-action pair extracted within the timespan related to a BMI variation is considered to be relevant for the weight loss goal. It is being assigned to a positive ($r = 1$) or negative reward ($r = -1$) according to the BMI alteration. Less significant gain or loss alterations of the BMI (less than 1%) are not considered. The largest allowed timespan between two BMI samples is set to $\Delta^{(bmi)}$ and is determined by tuning (see Sect. 6).

Once the reward function is defined, the Q-learning algorithm can be executed over the physical activity and BMI measures of a population of users over a period of time. The output consists of a policy represented by a Q-table. For each state, the table indicates the actions with highest cumulative (or longterm) reward, which may considerably differ from immediate rewards. The values are determined by the Bellman equation (Eq. 5). The RL execution usually consists of a number of episodes. In each episode the agent begins the exploration in one random state and performs a sequences of steps toward the terminal state, which represents the goal of the exploration. At each step the Bellman equation is computed and the Q-table updated. Since our setting does not

state						action				
						a_1	a_2	a_3	a_4	a_5
b_2	b_3	b_3	b_2	b_4	b_4	16.0	34.5	-25.5	46.3	**47.1**
b_2	b_3	b_3	b_2	b_4	b_5	-75.0	-27.3	**24.5**	-44.9	-86.4
b_2	b_3	b_3	b_2	b_5	b_1	-16.4	**65.2**	34.0	59.2	-30.8
b_2	b_3	b_3	b_2	b_5	b_2	2.2	7.0	-30.0	22.7	**64.8**
b_2	b_3	b_3	b_2	b_5	b_3	87.1	-63.4	-11.9	**90.1**	-70.1
b_2	b_3	b_3	b_2	b_5	b_4	-61.3	46.4	-41.5	**64.4**	27.7
...						...				

Fig. 10: A partial Q-Table built during the execution of the RL algorithm.

have any terminal state, every episode consists of a fixed number of steps, starting with a random state.

The computed Q-table in Figure 10 suggests that the best action to be taken in the state $< b_2^{(st)}, b_3^{(st)}, b_3^{(st)}, b_2^{(st)}, b_5^{(st)}, b_1^{(st)} >$ is a_2, which represents a number of steps in the 3905-5707 range (according with the user's bins in Table 5).

5.2.3 Exploration vs Exploitation

The ability of the recommender of suggesting behaviors that favorably affect the weight loss depends on the effectiveness of the system to explore and learn patterns of behavior from the available data. A good exploration strategy capitalizes on both the acquisition of new information (exploration) and the maximization of reward by taking advantage of previous knowledge at the same time (exploitation). Q-learning does not specify what the action should actually be taken in each state. A traditional greedy strategy in RL always chooses the best decision a according to the current Q values, that is, $\text{argmax}_a Q(s, a)$. But uncertain environment knowledge prevents from maximizing the long-term reward because suggested actions may not be optimal.

In order to balance opportunity-seeking and advantage-seeking, the approach considers the Boltzmann strategy over Q-values, where the probability of selecting action a in state s is defined as follows:

$$\pi(a|s) = Pr\{a_t = a|\mathbb{s}_t = \mathbb{s}\} = \frac{e^{\frac{q(\mathbb{s},a)}{\tau}}}{\sum_{a' \in \mathcal{A}(\mathbb{s})} e^{\frac{q(\mathbb{s},a')}{\tau}}} \tag{8}$$

where $\tau > 0$ is the temperature specifying how randomly values should be chosen that is repeatedly lowered by a constant factor $\alpha_\tau \in (0, 1)$. When τ is high, the exploration is favored by choosing the actions with a nearly random strategy. As the temperature is reduced, the highest-valued actions are more likely to be chosen (exploitation), with the best actions always chosen when τ is closed to the zero value.

6 Evaluation

The evaluation process is set in a comparative framework by considering various recommendation approaches.

The initial dataset of 11,615 users (Sect. 3) has been subjected to a ten-fold cross-validation. The data consisting of 379,794 state-action pairs is partitioned into equally sized segments or folds, and nine of them are used for training the recommenders, and one is left out for test. In order to evaluate the recommendation accuracy, the reinforcement learning is therefore trained offline on the fixed batch of training data consisting of state-action pairs, with the actual reward obtained by potential BMI alterations.

(R) Our baseline is a standard random policy, where an equal probability is assigned to each available action in A.

(MP) The approach is based on the RL state representation of the human behavior. But instead of exploring different sequences of states that might bring to better conditions, they limit the selection of the next action according to the current state. It recommends the most common action in the dataset given a certain state. Since the dataset consists of users who are supposed to regularly use and monitor their activities to improve their health status, we expect that this strategy reasonably recommends good actions to users. In terms of classification of the right action given an input instance, the MP approach correspond to the Zero Rule algorithm, which relies on the frequency of targets and predicts the majority target category.

