Improving Personalised Physical Activity Recommendation on the mHealth Information Service Using Deep Reinforcement Learning

Ji Fang\textsuperscript{a}, Vincent CS Lee\textsuperscript{b} and Haiyan Wang\textsuperscript{a,*}

\textsuperscript{a}School of Economics and Management, Southeast University, Nanjing, 210096, PR China
\textsuperscript{b}Faculty of Information Technology, Monash University, Melbourne, VIC, 3053, Australia

\begin{abstract}
Recently has seen the growth in the use of mobile health (mHealth) information services, which have rich guides on improving physical activity. These rich guides evolved from the consideration of various personal behavioural factors, which often deviate from the user’s health conditions. The behavioural factors include changing fitness preferences, adherence issues, and uncertainty about future fitness outcomes, which may all lead to a decline in the quality of the mHealth information services. Many of these mHealth information services provide limited fitness guidance owing to the dynamics of the user’s health conditions. This paper seeks an adaptive method using deep reinforcement learning to make personalised physical activity recommendations, which is learnt from retrospective physical activity data and can simulate realistic behaviour trajectories. We construct a real-time interaction model for the mHealth information service system based on scientific knowledge about physical activity to evaluate its exercise performance. The physical activity performance evaluation model is used to find the optimal exercise intensity considering the fitness and fatigue effects to avoid the lack of exercise or overload. The short-term activity plans are made using deep reinforcement learning and personal health conditions that change over time. Using this method, we can dynamically update the physical activity recommendation policy in accordance with the real implementation behaviour. Our DRL-based recommender policy was validated by comparison to other benchmark policies. Experimental results show that this adaptive learning algorithm can improve recommendation performance over 4.13 percent.
\end{abstract}

1. Introduction

Mobile health (mHealth) information services, such as Fitbit and TrainingPeaks, collect individual physiology data and daily exercise information through wearable devices with multiple sensors. These services provide short-term physical activity recommendations to their users on improving physical activity [29]. For decades, wearable activity trackers (e.g., smartwatches, fitness trackers, head-mounted displays, etc.) and fitness applications have been widely used in health self-management, making exercise under the guidance of a scientific training system both affordable and convenient [1]. Up to now, several systematic reviews and empirical studies have demonstrated the effectiveness of mHealth services in physical activity interventions [10, 11, 12, 13]. By using real-time situation triggered reminders, pushes, and notifications, the mHealth service can assist the individual to improve the effectiveness of daily exercise [2, 3].

Although the mHealth information services have enormous potential for promoting physical activity regardless of time and location constraints, the challenges remain in what, how and why the evaluation and improvement of the performance of the physical activity recommendations. Three factors may contribute to the inefficiency of these mHealth physical activity recommendation services [6, 7, 8].

To begin, the users’ physical body fitness and emotional profile drive the varying influences on the mHealth intervention effect. The majority of popular mHealth services push non-personalised and predetermined physical activity plans to their users. There is mounting evidence that such fixed daily exercise plans may result in demotivating an individual’s physical activity [6]. Second, health improvement requires an individual to take every effort to adhere to daily exercise plans [7]. The individual may struggle with self-control problems, resulting in failure to complete the daily tasks [8]. The pre-designed exercise programme is ineffective in dealing with the individual’s uncertain behaviour. Third, synchronising physical activity recommendations are time consuming to comply due to the long duration of mHealth information services arrival as well as the uncertainty of future health outcomes. Consequently, the heterogeneity of individuals’ physical characteristics, fitness goals, and exercise preferences add complexities to the realisation of physical activity recommendation. Besides, the gaps between planning and implementation of the user’s behaviour result in uncertain physical activity performance. The changing preferences of the individual and the uncertainty of future outcomes exacerbate the difficulty in providing of this long-term service.

This study proposes a knowledge-based framework for the personalised mHealth recommendation system. The physical activity recommendation policy is obtained based on the individual physiological data collected from the wearable trackers, which enables the system to tailor recommendations to each user. Moreover, we integrate the
scientific training principle into our mHealth information service to model the practise process and optimise its profile to maximise performance for the individual during the service. Then, we consider the real-time interaction environment between the decision maker and the people who participate in physical activity through the wearable trackers. An adaptive policy is designed to solve the mHealth optimisation problem in the mHealth recommendation system. We can dynamically alter the physical activity recommendation policy during the service period if the user’s behaviour varies from the fitness plan.

1.1. Existing Approaches for Physical Activity Recommender System

Recently, there has been a growing concern about improving intervention effects in personalised physical activity recommendation services. For example, Aswani et al. [5] considered the adherence to the prescribed plans for automated exercise management problems. Mintz et al. [34] studied the non-stationary phenomena in personalised healthcare adherence-improving interventions. In a similar vein, we consider how to design the intervention policy for the personalised mHealth information service. The mHealth service is useful in changing behaviour and health outcomes by selecting among multiple potential practise plans for the user. According to our investigation, a whole range of different approaches have been developed in the mHealth information service, including mathematical models (e.g., [16, 9, 15, 17, 18]) and data-driven models (e.g., [28, 38, 36, 37, 35]).

Mathematical optimisation methods rely on exact models to characterise the environment (e.g., system states, human behaviour). Evaluating exercise performance is vast in physical activity recommendation for mHealth information services by mathematical optimisation methods. Analytical research on practise performance was proposed by Calvert et al. [15]. In their study, practise performance is associated with training intensity with a positive effect ("fitness") and a negative effect ("fatigue"). The fitness-fatigue model has been used to predict physical activity performance in a variety of endurance sports, including running [16, 9], swimming [15], soccer [17]. The existing literature on practise performance is extensive and focuses particularly on the process optimisation perspective. For example, Topol [18] presented a dynamic model of the practise process to maximise performance. His study sheds important insights and guidance on the optimal amount of spacing or optimal intensity of each practise in the process configurations. However, the mHealth service can only provide persuasive recommendations to engage users in physical activity, leading to the practise performance optimisation model not being applied directly in mHealth service. To our knowledge, practise process optimisation has not considered human behavioural factors.

Another type of attractive approach for the physical activity recommendation area is the data-driven approach, such as artificial neural networks (e.g., [28, 38]), support vector machines (e.g., [36, 37]), random forest approach (e.g., [35]), etc. However, these methods cannot solve the physical activity recommendation problem using the knowledge-based system because there have not been apparent methods to integrate them into an optimisation model [5]. Hence, a state-of-the-art learning method is needed to solve the personalised mHealth physical activity recommendations problem, which can establish a framework to achieve a knowledge-based mHealth information recommendation system, and enhance learning capability to train decision policies to adapt to the individual’s time-varying characteristics and the dynamic health improvement process.

For the purpose of providing low-effort to procedure outcomes with a high level of intervention effect, a reinforcement learning method is needed, which can constantly monitor the user activities and adjust the recommendation policy during the service period [20]. To test the effects of reinforcement learning technology in mHealth services, Yom et al. [21] experimented with a reinforcement learning approach on 27 sedentary type 2 diabetes patients to encourage them to increase the level of their physical activity. The results showed that the reinforcement learning improved gradually in predicting which intervention messages would lead patients to exercise. Forman et al. [20] evaluated the feasibility, acceptability, cost savings, and effectiveness of reinforcement learning for weight loss intervention in behavioural weight-loss trials. They observed that the reinforcement learning system can considerably lower costs in the optimal conditions while still producing a satisfying intervention effect. Reinforcement learning has already been considered for the physical activity recommendation service. Our approach differs through incorporating with uncertainties of users behavioural and personal health goals.

1.2. Limitations of Existing Approaches

This research aims to improve the performance of the personalised physical activity recommendation systems by addressing the shortcomings of the existing approaches. First, the typical mathematical optimisation methods in health service require a well-represented mathematical model for the decision processes. However, it is frequently difficult to develop an accurate model for the individual due to differences in activity preference, time-varying characteristics, and uncertainty regarding future health outcomes. In contrast, the physical activity recommender system proposed in our study uses prior experiences to develop an optimal execution strategy on its own. Second, the physical activity recommender system in this study is built on a commonly adopted model that uses professional physical activity expertise to describe exercise performance, improving the interpretability of machine learning methods. Third, no prior research focuses on the impact of exercise intensity among potential practise plans on the outcomes in health recommender systems (e.g., over exercise). Based on the research of the reinforcement learning method in physical activity recommender systems, we introduce a practise theory to
evaluate the performance of physical activities in health recommender systems. In the mHealth service, we fill a gap because there is limited guidance on the optimal intensity throughout the intervention period from the practice process optimization perspective.

1.3. Objective and Contributions

This study sheds new insights into physical activity recommendation research in the mHealth information service. The scientific training principle is introduced into our mHealth information service to evaluate physical activity performance, focusing on the interaction between the trainer and associated physical activity. The recommendation policy is derived based on the scientific training principle using data from the wearable trackers, which allows sequential evolution to tailor the individual’s changing health condition. The exercise plans over time adapt to the individual’s changing health conditions, which can be tracked via mobile device interaction traces. In contrast to conventional physical activity recommendations that are incapable of learning, our proposed dynamic system is a DRL-based recommender system that can learn from input sequential data and simulate realistic future health conditions. Then, this dynamic system uses the estimated model to generate physical activity recommendations that maximise the individual’s physical activity performance.

We then conduct experiments to test how our policy would have performed in real-world situations. The model is calibrated using sampling data, published data, and real-world data.

The findings indicate that our DRL policy outperforms fixed practise plans. While implementing an AI system for mHealth service providers requires an initial investment, these investments pay off over time with increased user satisfaction. The more satisfied the users are, the higher the purchase rate of the service; at the same time, the satisfaction. The more satisfied the users are, the higher these investments pay off over time with increased user satisfaction.

We suppose that the user can adhere to the exercise plan exactly, which the user’s exercise behaviour through the mobile application communication with wearable devices and then selects a short-term exercise plan from a pool of possible exercise plans based on exercise intensity calculations and an evaluation of the physical activity performance as a reward following the process optimization. The process optimisation relies on the retrospective physical activity data and the simulation of the realistic behaviour trajectories. Then, the DRL-based recommender agent provides fitness guidance to the user through the mobile application. The user makes an effort to implement the exercise plan, and the health condition is changed. After that, the DRL-based recommender agent will receive an observation of the user’s actual exercise intensity. The physical activity recommendation policy will be updated when the user’s implementation behaviour deviates from the exercise plan. The overview of the system architecture is depicted in Fig. 1.

2. Methodology

In this section, we establish the framework for the physical activity recommender system in mHealth information service by introducing the physical activity performance evaluation model. Then, we state the problem of physical activity recommendation in terms of the Markov Decision Process (MDP). Next, we introduce a deep reinforcement learning algorithm to train the recommender agent to provide physical activity recommendations.

2.1. The mHealth Physical Activity Recommender System

The personalised mHealth information service is described as a dynamic system with a sequential decision process. At the beginning of each decision epoch, the mHealth service provider (i.e., DRL-based recommender agent) monitors the user’s exercise behaviour through the mobile application communication with wearable devices and then selects a short-term exercise plan from a pool of possible exercise plans based on exercise intensity calculations and an evaluation of the physical activity performance as a reward following the process optimization. The process optimisation relies on the retrospective physical activity data and the simulation of the realistic behaviour trajectories. Then, the DRL-based recommender agent provides fitness guidance to the user through the mobile application. The user makes an effort to implement the exercise plan, and the health condition is changed. After that, the DRL-based recommender agent will receive an observation of the user’s actual exercise intensity. The physical activity recommendation policy will be updated when the user’s implementation behaviour deviates from the exercise plan. The overview of the system architecture is depicted in Fig. 1.

![Figure 1: The system architecture of the physical activity recommendation in individual behaviour (black arrows) and process optimization (red arrows)](image-url)

In our DRL-based recommender system, the agent can predict the physical activity performance based on the simulation of the realistic behaviour trajectories when push the specific exercise plan (action) \( a_t \) to the user. We suppose that the user can adhere to the exercise plan exactly, which can avoid overestimating or underestimating the user’s behaviour. When the user does not complete the exercise plan,
the DRL-based recommender system will revise the estimated physical activity performance according to the actual exercise intensity collected through the wearable devices and then update the following fitness guidance during the service period.

The physical activity performance evaluation assumes that the physical activity is divided into several practise sessions. Each practise session utility function for the individual is considered to have two components: (1) the user performance of exercise effort and (2) the impact of the mHealth recommendation intervention. According to [9], practise performance includes the effects of each workout on the "stock of fitness" $f_t$ and "stock of fatigue" $g_t$, which are measured in reference to an individual-specific baseline level of exercise $b_t$. Both stocks depreciate exponentially in the time interval between practise sessions. In addition, the formulation of the mHealth intervention service was inspired by [5], who examined the impact of achieving and failing to complete exercise goals. We combine the results of the two studies to build a model for evaluating the performance of the mHealth information service.