(G) It is similar to the MP strategy, but instead of the most common action in the dataset, the selection is based on the frequency of success, that is, the times that the given state-action pair has led to a negative alteration of weight. It is a typical greedy strategy that tries to maximize the reward without acquiring new knowledge.

(AT) A second baseline simulates the kind of recommendation of fitness mobile apps. Many users use mobile apps that collect data from fitness trackers and help them developing exercise routines with achievable, slightly challenging goals based on daily averages. In this scenario, one user shall be deemed to meet the app recommendations if she increases the number of steps per day with respect to the steps taken on average in the previous N_{AT} days. This simulation is implemented in the AT recommender with $N_{AT} = 3$.

(RL) The proposed approach based on RL.

The hyperparameter optimization based on a traditional grid search on a distinct validation dataset determined the tuning of the model, with the optimal configuration reported in Table 6. A six-day pattern analysis of the RL states achieves a slightly better performance in terms of recommendation accuracy compared to the seven-day timespan considered in the clustering analysis.

In our scenario, the usefulness of one recommendation is determined by the achievable cumulative reward in terms of effective weight alterations after having considered the suggested actions. Offline RS evaluations measure the accuracy of recommendations without the involvement of actual users but using existing real-world data as ground truth (Shani and Gunawardana, 2011). Good performances are attributed to RS that match the ground truth, especially when the recommendations are of user interest.

More specifically, the dataset is apportioned into training and test sets. After the initial learning phase on the training set, for each state in the test set, we collect the recommended actions (in terms of range of steps) from each considered approach. The actions are the compared with the action actually taken by the user. State-action instances are also been attributed to BMI positive or negative alterations. The performances depend on how many actions match the ground truth in case of negative alternations of BMI.

We indicate with $W^{(\downarrow,=)}$ and $W^{(\uparrow,=)}$ the number of test state-action pairs that match the recommendations, which are characterized by a weight loss and gain alteration, respectively. Likewise, the metrics $W^{(\downarrow,\neq)}$ and $W^{(\uparrow,\neq)}$ indicates the number of test state-action pairs that do not match the recommendations.

We can cast the evaluation to the traditional set-based measures of precision and accuracy, as follows:

$$Pr = \frac{W^{(\downarrow,=)}}{W^{(\downarrow,=)} + W^{(\downarrow,\neq)}} \tag{9}$$

$$Acc = \frac{W^{(\uparrow,=)} + W^{(\downarrow,=)}}{W^{(\uparrow,=)} + W^{(\downarrow,=)} + W^{(\uparrow,\neq)} + W^{(\downarrow,\neq)}} \tag{10}$$

Table 6: Quantitative parameters and references.

Number of bins = 5	: Sect. 5.2.1
$\alpha = 0.30$	: Eq. 5 (learning rate)
$\gamma = 0.90$	: Eq. 5 (discount factor)
$\Delta^{(s_{st})} = 6$	: Eq. 6 (sliding window size)
$\Delta^{bmi} = 11$	: Sect. 5.2.2 (largest timespan for BMI variations)
1,000 episodes	: Section 5.1 (RL training episodes)

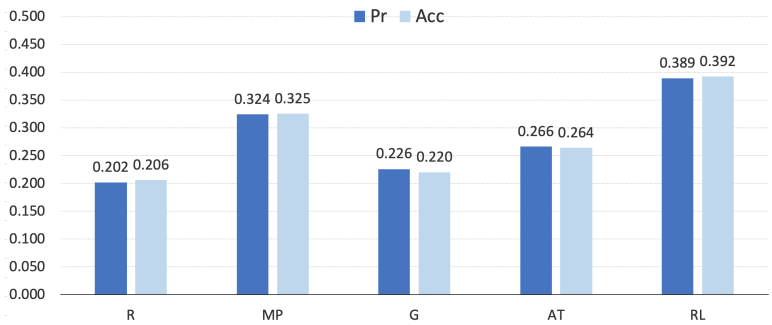

Fig. 11: Precision and accuracy measures of the recommendation on the test set.

In other words, when the user undertakes an action that proves to be beneficial for weight loss, but the recommender system fails to match it, the precision is negatively affected. The accuracy metric seeks to measure how much the recommender is able to match the actions that influence the user BMI on overall. If it is less significant for the weight-loss purpose, it gives us an additional dimension to the performance of the algorithms in terms of prediction of good and bad actions. If the recommender is able to identify actions that have the chance to increase the weight, they might however be notified to the user, who can proactively alter her behavior accordingly to avoid them. A recent study proves the associations between mortality risk and BMI levels below 18.5 kg/m^2 (Bhaskaran et al, 2018), so recommendations for gaining weight cannot be totally ignored.