Compared with [9], we focus on the ongoing decision-making process for personalised physical activity plans, which is the real-time communication between the planner and the executor. Our research focuses on three aspects of improving personalised physical activity performance. To begin, the user’s changing preferences may have an effect on the exercise behaviour and thereby influence the health outcome for the mHealth recommendation policy. Second, we observe that the mHealth information service can only decide the exercise plans notification, not the actual behaviour of the user. Thirdly, we notice that the same intensity of activity recommendations can result in varying degrees of pressure and difficulty for different individuals. When evaluating the impact of mHealth intervention services, it is necessary to consider the individual motivation effect on exercise plans. To accomplish these goals, we train a learning agent that can detect an individual’s health condition through wearable trackers and adapt its actions to provide fitness guidance considering the user’s changing health condition. The wearable trackers can monitor and visualise the user’s health condition in real time. The objective of our DRL-based recommender agent is to learn the optimal recommendation policy that maximises both the individual physical activity performance and the recommendation effect during the service period.

2.2. Problem Formulation

We model the physical activity recommendation problem using the MDP model as a sequential decision process in an uncertain environment, considering the deviation between the fitness plan and the user’s implementation behaviour. Our objective is to maximise the individual’s cumulative physical activity performance and the mHealth recommendation effect during the service period. The physical activity is separated into two stages in the personalised mHealth information service [9], the stage of skill acquisition and the stage of skill retention. Repeated practise has a compounding effect on memories via a “reconsolidation” process, whereas practise cannot have a compounding effect until early skill acquisition occurs. As a result, the additive model is considered to be more appropriate for the stage of exercise skill acquisition, whereas the multiplicative model is considered to fit the exercise skill retention stage. We suppose that the skill stage of physical activity for the user during the service period is fixed, and the information is provided by the user. In the mHealth information service, the DRL-based recommender cannot identify the user’s fitness guidance demand. Therefore, the user chooses their exercise type between the beginner user and the advanced user before using the mHealth information service.

The DRL-based recommender agent seeks a recommendation policy to maximize the total reward during a finite service period. For the cumulative rewards of the physical activity performance, the impact of the previous service is included in the state definition because the previous exercise frequently has long-lasting effects, which is included in the state definition following the Markov property. Fig. 2 depicts the sequence of events within a single decision epoch. At the beginning of each decision epoch $t$, the DRL-based recommender agent chooses physical activity plan with exercise intensity $a_t$. The DRL-based recommender agent will then evaluate a reward based on the physical activity performance evaluation function with the assumption that the user adhere the physical activity plan. The physical activity performance evaluation function considers the exogenous information associated with service content about the individual’s physical activity (i.e., in the skill acquisition stage or in the skill retention stage). The reward $r_t$ is also related to the individual’s former state $s_{t-1}$ at period $t - 1$. After this recommendation action, the individual will make effort to complete the physical activity plan, and the user’s real exercise intensity level $e_t$ will be observed by the DRL-based recommender agent via sensors. The evaluation of health condition $s_t$ will transition according to the Markov property.

A subscript $t$ denotes the value of a variable on the $t$-th decision epoch ($t = 0, 1, ..., T$). At first, the decision maker decides the exercise plan with exercise intensity $a_t$ among potential exercise plans in time $t$ in the mHealth information service, where receive a reward $r_t$ according to physical activity performance evaluation model with the assumption that the user can complete the exercise plan to avoid excess intervention the user behaviour. We adopt $s_t = [e_t, b_t, f_t, g_t]$ to represent the state variables of the individual, containing all the necessary information for decision making. The $e_t \geq 0$ is random variable represents the intensity of the workout in time $t$. The $b_t > 0$ is defined as the base level of past practices. The $f_t, g_t \geq 0$ represent the fitness and the fatigue at time instance $t$ respectively. We refer to $\chi = (\alpha, \beta, \lambda, \mu, \delta, k_f, k_g)$ as the type of the individual. The parameters $\alpha, \beta \in (0, 1)$ represent the decay rates of fitness and fatigue respectively causing by forgetting (for fitness)
or recovery (for fatigue), $\lambda, \mu > 0, \lambda \leq 1, \mu \geq 1$ represent the nonlinearities in the response of fitness and fatigue to exercise effect, including the concave effect on fitness and the convex effect on fatigue, and $0 < \delta \leq 1$ measures some decay in the base level. Another set of parameters are used in the utility function. These include $k_f, k_g > 0$ which represents the marginal utility of $f_t$ and $g_t$. The $m \geq 0$ captures the increase in utility experienced if the exercise plan is achieved. Based on reasonable assumptions, it is supposed that regardless of how much physical activity intensity level a user achieved, the increase in utility for goal achievement is the same. The $f \geq 0$ represents the impact of failing the exercise plan. The greater the degree of the unfinished physical activity there is, the more decrease in utility experienced caused by frustration.

We define the various MDP elements, culminating in the presentation of the optimality recommendation policy expression.

### 2.2.1. Decision epochs

The decision epochs in our problem are denoted by $t = 0, 1, ..., T$, where $T$ denotes the length of service. The mHealth service provider is considered to deliver practise recommendations on a periodic basis in time $t$.

### 2.2.2. State space

The state at period $t$, denoted by $s_t \in S$, where $s_t = [e_t, b_t, f_t, g_t]$. The state $s_t$ contains the individual exercise status and the accumulated impacts on practice performance. $S$ represents the set of states $s_t$.

The individual exercise state at each period is denoted by $e_t \in E$, where $E$ contains discrete exercise intensity levels determined by the precision of wearable devices. The precision of mHealth system in our experiment is 0.01, so $E = \{E_0, E + 0.01, E + 0.02,..., \bar{E}\}$, where $E_0$ and $\bar{E}$ are the minimum and the maximum exercise intensity levels recorded by the wearable devices, respectively.

### 2.2.3. Action space

At each decision epoch $t$, the mHealth service determines an exercise plan with intensity $a_t \in A$, assumed to be uni-dimensional, where $A = \{0\} \cup \{A_1, A_2, ..., A_n\}$ contains no exercise plan, and discrete exercise intensity $A_i (i = 1, 2, ..., n - 1)$.

### 2.2.4. Reward functions

The immediate reward function $r_t(s_t, a_t)$ captures the one-period performance of mHealth services and effect of mHealth interventions when the individual takes the exercise intensity level $e_t$ after receiving the mHealth recommendation $a_t$ in the state $s_t$. Combining the research on [9, 5], the following reward function was selected, the utility of each practise session for the individual includes the fitness effect, fatigue effect, and the impact of mHealth recommendation intervention. We assumed that the base level $b_t$ is the max mean intensity of past practices based on the study in [9]. The less physical activity effort is required if one has practiced at high intensity level in the past, especially in the recent past. Thus, $b_t = \delta \cdot \max(e_{t-1}, b_{t-1})$, where $0 < \delta \leq 1$ represents the decay rate of the base level. When the individual fulfills the physical activity plan, the increase $m$ in utility experienced will be occurred, where $\mathbb{I}(\cdot)$ is the indicator function (i.e., when $e_t \geq a_t$, $\mathbb{I}(e_t \geq a_t) = 1$, otherwise it is equal to zero).

- **Skill acquisition stage (additive model):**

  $$
  r_t(s_t, a_t) = r_0 + k_f \left( a f_{t-1} + \left( \frac{a}{b_t} \right)^{\lambda} \right) - k_g \left( b g_{t-1} + \left( \frac{a}{b_t} \right)^{\mu} \right) + m \cdot \mathbb{I}(e_t \geq a_t) + l \cdot \min(0, e_t - a_t) \tag{1}
  $$

- **Skill retention stage (multiplicative model):**

  $$
  r_t(s_t, a_t) = r_0 + \frac{k_{f} f_{t-1} \left( a + \left( \frac{a}{b_t} \right)^{\lambda} \right)}{k_{g} g_{t-1} \left( a + \left( \frac{a}{b_t} \right)^{\mu} \right)} + m \cdot \mathbb{I}(e_t \geq a_t) + l \cdot \min(0, e_t - a_t) \tag{2}
  $$

The purpose of the decision maker is to provide effective mHealth service, improving the individual’s utility during the service period. The objective function can be expressed as follows:

$$
\max_{a_t} \sum_{t=1}^{T} r_t \tag{3}
$$

### 2.3. Deep Reinforcement Learning Algorithm

In the personalised mHealth optimisation problem, no assumption is made on the individual’s exercise behaviour, considering the unexpected effects of external influences. Additionally, because the individual’s state space and the

---

Figure 2: The sequential of events with a single decision epoch

---
mHealth alternative recommendation action spaces are discrete and quite large, solving this practise plan optimisation problem will require a significant amount of computation time and storage space. Furthermore, an individual’s exercise preference frequently changes randomly depending on his or her state of mind and mood. Hence, the resulting optimisation problem is intractable using traditional methods such as dynamic programming or genetic algorithms with a time-dependent fitness function.

As an alternative, deep reinforcement learning (DRL) is a promising tool for efficiently solving problems. Instead of relying on prior knowledge of an individual’s changing behaviour, the DRL methodology can learn from experience. This method exacts sequences of states, actions, and rewards from real-world interactions. In addition, we improve the DRL methodology by constructing a hybrid neural network that can capture temporal information about an individual’s exercise history using a long short-term memory (LSTM) neural network and improve the fitting effects of the fully connected neural network.

In this study, the stochastic decision policy is trained using a combination of the generalisation method (actor network) and the parameterized policy method (critic network) [26]. The DRL algorithm trains our hybrid neural network using two major techniques: asynchronous training work) and the parameterized policy method (critic network) and the parameterized policy method (critic network) using a combination of the generalisation method (actor network) and the parameterized policy method (critic network).

2.4. The Networks Architecture

The following describes the training process. To begin, the “Critic” network’s parameters $\theta$ can be trained using gradient descent to minimise the mean square error. The “Actor” then updates the policy distribution in the direction suggested by the “Critic”. The complete architecture network aims to realise the dual evolution: exploitation and exploration principle, which is embedded in the Deep Reinforcement Learning cycle.

### 2.4.1. The critic network

The critic network combines an LSTM layer and a deep neural network with multiple fully connected hidden layers to create a hybrid forecaster model. The main structure of the critic network is depicted in Fig. 4. The critic network is composed of an input layer that feeds data into a single LSTM layer. The input data for our personalised mHealth optimisation problem is a three-dimensional vector consisting of state characteristics, the stack of the previous state, and time. The state characteristics are the individual’s exercise intensity $e$, the practice base level $b$, the fitness level $f$, and the fatigue level $g$ in period $t$. The stack of the previous state is constructed because the reward is conditional on both the current and the previous states. The term “time” refers to discrete-time, which is equal to one. The LSTM is equipped with feedback connections that enable it to learn and model the sequential relationships present in the time series data. Then, we construct a deep neural network for the LSTM, which can learn from the previous layer and develop new representations for the approximation of the action-value function. Our critic network has $l$ hidden layers, where the first hidden layer is with the LSTM cells, and the others are fully connected layers. The number of neural cells in the hidden layer $i$ is $y_i$, $i = 1, \ldots, l$. The $h_{k|i}$ represent the outputs of the LSTM cell in time $t$, where $k = 1, \ldots, y_1$, and $C_{t|i}$ represents the long-term memory, where $t = 0, 1, \ldots, T$. Besides, the outputs of neural cells in the hidden layer $i$ are denoted as $h_{k|i}$ where $k = 1, \ldots, y_i$, $i = 2, \ldots, l$. The parameters used in this critic network include the weight parameters of the neural network $\phi$ and the bias parameters $\xi$. The $\sigma(\cdot)$ represent the activation functions in the critic network.
An illustration of the LSTM cell is given in Fig. 4, where the grey circles represent pointwise operations while the white boxes are learned neural network layers. The LSTM can remove or add information to the cell state by a hidden state \( h_t \) to store the short-term memory and a cell state \( C_t \) to store the long-term memory in time \( t \). Specifically, the LSTM cell is regulated by three gates. The forget gate \( f_t \) controls the range to which a value remains in the LSTM cell. The input gate \( i_t \) controls the recording of the latest information in the individual’s state. The output gate \( o_t \) selectively selects information to compute the output activation of the LSTM unit. For instance, the calculation process of the critic network at time \( t \) is defined as follows. The architecture of the critic network consists of one LSTM layer and \( l-1 \) fully connected layers. In the LSTM cell \( k(k = 1, 2, ..., y_1) \), the first step is to transform the three-dimensional input data into a matrix \( X_t \). Then, decide what information will be thrown away from the cell state controlled by the forget gate \( j_t \), taking the previous hidden state \( h_{k,(t-1)} \) and the current information \( X_t \) as input and generating a number between 0 and 1 for each number in the cell state \( C_{k,(t-1)} \) as output by the sigmoid activation function \( \sigma_{\text{sigmoid}}(\cdot) \).

\[
i_t = \sigma_{\text{sigmoid}}(\varphi_f \cdot [h_{k,(t-1)}, X_t] + \xi_f).
\]