For statistical validation, Wilcoxon signed-ranks test on each pair of classifiers at confidence level of 0.01 has been computed to reject the null-hypothesis (i.e., two algorithms perform equally well).

Figure 11 reports the outcomes of the experimental evaluation. Obvious results are to be found in the baseline (R): the random approach chooses one of the five available actions, gaining 20% of success on average.

In terms of percentage difference of precision averaged over the 10 folds, the proposed RL-based recommender, which selects candidate actions by analyzing the recent activity patterns, reaches the highest increments, namely: 20.06%, 72.12% and 46.24%, in comparison with MP, G and AT approaches. Increments of accuracy show comparable increments.

But a significant observation is related to the recommendations provided by traditional apps associated with activity trackers (AT), which follow a locally optimal choice, similar to a greedy strategy, therefore they often obtain suboptimal outcomes in terms of weight loss. Whenever the user decides to adhere to that strategy, trying to steadily increment the count of steps every day, the chances of success decrease.

A sort of wisdom of the crowd is manifested when the recommendations are based on the most common action given the current state (MP). Regardless the past behavior of users, the most-popular strategy suggests valid routines toward weight loss one-third of the times, with higher accuracy with respect to purely greedy strategies.

These performance estimates on the considered real world dataset confirm that weight loss personalized strategies can be devised by taking into consideration behavior patterns related to the physical activity, which has a proxy in the count of steps sampled by the consumer-grade activity trackers.

7 Limitations

Our work has some limitations which we summarized below.

Limited explanatory power. Our study has only a partial view over the multitude of factors that impact health outcomes and, specifically, loss or gain of body weight. While physical activity is hugely important for overall health and controlling weight, other factors including food consumption, medications, and overall physical and mental health conditions, play a major role. Important confounders are not limited to the personal sphere of the individuals under study but also pertain the context around them. Ideally, social influence and environmental factors (e.g., weather, access to amenities, exogenous events) should be considered. Gathering data that is both high dimensional and large-scale is challenging, not least because of the privacy issues that collecting very granular data entails. Our work makes a first step in that direction by considering three dimensions (sleep, activity, and weight) that are rarely studied jointly, and does so on a rather large population. In the future, additional studies are needed to bring important determinants of body weight into the equation. As additional environmental or human dimensions become available through wearable sensors or self-reports, the state spaces can be extended with more configurations.

Incomplete data. The restricted set of signals used in our model is also a limitation, mainly caused by the nature of the data collection process. The information gathered from commercial activity trackers deployed “in the wild” is naturally noisy and incomplete: the signal is rarely observable for long and uninterrupted periods of time. To partially tackle this intrinsic limit, we proposed to investigate significant weekly patterns instead of focusing on longer periods of time. Our experimental results show that this approach is effective, yet it misses important temporal signals such as seasonality and other long-term patterns of the temporal traces.

Representativeness. Analyzing the behaviour of the customer base of a consumer product introduces also a problem of representativeness. The cohort of owners of wearable and tracking devices is a self-selected sample of people which might not be representative of the global population. Users of popular wearables and tracking apps tend to be biased toward young, more affluent, gender-skewed populations.

Causality. Our study is purely observational. Our method relies on longitudinal data only and it is evaluated in an off-line fashion, with no actual recommendation given to real users. This is the most common setup adopted by most studies in the recommender system domain, primarily because it is often not possible to get direct access to the user base to deliver interventions and to monitor their effects over time. As a result, our framework has a good predictive power—it is able to anticipate which sequences of actions lead to the desired outcome—but it cannot speak to the underlying causal processes that lead to weight loss. To partly address this shortcoming, one could experiment with statistical frameworks for causal modeling and possibly corroborate the findings from the large-scale study with small-scale experiments conducted with volunteers.

8 Conclusions

Contemporary technology provides us an unprecedented opportunity for the use of consumer-grade activity trackers to both collect wealthy of data and investigate the many aspects of the user behavior. An exploratory cluster analysis revealed some significant patterns in terms of physical activity and type of sleep. The outcomes were achieved from a real-life dataset of signals collected over an extended period of time (1 year). A novel health recommender system based on the Reinforcement learning paradigm has been introduced to prove that specific patterns of physical activity, in terms of step count, are also useful in the weight loss scenario. Whereas Reinforcement learning, Case-based reasoning and optimization techniques have already be considered for improving people's personal habits and attitudes, to the best of our knowledge none of these paradigms have been evaluated in a real-world scenario related to weight loss. Potential benefits can be obtained if the proposed RL-based approach is implemented in fitness apps, for providing less challenging routines which, however, show effects on weight loss maintenance over time.