The next step is to decide what new information will be stored in the cell state controlled by the input gate \( i_t \). In particular, a neural network layer \( i_t \) with the sigmoid activation function decides what values will be updated in the cell stat. Then, a neural network layer with a hyperbolic tangent activation function \( \sigma_{\text{tanh}}(\cdot) \) generates a vector of new candidate values \( \tilde{C}_{k,(t)} \), which could be added to the cell state \( C_{k,(t)} \). Finally, the old cell state \( C_{k,(t-1)} \) will be updated into the new cell state \( C_{k,(t)} \) by multiplying \( C_{k,(t-1)} \) by \( j_t \), forgetting the information predetermined by the forget gate, and then adding new information with \( \tilde{C}_{k,(t)} \) scaling by \( i_t \).

\[
i_t = \sigma_{\text{sigmoid}}(\varphi_f \cdot [h_{k,(t-1)}, X_t] + \xi_f), \quad \tilde{C}_{k,(t)} = \sigma_{\text{tanh}}(\varphi_C \cdot [h_{k,(t-1)}, X_t] + \xi_C),
\]

\[
C_{k,(t)} = j_t \cdot C_{k,(t-1)} + i_t \cdot \tilde{C}_{k,(t)}.
\]

After that, a neural network layer \( o_t \) with the sigmoid activation function decides what parts of the LSTM cell will be output. Then, the cell state \( C_{k,(t)} \) can be normalized between -1 and 1 through hyperbolic tangent function and multiply it by the \( o_t \) to obtain the output \( h_{k,(t)} \) of the LSTM cell.

\[
o_t = \sigma_{\text{sigmoid}}(\varphi_o \cdot [h_{k,(t-1)}, X_t] + \xi_o), \quad h_{k,(t)} = o_t \cdot \text{tanh}(C_{k,(t)}).
\]

Then, the outputs from the LSTM layer \( h_{k,(t)}, k = 1, ..., y_1 \) in time \( t \) are connected to a fully connected network of the hidden layers. Mathematically, the output of the neural cell \( k \) in the layer \( i \) is given by the following equation:

\[
h^2_k = \sigma_{\text{ReLU}} \left( \frac{1}{y_l} \sum_{j=1}^{y_l} \varphi_{h_{j,k}} h^1_{j,(t)} + \xi_{h^1_k} \right), k = 1, 2, ..., y_2,
\]

\[
h^1_k = \sigma_{\text{ReLU}} \left( \frac{1}{y_i} \sum_{j=1}^{y_i} \varphi_{h^0_{j,k}} h^0_{j,(t)} + \xi_{h^0_k} \right), k = 1, 2, ..., y_1, i = 3, ..., l.
\]

Finally, the critic network produces the output of the estimated value of the value function:

\[
v_x(s_t; \varphi) = \sum_{j=1}^{y} \varphi_{h^0_{j}} h^{1}_{j(t)} + \xi_{h^0_j}.
\]
The activation function \( \sigma(\cdot) \) includes sigmoid function \( \sigma_{\text{sigmoid}} \), hyperbolic tangent function \( \sigma_{\tanh} \), and ReLU function \( \sigma_{\text{ReLU}} \). The activation function can be formulated as follows:

\[
\begin{align*}
\sigma_{\text{sigmoid}}(x) &= \frac{1}{1 + e^{-x}}, \\
\sigma_{\tanh}(x) &= \frac{e^x - e^{-x}}{e^x + e^{-x}}, \\
\sigma_{\text{ReLU}}(x) &= \max(x, 0).
\end{align*}
\] (9)

2.4.2. The actor network

In our DRL methodology, an actor will learn a policy to select an action in a given state \( s_t \). The actor network is constructed to learn the parameterized policy \( \pi(a_t|s_t; \theta) = P[a_t|s_t; \theta] \). The actor network is similar to the critic network, combining an LSTM layer and a deep neural network with fully connected layers.

The actor network has the same architecture as the critic network with the weight parameters of the neural network \( \theta \). Because the daily personal practice behaviour may be subject to random disturbance from external sources, the stochastic policy is better than the determined policy in the personalized practice optimization problem. The only difference between the actor network and the critic network is computing in the output layer. We introduce the Softmax function as the activating function in the neural cells in the output layer.

\[
\pi_i(a_t|s_t; \theta) = \frac{\exp(h_i^t)}{\exp\sum_{j=1}^{n} h_j^t}, \quad i = 1, \ldots, n
\] (10)

where \( n \) represents the number of action space \( \mathcal{A} \).

2.5. Algorithm of the Proposed Model

Since the DRL methodology has two networks, including the critic network and the actor network. The critic network is constructed to estimate the state-value function \( v(s_t; \phi) \) and \( v(s_{t+1}; \phi) \) for current policy \( \pi \), which can be trained by the mean square of the advantage function. The advantage function is introduced to assess how much better it is to take action than the average, general action at the state \( s_t \). It is considered the state-value function \( v_x(s_t) \) as a baseline to evaluate an action for current policy \( \pi \) is given as follows:

\[
A_x(s_t, a_t) = q_x(s_t, a_t) - v_x(s_t)
\] (11)

We adopt the sampling method to evaluate the action-value \( q_x(s_t, a_t) \). Therefore, the advantage function can be express as follows:

\[
A_x(s_t, a_t) = r_t + v_x(s_{t+1}) - v_x(s_t) \approx r_t + v(s_{t+1}; \phi) - v_x(s_t; \phi) \] (12)

Therefore, the accumulated gradient of the weight \( \phi \) of the critic network is obtained as follows.

\[
d\phi \leftarrow d\phi + \frac{\partial (A_x(s_t, a_t))^2}{\partial \phi}
\] (13)

On the other hand, the actor updates the policy distribution in the direction suggested by the critic. According to the policy gradient theorem, using the gradient descent with root mean square propagation (RMSProp) algorithm to update the parameters \( \theta \) [39, 40].

\[
d\theta \leftarrow d\theta + \nabla_{\theta} \log \pi(a_t|s_t; \theta) A_x(s_t, a_t)
\] (14)

To train the networks, we employ the forward view of an \( n \)-step method and an asynchronous technique. The forward view of \( n \)-step method may improve convergence speed in the DRL methodology. The asynchronous approach used in our algorithm enables significantly quick exploration speed of a larger state-action space proposed by [26], which designs multiple independent agents interacting with different copies of the environment in parallel. In our algorithm, the shared parameters in the global network are updated periodically through the parallel training of the agents. After each update, the agents reset their network parameters to those of the global network and continue the exploration and training for \( n \) steps to the next update. The pseudo-code of the personalized physical activity recommendation algorithm is given below.

Algorithm 1: Physical activity recommendation algorithm.

**Input:** Physical activity environment, DRL-based agent neural networks weights \( \theta, \phi \). Upper bound of global shared step counter \( T_{\text{max}} \), Learning rate \( \alpha > 0, \omega > 0 \), Thread step counter \( n \) in each thread.