The suggestions provided by the recommender are tailored to the particular behavior of the user, but they do not take into consideration individual characteristics, such as the exercise capacity, specific health risks, current medical conditions or, longer simply, the age.

Future work shall consider stereotype-based user profiling to identify different clusters of people with similar characteristics. For each cluster, a recommendation sub-strategy is being conceived by the subset of data that better characterize the users in the cluster. For instance, if the population is categorized by age into adolescents, adults and elders; we can expect different physical activity behaviors between groups in terms of frequency, intensity and continuity. By separately training the RL approach on each subset of the population, we can expect more tailored strategies based on the singular characteristics of individuals.

References

Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, Greer JL, Vezina J, Whitt-Glover MC, Leon AS (2011) 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc 43(8):1575–1581

Althoff T (2017) Population-scale pervasive health. IEEE Pervasive Comput 16(4):75–79, DOI 10.1109/MPRV.2017.3971134, URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5951162/

Althoff T, White RW, Horvitz E (2016) Influence of pokémon go on physical activity: study and implications. Journal of medical Internet research 18(12)

Althoff T, Sosic R, Hicks JL, King AC, Delp SL, Leskovec J (2017) Large-scale physical activity data reveal worldwide activity inequality. Nature 547:336 EP –, URL http://dx.doi.org/10.1038/nature23018

Baranowski T (1988) Validity and reliability of self report measures of physical activity: An information-processing perspective. Research Quarterly for Exercise and Sport 59(4):314–327, DOI 10.1080/02701367.1988.10609379, URL https://doi.org/10.1080/02701367.1988.10609379, https://doi.org/10.1080/02701367.1988.10609379

Baron J (2010) Nudge: Improving decisions about health, wealth, and happiness. thaler, richard h. and sunstein, cass r. (eds.), yale university press: New haven, ct, 2008). Journal of Behavioral Decision Making 23(2):224–226

Berkovsky S, Freyne J, Coombe M (2012) Physical activity motivating games: Be active and get your own reward. ACM Trans Comput-Hum Interact 19(4):32:1–32:41

Bhaskaran K, dos Santos-Silva I, Leon DA, Douglas IJ, Smeeth L (2018) Association of bmi with overall and cause-specific mortality: a population-based cohort study of 3.6 million adults in the uk. The Lancet Diabetes & Endocrinology 6(12):944–953, DOI 10.1016/S2213-8587(18)30288-2, URL https://doi.org/10.1016/S2213-8587(18)30288-2

Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, Stave CD, Olkin I, Sirard JR (2007) Using pedometers to increase physical activity and improve health: a systematic review. JAMA 298(19):2296–2304

Brodie M, Pliner E, Ho A, Li K, Chen Z, Gandevia S, Lord S (2018) Big data vs accurate data in health research: Large-scale physical activity monitoring, smartphones, wearable devices and risk of unconscious bias. Medical Hypotheses 119:32 – 36, DOI https://doi.org/10.1016/j.mehy.2018.07.015, URL http://www.sciencedirect.com/science/article/pii/S0306987718301750

Bulling A, Blanke U, Schiele B (2014) A tutorial on human activity recognition using body-worn inertial sensors. ACM Comput Surv 46(3):33:1–33:33, DOI 10.1145/2499621, URL http://doi.acm.org/10.1145/2499621

Catenacci VA, Ogden LG, Stuht J, Phelan S, Wing RR, Hill JO, Wyatt HR (2008) Physical activity patterns in the national weight control registry. Obesity 16(1):153–161

Chekroud SR, Gueorguieva R, Zheutlin AB, Paulus M, Krumholz HM, Krystal JH, Chekroud AM (2018) Association between physical exercise and mental health in 1·2 million individuals in the usa between 2011 and 2015: a cross-sectional study. The Lancet Psychiatry DOI https://doi.org/10.1016/S2215-0366(18)30227-X, URL http://www.sciencedirect.com/science/article/pii/S221503661830227X

Connolly J, Romano T, Patruno M (1999) Effects of dieting and exercise on resting metabolic rate and implications for weight management. Family Practice 16(2):196–201, DOI 10.1093/fampra/16.2.196, URL http://dx.doi.org/10.1093/fampra/16.2.196, /oup/backfile/content_public/journal/fampra/16/2/10.1093_fampra_16.2.196/1/160196.pdf