1. Initialize global shared step counter \( T \leftarrow 0 \)
2. while \( T < T_{\text{max}} \) do
3. while each thread \( i \leftarrow 1 \) to \( N \) do
4. Initialize thread step counter \( t \leftarrow 1 \)
5. Reset accumulate gradients: \( d\theta \leftarrow 0, d\phi \leftarrow 0 \)
6. Synchronize thread-specific parameters: \( \theta' \leftarrow \theta, \phi' \leftarrow \phi \)
7. \( t_{\text{start}} \leftarrow t \)
8. Initialize state \( s_t \)
9. while \( s_t \) is terminal or \( t - t_{\text{start}} \equiv n \) do
10. Take action \( a_t \) according to policy \( \pi_{\theta}(a_t|s_t) \)
11. Compute reward \( r_t \) using (1) or (2)
12. Observe new state \( s_{t+1} \)
13. \( t \leftarrow t + 1, T \leftarrow T + 1 \)
14. end while
15. \( R = \left\{ \begin{array}{ll} 0 & \text{for terminal } s_t \\ V(s_t; \phi') & \text{for non-terminal } s_t \end{array} \right. \)
16. while \( j \leftarrow t - 1 \) to \( t_{\text{start}} \) do
17. \( R = r_j + R \)
18. \( d\theta \leftarrow d\theta + \nabla_{\theta} \log \pi_{\phi'}(a_t|s_t)(R - V(s_t; \phi')) \)
19. \( d\phi \leftarrow d\phi + \partial((R - V(s_t; \phi'))^2) / \partial \phi' \)
20. end while
21. Perform update of \( \theta \) using \( d\theta \) and of \( \phi \) using \( d\phi \)
22. end while
23. end while
3. Experimentation

DRL-based physical activity recommender system is implemented using TensorFlow in Python. We conduct experiments to evaluate the effectiveness of the proposed DRL-based recommender system in mHealth information services. The experiments test different types of individual physical activity increasing responses in the construction of the environment due to the individual’s responding differently to the mHealth recommender service. We split the time series data for in-sample and out-of-sample as 80% : 20% for training and simulation, respectively.

3.1. Data Set

The data set contains personalised characteristic information as well as daily exercise intensity data. The former individual-specific parameters can be estimated using the statistical techniques [9]. The raw acceleration data can be handled in the structured physical activity intensity data [28].

The first study’s data is generated from the literature with different participant groups. We randomise the individual-specific parameters based on the configuration of the parameters in Section 6.1. The daily exercise intensity is generated randomly following the mean values and standard deviations in the literature. The second study’s data is produced from an open-source dataset [44] collecting from 16 persons for five months, which combines lifelogging data from the PMSys sports logging smartphone application and sports-activity data from Fitbit smartwatch. We considered the participants perceived exertion level as the exercise intensity level excluded for 5 participants data missing. The third study’s data is performed by our research team, collected from 5 adult participants running data in age range from 28 to 31 during 12 weeks without any physical activity recommendations. The wearable device Fitbit is provided to measure raw activity data during the running. The raw data was preprocessed into structural data based on the professional knowledge of running science in Section 6.2. We selected Training IMPulses to estimate the running activity intensity, the $VO_{2}\text{Max}$ to evaluate the physical activity performance [45]. Then, the individual-specific parameters were estimated by the nonlinear least-squares method in Python to minimize the residual sum of squares.

### Table 1
Details of the Data Set

| Dataset        | Participants | Intensity Measures | Resources or sensors |
|----------------|--------------|--------------------|----------------------|
| G1 Simulation from literature | 50 participants | Walking steps (mean ± SD) 6274 ± 2106 | [46] |
| G2 Open-Source data | 11 participants | Perceived exertion (mean ± SD) 2.41 ± 3.02 | [44] |
| G3 Real-World data | 5 participants | Training IMPulses (mean ± SD) 23.38 ± 15.17 | Fitbit |

3.2. Environment Generation

Many researchers have studied the effects of increasing physical activity levels through mobile device interventions [51, 52]. We review the existing scientific literature to construct the interactive environment for the DRL-based recommender agent in order to simulate the personalised behaviour change towards the interventions. Studies investigating efforts to increase physical activity level through mHealth information services are divided into four types. Some studies showed that the mHealth intervention was not effective in increasing physical activity [46, 48, 53]. Kirwan et al. [46] compared the intervention group using a website-delivered physical activity programme with the matched group during a 90-day study, showing a nonsignificant relationship between perceived usability and usefulness of the mHealth service. Most of the interventions showed significant increases in physical activity levels through mHealth services reported in [52, 49, 47]. Taking into account the physical activity level increases in proportions, two degrees of increase are considered in this study. The final environment is designed based on Gonce et al. [50], they observed a decrease and an increase trend in the number of steps per day throughout the intervention in the sequential multiple assignment randomised trial. The environmental setting of this study is shown in Table. 2.

### Table 2
Individual physical activity increasing response

| Physical activity increasing trend | Study | Assumption of the physical activity intensity trend in our study |
|----------------------------------|-------|---------------------------------------------------------------|
| E1 No effect                     | [46]  | Remain stable during 12 weeks                                  |
| E2 Slighty increase              | [47]  | Increases by 40% after 8 weeks                                |
| E3 Highly increase               | [46]  | A two-fold increase after 6 weeks                             |
| E4 A down and an increase trend  | [50]  | Decreases by 20% during 6 weeks, increases by 60% during 6 – 12 weeks |

3.3. Results

All simulations are run in Python using TensorFlow and run on a PC with the Intel (R) Core (TM) i7 CPU with 32.00 GB of RAM. To evaluate the performance of the proposed DRL-based recommender algorithms and compare them against the performance without any mHealth intervention service, we tested three experiment objectives G1, G2, and G3 with different assumptions that the personalised reaction four types in physical activity will increase levels (Environment E1, E2, E3 and E4). Each simulation is run 100 times for statistical purposes. We observe that each simulation converges in 6500 steps, and takes less than 355 seconds.

For demonstration purposes, Fig. 5, 6 display the daily rewards and the clustering of daily physical activity recommendations during the 30 day service period. It can be observed that our DRL-based policy can adapt to the individual’s physical activity behaviour because our DRL-based policy causes larger fluctuations and performs better of the daily physical activity reward than without the mHealth service. Besides, it can be observed that when the individual...
is in a high level of physical activity intensity increase (i.e., Environment E3), the DRL-based policy may cause slightly better performance and it is more inclined to choose high intensity physical activity plans than in other environments.