Dehghan M, Merchant AT (2008) Is bioelectrical impedance accurate for use in large epidemiological studies? Nutr J 7:26

Demura S, Sato S (2015) Comparisons of accuracy of estimating percent body fat by four bioelectrical impedance devices with different frequency and induction system of electrical current. The Journal of sports medicine and physical fitness 55(1-2):68—75, URL http://europepmc.org/abstract/MED/25642685

Doherty A, Jackson D, Hammerla N, Plötz T, Olivier P, Granat MH, White T, van Hees VT, Trenell MI, Owen CG, Preece SJ, Gillions R, Sheard S, Peakman T, Brage S, Wareham NJ (2017) Large scale population assessment of physical activity using wrist worn accelerometers: The uk biobank study. PLOS ONE 12(2):1–14, DOI 10.1371/journal.pone.0169649, URL https://doi.org/10.1371/journal.pone.0169649

Evenson KR, Goto MM, Furberg RD (2015) Systematic review of the validity and reliability of consumer-wearable activity trackers. International Journal of Behavioral Nutrition and Physical Activity 12(1):159, URL https://doi.org/10.1186/s12966-015-0314-1

Ferguson T, Rowlands AV, Olds T, Maher C (2015) The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. Int J Behav Nutr Phys Act 12:42

Finkelstein EA, Haaland BA, Bilger M, Sahasranaman A, Sloan RA, Nang EEK, Evenson KR (2016) Effectiveness of activity trackers with and without incentives to increase physical activity (trippa): a randomised controlled trial. The Lancet Diabetes and Endocrinology 4(12):983 – 995, DOI https://doi.org/10.1016/S2213-8587(16)30284-4, URL http://www.sciencedirect.com/science/article/pii/S2213858716302844

Fukuoka Y, Zhou M, Vittinghoff E, Haskell W, Goldberg K, Aswani A (2018) Objectively measured baseline physical activity patterns in women in the mped trial: Cluster analysis. JMIR public health and surveillance 4(1)

Gilbert RO (1987) Statistical methods for environmental pollution monitoring. New York : Van Nostrand Reinhold Co, includes index

Graells-Garrido E, Ferres L, Bravo L (2016) The effect of pokémon go on the pulse of the city: A natural experiment. arXiv preprint arXiv:161008098

Greenlaw R, Kantabutra S (2013) Survey of clustering: Algorithms and applications. Int J Inf Retr Res 3(2):1–29

Jakicic J, Clark K, Coleman E, Donnelly J, Foreyt J, Melanson E, Volek J, Volpe S, American College of Sports Medicine (2001) American college of sports medicine position stand. appropriate intervention strategies for weight loss and prevention of weight regain for adults. Medicine and science in sports and exercise 33(12):2145–2156

Jakicic JM, Davis KK, Rogers RJ, King WC, Marcus MD, Helsel D, Rickman AD, Wahed AS, Belle SH (2016a) Effect of Wearable Technology Combined With a Lifestyle Intervention on Long-term Weight Loss: The IDEA Randomized Clinical Trial. JAMA 316(11):1161–1171

Jakicic JM, Davis KK, Rogers RJ, King WC, Marcus MD, Helsel D, Rickman AD, Wahed AS, Belle SH (2016b) Effect of wearable technology combined with a lifestyle intervention on long-term weight loss: the idea randomized clinical trial. Jama 316(11):1161–1171

Kamišalić A, Fister I, Turkanović M, Karakatič S (2018) Sensors and functionalities of non-invasive wrist-wearable devices: A review. Sensors 18(6), DOI 10.3390/s18061714, URL http://www.mdpi.com/1424-8220/18/6/1714

Kaufman L, Rousseeuw P (1987) Clustering by Means of Medoids. Delft University of Technology : reports of the Faculty of Technical Mathematics and Informatics, Faculty of Mathematics and Informatics

King NA, Caudwell P, Hopkins M, Byrne NM, Colley R, Hills AP, Stubbs JR, Blundell JE (2007) Metabolic and behavioral compensatory responses to exercise interventions: Barriers to weight loss. Obesity 15(6):1373–1383, DOI 10.1038/oby.2007.164, URL https://onlinelibrary.wiley.com/doi/abs/10.1038/oby.2007.164, https://onlinelibrary.wiley.com/doi/pdf/10.1038/oby.2007.164

Konopka AR, Harber MP (2014) Skeletal muscle hypertrophy after aerobic exercise training. Exerc Sport Sci Rev 42(2):53–61

Leijdekkers P, Gay V (2015) Improving user engagement by aggregating and analysing health and fitness data on a mobile app. In: Geissbühler A, Demongeot J, Mokhtari M, Abdulrazak B, Aloulou H (eds) Inclusive Smart Cities and e-Health, Springer International Publishing, Cham, pp 325–330

Lewis ZH, Lyons EJ, Jarvis JM, Baillargeon J (2015) Using an electronic activity monitor system as an intervention modality: A systematic review. BMC Public Health 15:585

Liebman E, Saar-Tsechansky M, Stone P (2015) Dj-mc: A reinforcement-learning agent for music playlist recommendation. In: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, International Foundation for Autonomous Agents and Multiagent Systems, pp 591–599

Lim WK, Davila S, Teo JX, Yang C, Pua CJ, Blöcker C, Lim JQ, Ching J, Yap JJL, Tan SY, Sahlén A, Chin CWL, Teh BT, Rozen SG, Cook SA, Yeo KK, Tan P (2018) Beyond fitness tracking: The use of consumer-grade wearable data from normal volunteers in cardiovascular and lipidomics research. PLOS Biology 16(2):1–18, DOI 10.1371/journal.pbio.2004285, URL https://doi.org/10.1371/journal.pbio.2004285

Lung, NH, Institute B (2018) Classification of overweight and obesity by bmi, waist circumference, and associated disease risks. https://www.nhlbi.nih.gov/health/educational/lose_wt/BMI/bmi_dis.htm, online; accessed May 1, 2018

Maier E, Reimer U, Laurenzi E, Ridinger M, Ulmer T (2014) Smartcoping - A mobile solution for stress recognition and prevention. In: Bienkiewicz M, Verdier C, Plantier G, Schultz T, Fred ALN, Gamboa H (eds) HEALTHINF 2014 - Proceedings of the International Conference on Health Informatics, ESEO, Angers, Loire Valley, France, 3-6 March, 2014, SciTePress, pp 428–433

Majumder S, Mondal T, Deen MJ (2017) Wearable Sensors for Remote Health Monitoring. Sensors (Basel) 17(1)

Martin SS, Feldman DI, Blumenthal RS, Jones SR, Post WS, McKibben RA, Michos ED, Ndumele CE, Ratchford EV, Coresh J, Blaha MJ (2015) mActive: A Randomized Clinical Trial of an Automated mHealth Intervention for Physical Activity Promotion. J Am Heart Assoc 4(11)

McDaniel M, Anwar M (2017) Zen_space: A smartphone app for individually tailored stress management support for college students. In: Chen H, Zeng DD, Karahanna E, Bardhan I (eds) Smart Health: International Conference, ICSH 2017, Hong Kong, China, June 26-27, 2017, Proceedings, Springer International Publishing, Cham, pp 123–133

Melanson EL, Keadle SK, Donnelly JE, Braun B, King NA (2013) Resistance to exercise-induced weight loss: compensatory behavioral adaptations. Med Sci Sports Exerc 45(8):1600–1609

Mirkin B (2011) Choosing the number of clusters. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 1(3):252–260

Oyibo K, Adaji I, Orji R, Olabenjo B, Azizi M, Vassileva J (2018) Perceived persuasive effect of behavior model design in fitness apps. In: Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization, ACM, New York, NY, USA, UMAP '18, pp 219–228, DOI 10.1145/3209219.3209240, URL http://doi.acm.org/10.1145/3209219.3209240

Pantelopoulos A, Bourbakis N (2008) A survey on wearable biosensor systems for health monitoring. In: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, pp 4887–4890

Patel SR, Hu FB (2008) Short sleep duration and weight gain: a systematic review. Obesity (Silver Spring) 16(3):643–653

Pontzer H, Durazo-Arvizu R, Dugas LR, Plange-Rhule J, Bovet P, Forrester TE, Lambert EV, Cooper RS, Schoeller DA, Luke A (2016) Constrained Total Energy Expenditure and Metabolic Adaptation to Physical Activity in Adult Humans. Curr Biol 26(3):410–417

Rabbi M, Aung MH, Zhang M, Choudhury T (2015) Mybehavior: Automatic personalized health feedback from user behaviors and preferences using smartphones. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, ACM, New York, NY, USA, UbiComp '15, pp 707–718, DOI 10.1145/2750858.2805840, URL http://doi.acm.org/10.1145/2750858.2805840

Reimer U, Maier E (2016) An application framework for personalised and adaptive behavioural change support systems. In: Röcker C, Ziefle M, O'Donoghue J, Maciaszek LA, Molloy W (eds) Proceedings of the 2nd International Conference on Information and Communication Technologies for Ageing Well and e-Health, ICT4AgeingWell 2016, Rome, Italy, April 21-22, 2016, SCITEPRESS, pp 152–159