4. Discussion

4.1. Benchmark Policies

We compare our DRL-based policy with the current recommendation policies, including without any physical activity recommendation, weak fixed physical activity intensity recommendation policy, slightly weak fixed physical activity intensity recommendation policy, slightly strong fixed physical activity intensity recommendation policy, and strong fixed physical activity intensity recommendation policy.

The results of all three experiments show in Table 3, which indicates that our DRL-based policy is superior to the other fixed policies and without mHealth service on average. The third line for each benchmark policies shows how much improvement DRL-based recommendation policy provides compared to the other benchmark policies. It can be observed that DRL-based policy improving the performance by more than 4.13%.

We show the experiment results for G3 in the skill acquisition stage in Fig. 7 as an example. The results for other experiment objectives G1, G2 and in the skill retention stage are in accordance. However, it can be observed that the fixed policies with weak physical activity intensity perform better than our DRL policy in some instances. This result may be explained by the fact that the individual is physically inactive in some special situations, where physical activity plans with low intensity are more likely to be completed. In addition, it may cause the individual to be discouraged when the individual challenges and difficulties completing the high-intensity physical activity plans.
4.2. Comparison With State-of-the-Art

We compare our DRL-based recommender algorithm with other state-of-the-art algorithms. The competitor algorithms are implemented on the OpenAI-Baseline [41]. Each simulation is run 100 times for statistical purposes. Two of the six other algorithms are in the same asynchronous advantage actor-critic (A3C) learning algorithm with different network architectures. Others are in the different reinforcement learning algorithms with the same network architectures.

4.2.1. Our A3C algorithm on different network architectures

In this category, all reinforcement learning approaches use the same A3C learning algorithm but employ different network architectures. The network architectures include multi-layer perceptrons (MLPs), long short-term memory (LSTM) recurrent neural networks [42]. The number of neurons in the hidden layers of the A3C-MLPs, A3C-LSTM are determined by empirical experiments: the A3C-MLPs has two hidden layers with 64 neurons, the A3C-LSTM has a hidden layer with 32 neurons.

4.2.2. Our hybrid network architecture but different reinforcement learning algorithm

In this category, all reinforcement learning approaches use exactly the same hybrid network architecture as used in our approach. However, instead of using the A3C reinforcement learning algorithm, these models use different learning algorithms, namely DQN, ACKTR, PPO, and GAIL [43].

Table 4 shows the results for all the algorithms. It can be observed that our DRL methodology consistently outperforms all competitors in terms of computing complexity and average speed. Consider using the A3C learning method on different network architectures (MLP, and LSTM). Our DRL methodology outperforms these algorithms significantly in averaging speed. This shows that the hybrid neural network architecture is useful for learning the user’s exercise behaviour in the personalised mHealth service system. On the other hand, our DRL methodology outperforms others with the same network architecture in different reinforcement learning algorithms (DQN, ACKTR, PPO, and GAIL). This shows the effectiveness of the A3C algorithm used to train our hybrid network, and find the optimisation recommendations for the mHealth information service.

4.3. Sensitivity Analysis

The parameters related to the mHealth intervention service include the impacts of completing the exercise plan m and the marginal disutility of failing the exercise plan l. These exogenous factors can be assessed through a scale of experimental factors. Because of the conditions, we analyse the sensitivity of the parameters of the effect of mHealth intervention services. The three experiment objectives in different environments show the same tendency in the parameters m, l sensitivity analysis. We take experiment objective G3 in environment E4 as an example to show the sensitivity analysis.

Figure 8(a),(b) shows that increasing the impacts of completing the exercise plan m will increase the average performance of the personalised utility and may cause a more
Table 4
Results of the proposed DRL methodology and the competitor algorithms

| Algorithms     | Model       | Average running time (second) | Average converging steps |
|----------------|-------------|------------------------------|--------------------------|
| DRL method     | skill acquisition | 351.09                       | 4500                     |
|                | skill retention     | 367.46                       | 5100                     |
| A3C-MLP        | skill acquisition     | 426.71                       | 5400                     |
|                | skill retention       | 410.38                       | 5340                     |
| A3C-LSTM       | skill acquisition     | 431.08                       | 4960                     |
|                | skill retention       | 405.91                       | 5250                     |
| DQN-hybrid     | skill acquisition     | 422.13                       | 9600                     |
|                | skill retention       | 449.12                       | 10000                    |
| ACKTR-hybrid   | skill acquisition     | 361.61                       | 6800                     |
|                | skill retention       | 374.27                       | 6100                     |
| PPO-hybrid     | skill acquisition     | 349.56                       | 6400                     |
|                | skill retention       | 348.62                       | 6260                     |
| GAIL-hybrid    | skill acquisition     | 368.43                       | 9100                     |
|                | skill retention       | 380.19                       | 5940                     |

5. Conclusion

Mobile health information services are a new trend in self-health management, particularly during the COVID-19 pandemic. The recommender system in the mHealth information service records and visualises the individual health conditions through wearable trackers and then delivers a set of physical activity plans for daily exercise with motivational messages. Throughout the service, this artificial intelligence system can adapt to the user’s ongoing exercise preferences, implementation behaviour, and uncertain health outcomes. The adaptiveness and learning capability of the health management system have been improved by over 4.13 percent comparing with the benchmark policies pushing the fixed exercise recommendations to the user. The experiment results show that our DRL methodology can be tailored to fit personalised behaviour. The interaction between the mHealth DRL-based recommender agent and the individual physical activity environment through wearable devices enables the ongoing learning of recommender policy using observed information from wearable devices.

The DRL-based recommender system considers the personal behaviour factors. Through better awareness, the DRL-based recommender agent can provide better physical activity guidance, and increase the user’s service satisfaction. For example, exercise equipment and media companies could provide customised physical activity courses for the digital membership through a wireless device installed in the equipment (record the member exercise progress). The customised classes could potentially be scheduled about physical activity intensity level using the proposed DRL-based approach to improve the guidance effect of the users’ physical activity and thus increase users’ dependency and engagement. Although, the artificial intelligence system may cost the fixed investment at the beginning. It will improve the user’s satisfaction during the healthcare information service, improve customer loyalty. Over the long term, it will be valuable to update the system for the mHealth information service.

A limitation of this study is the lack of measurement of parameter values for mHealth interventions, including the impacts of completing the exercise plan and the marginal disutility of failing the exercise plan. Notwithstanding this limitation, the study suggests that the incentives for completing the exercise plans can improve the expected average performance of the mHealth service. Meanwhile, low incentives may cause mHealth users’ satisfaction to be less than those who do not use the mHealth service. On the other hand, the study suggests that low punishment for failing the exercise plan does not improve individual satisfaction during the service. Therefore, the service provider needs to use appropriate rewards and punishments to encourage the individual to participate in physical activity. An extension of the current study is to employ the behavioural experiments to qualify the impacts of the mHealth interventions and set the optimal levels of rewards and punishments.
6. Appendix

6.1. Personalised-Parameters

There are differences among individuals in their physical characteristics. We randomly generate a set of parameters representing the specific characteristics for the experiment objectives G1 and G2. The configuration of the parameters in the initial set of experiments in shown is Table 5 based on the model assumption in the Section 2.2.