Reimer U, Laurenzi E, Maier E, Ulmer T (2017) Mobile stress recognition and relaxation support with smartcoping: User-adaptive interpretation of physiological stress parameters. In: 50th Hawaii International Conference on System Sciences, HICSS 2017, Hilton Waikoloa Village, Hawaii, USA, January 4-7, 2017, AIS Electronic Library (AISeL)

Ricci F, Rokach L, Shapira B, Kantor PB (2010) Recommender Systems Handbook, 1st edn. Springer-Verlag, Berlin, Heidelberg

Rolls BJ, Drewnowski A, Ledikwe JH (2005) Changing the energy density of the diet as a strategy for weight management. J Am Diet Assoc 105(5 Suppl 1):98–103

Romieu I, Dossus L, Barquera S, Blottière HM, Franks PW, Gunter M, Hwalla N, Hursting SD, Leitzmann M, Margetts B, Nishida C, Potischman N, Seidell J, Stepien M, Wang Y, Westerterp K, Winichagoon P, Wiseman M, Willett WC (2017) Energy balance and obesity: what are the main drivers? Cancer Causes & Control 28(3):247–258

Sallis JF, Saelens BE (2000) Assessment of physical activity by self-report: status, limitations, and future directions. Res Q Exerc Sport 71(2 Suppl):1–14

Schäfer H, Hors-Fraile S, Karumur RP, Calero Valdez A, Said A, Torkamaan H, Ulmer T, Trattner C (2017) Towards health (aware) recommender systems. In: Proceedings of the 2017 International Conference on Digital Health, ACM, New York, NY, USA, DH '17, pp 157–161

Shameli A, Althoff T, Saberi A, Leskovec J (2017) How gamification affects physical activity: Large-scale analysis of walking challenges in a mobile application. In: Proceedings of the 26th International Conference on World Wide Web Companion (WWW), pp 455–463

Shani G, Gunawardana A (2011) Evaluating Recommendation Systems, Springer US, Boston, MA, pp 257–297. DOI 10.1007/978-0-387-85820-3_8, URL https://doi.org/10.1007/978-0-387-85820-3_8

Smyth B, Cunningham P (2018) Marathon race planning: A case-based reasoning approach. In: Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18, International Joint Conferences on Artificial Intelligence Organization, pp 5364–5368, DOI 10.24963/ijcai.2018/754, URL https://doi.org/10.24963/ijcai.2018/754

Straiton N, Alharbi M, Bauman A, Neubeck L, Gullick J, Bhindi R, Gallagher R (2018) The validity and reliability of consumer-grade activity trackers in older, community-dwelling adults: A systematic review. Maturitas 112:85 – 93, DOI https://doi.org/10.1016/j.maturitas.2018.03.016, URL http://www.sciencedirect.com/science/article/pii/S0378512218301828

Sutton RS, Barto AG (1998) Introduction to Reinforcement Learning, 1st edn. MIT Press, Cambridge, MA, USA

Swift DL, Johannsen NM, Lavie CJ, Earnest CP, Church TS (2014) The role of exercise and physical activity in weight loss and maintenance. Progress in Cardiovascular Diseases 56(4):441–448

Tang LM, Meyer J, Epstein DA, Bragg K, Engelen L, Bauman A, Kay J (2018) Defining adherence: Making sense of physical activity tracker data. Proc ACM Interact Mob Wearable Ubiquitous Technol 2(1):37:1–37:22, DOI 10.1145/3191769, URL http://doi.acm.org/10.1145/3191769

Thibault R, Genton L (2014) Accuracy of bioelectrical impedance analysis to measure skeletal muscle mass. Clinical Nutrition 33(6):1157, DOI 10.1016/j.clnu.2014.08.004, URL http://dx.doi.org/10.1016/j.clnu.2014.08.004

Thomson C, Nash J, Maeder A (2016) Persuasive design for behaviour change apps: Issues for designers. In: Proceedings of the Annual Conference of the South African Institute of Computer Scientists and Information Technologists, ACM, New York, NY, USA, SAICSIT '16, pp 43:1–43:10, DOI 10.1145/2987491.2987535, URL http://doi.acm.org/10.1145/2987491.2987535

Vildjiounaite E, Kallio J, Kyllönen V, Nieminen M, Määttänen I, Lindholm M, Mäntyjärvi J, Gimel'farb G (2018) Unobtrusive stress detection on the basis of smartphone usage data. Personal and Ubiquitous Computing 22(4):671–688, DOI 10.1007/s00779-017-1108-z, URL https://doi.org/10.1007/s00779-017-1108-z

Watkins CJCH, Dayan P (1992) Q-learning. Machine Learning 8(3):279–292

Watson NF, Badr MS, Belenky G, Bliwise DL, Buxton OM, Buysse D, Dinges DF, Gangwisch J, Grandner MA, Kushida C, Malhotra RK, Martin JL, Patel SR, Quan SF, Tasali E (2015) Joint Consensus Statement of the American Academy of Sleep Medicine and Sleep Research Society on the Recommended Amount of Sleep for a Healthy Adult: Methodology and Discussion. J Clin Sleep Med 11(8):931–952

Wiesner M, Pfeifer D (2014) Health recommender systems: Concepts, requirements, technical basics and challenges. International Journal of Environmental Research and Public Health 11(3):2580–2607

Wing RR, Phelan S (2005) Long-term weight loss maintenance. The American Journal of Clinical Nutrition 82(1):222S–225S, DOI 10.1093/ajcn/82.1.222S, URL http://dx.doi.org/10.1093/ajcn/82.1.222S

World Health Organization (2010) Global Recommendations on Physical Activity for Health. World Health Organization

Yom-Tov E, Feraru G, Kozdoba M, Mannor S, Tennenholtz M, Hochberg I (2017) Encouraging physical activity in patients with diabetes: intervention using a reinforcement learning system. Journal of medical Internet research 19(10)

Zhao Y, Kosorok MR, Zeng D (2009) Reinforcement learning design for cancer clinical trials. Statistics in medicine 28(26):3294–3315, DOI 10.1002/sim.3720, URL http://europepmc.org/articles/PMC2767418

Zheng G, Zhang F, Zheng Z, Xiang Y, Yuan NJ, Xie X, Li Z (2018) Drn: A deep reinforcement learning framework for news recommendation. In: Proceedings of the 2018 World Wide Web Conference on World Wide Web, International World Wide Web Conferences Steering Committee, pp 167–176

Zhou M, Fukuoka Y, Mintz Y, Goldberg K, Kaminsky P, Flowers E, Aswani A (2018) Evaluating machine learning-based automated personalized daily step goals delivered through a mobile phone app: Randomized controlled trial. JMIR Mhealth Uhealth 6(1):e28, DOI 10.2196/mhealth.9117, URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5806006/

Fabio Gasparetti is Assistant Professor at the Artificial Intelligence Laboratory of the Department of Engineering, Roma Tre University. He received his PhD in 2005 doing research on Adaptive Web Search. He is member since 2017 of the A.I. Task Force promoted by AgID (Agency for Digital Italy) whose goal is enhancing the quality of the public administration services offered to citizens by the recent AI advances. He co-authored 60+ peer-reviewed papers. He has collaborated and is still collaborating in National and International research projects regarding Internet technologies. His current research interests include user modeling and user adapted interaction; recommender systems and context-awareness, social network analysis and intelligent tutoring systems.

Luca Maria Aiello is a Senior Research Scientist in the Social Dynamics group at Bell Labs in Cambridge, UK and Resarch Fellow of the ISI Foundation in Torino, Italy. He received his PhD in Computer Science from the University of Torino in 2012 and he has been a Research Scientist at Yahoo Labs for almost 5 years. He conducts interdisciplinary research in network science, computational social science, and urban informatics. He co-authored 60+ peer-reviewed papers and his work has been covered by hundreds of news articles published by prestigious news outlets worldwide including Wired, Wall Street Journal, and BBC. He was General Chair of SocInfo in 2014 and covered different organizational roles in ICWSM, ACM Hypertext, ACM Multimedia, UMAP. He is among the founding members of GoodCityLife.org, a global network of scientist with the goal of giving a good life to city dwellers.

Daniele Quercia is Department Head of Social Dynamics at Nokia Bell Labs Cambridge (UK) and Professor of Urban Informatics at the Center for Urban Science and Progress (CUSP) at King's College London. He has been named one of Fortune magazine's 2014 Data All-Stars, and spoke about "happy maps" at TED. His research has been focusing in the area of urban informatics and received best paper awards from Ubicomp 2014 and from ICWSM 2015, and an honourable mention from ICWSM 2013. He was Research Scientist at Yahoo Labs, a Horizon senior researcher at the University of Cambridge, and Postdoctoral Associate at the department of Urban Studies and Planning at MIT. He received his PhD from UC London. His thesis was sponsored by Microsoft Research and was nominated for BCS Best British PhD dissertation in Computer Science.