6.2. Running Data Preprocess

The professional knowledge of running is applied in running data preprocessing inspired by [9]. We estimate the runner’s exercise intensity by Training IMPulses (TRIMPs) provided by [45]. Each practise session is characterised as easy running (E), marathon-pace running (M), threshold running (T), interval training (I), and repetition training (R). The recording points associated with various intensities of training can be calculated by the following rough estimates: 0.2 points/min for E running, 0.4 points/min at M pace, 0.6 points/min at T pace, 1.0 points/min at I pace, and 1.5 points/min at R pace [45] in page 90. The speeds of different types of training can be estimated using the VDOT training tables, where the VDOT value is based on the runner’s most recent running time in [45].

The running performance without the mHealth intervention can be measured on the raw data by using the runner VO2Max, which is the maximum oxygen consumption (in milliliters) per minute per kilogram of body weight. As [9] states: ‘the VO2Max is considered as a good, though imperfect, predictor of performance’. The VO2Max is collected by the Fitbit wearable tracker at the end of practise sessions. [45] has established the VO2Max running calculator gathered from years of testing many runners of various ability levels.

6.3. Hyper-Parameters

In Table 6, we list the hyper-parameters performing limited tuning that have been used in our DRL-based recommender agent. We tested the learning performance of the algorithm for the neural network with the number of hidden layers and the number of neurons in the networks. Other hyper-parameters are used in the physical activity recommendation algorithm 1 which have an impact on the speed of convergence and the performance of learning.

CRediT authorship contribution statement

Ji Fang: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing-Original draft preparation. Vincent CS Lee: Supervision, Validation, Writing-Reviewing & Editing. Haiyan Wang: Supervision, Project administration, Funding acquisition, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Y. Liu and M. Avello, “Status of the research in fitness apps: A bibliometric analysis,” Telematics and Informatics, vol. 57, p. 101506, 2021.
[2] W. Hardeman, J. Houghton, K. Lane, A. Jones, and F. Naughton, “A systematic review of just-in-time adaptive interventions (jitais) to promote physical activity,” International Journal of Behavioral Nutrition and Physical Activity, vol. 16, no. 1, pp. 1–21, 2019.
[3] S. Scheppe, S. Alley, A. L. Rebar, M. Hayman, N. A. Bray, W. Van Lippevelde, J.-P. Gnam, P. Bachert, A. Direito, and C. Van delanotte, “Apps to improve diet, physical activity and sedentary behaviour in children and adolescents: a review of quality, features and change behaviour techniques,” International Journal of Behavioral Nutrition and Physical Activity, vol. 14, no. 1, pp. 1–10, 2017.
[4] F. Gasparetti, L. M. Aiello, and D. Quercia, “Personalized weight loss strategies by mining activity tracker data,” User Modeling and User-Adapted Interaction, vol. 30, no. 3, pp. 447–476, 2020.
[5] A. Aswani, P. Kaminsky, Y. Mintz, E. Flowers, and Y. Fukuoka, “Behavioral modeling in weight loss interventions,” European journal of operational research, vol. 272, no. 3, pp. 1058–1072, 2019.
[6] M. Zhou, Y. Fukuoka, Y. Mintz, K. Goldberg, P. Kaminsky, E. Flowers, and A. Aswani, “Evaluating machine learning–based automated personalized daily step goals delivered through a mobile phone app: Randomized controlled trial,” JMIR mHealth and uHealth, vol. 6, no. 1, p. e9117, 2018.
[7] S. Jain, “Self-control and optimal goals: A theoretical analysis,” Marketing science, vol. 28, no. 6, pp. 1027–1045, 2009.
[8] T. O’Donoghue and M. Rubin, “Doing it now or later,” American economic review, vol. 89, no. 1, pp. 103–124, 1999.
[9] G. Roels, “High-performance practice processes,” Management Science, vol. 66, no. 4, pp. 1509–1526, 2020.
[10] I. Adjerid, G. Loewenstein, R. Purta, and A. Striegel, “Gain-loss incentives and physical activity: the role of choice and wearable health tools,” Management Science, 2021.
10,000 steps program: a matched case–control trial,” *Journal of medical Internet research*, vol. 14, no. 2, p. e1950, 2012.

[47] E. C. Voth, N. D. Oelke, and M. E. Jung, “A theory-based exercise app to enhance exercise adherence: a pilot study,” *JMIR mHealth and uHealth*, vol. 4, no. 2, p. e4997, 2016.

[48] R. Sirriyeh, R. Lawton, and J. Ward, “Physical activity and adolescents: an exploratory randomized controlled trial investigating the influence of affective and instrumental text messages,” *British journal of health psychology*, vol. 15, no. 4, pp. 825–840, 2010.

[49] B. H. Kim and K. Glanz, “Text messaging to motivate walking in older african americans: a randomized controlled trial,” *American journal of preventive medicine*, vol. 44, no. 1, pp. 71–75, 2013.

[50] B. D. B. Gonze, R. D. C. Padovani, M. D. S. Simoes, V. Lauria, N. L. Proença, E. F. Sperandio, T. L. V. D. P. Ostolin, G. A. D. O. Gomes, P. C. Castro, M. Romiti et al., “Use of a smartphone app to increase physical activity levels in insufficiently active adults: feasibility sequential multiple assignment randomized trial (smart),” *JMIR research protocols*, vol. 9, no. 10, p. e14322, 2020.

[51] J. Fanning, S. P. Mullen, E. McAuley et al., “Increasing physical activity with mobile devices: a meta-analysis,” *Journal of medical Internet research*, vol. 14, no. 6, p. e2171, 2012.

[52] A. Muntaner, J. Vidal-Conti, and P. Palou, “Increasing physical activity through mobile device interventions: a systematic review,” *Health informatics journal*, vol. 22, no. 3, pp. 451–469, 2016.

[53] J. R. Shapiro, S. Bauer, R. M. Hamer, H. Kordy, D. Ward, and C. M. Bulik, “Use of text messaging for monitoring sugar-sweetened beverages, physical activity, and screen time in children: a pilot study,” *Journal of nutrition education and behavior*, vol. 40, no. 6, pp. 385–391, 2008.