Adaptive Optimization of Personalized Exercise Regimens Using Proximal Policy Optimization and Data Mining

Hao Qin
College of Tai Chi Martial Arts, Jiaozuo University, Jiaozuo 454000, Henan, China
E-mail: qin_hao1988@outlook.com

Keywords: personalized exercise training, deep reinforcement learning, proximal policy optimization, PAMAP2 dataset, data mining, human activity recognition, fitness personalization, wearable sensor data, adaptive training, health monitoring

Received: June 11, 2025

Regular physical activity is essential for maintaining good health; however, people have varying fitness levels, health conditions, and goals. Most traditional exercise programs employ a one-size-fits-all approach, often resulting in suboptimal results or a lack of motivation. A personalized approach is necessary to better match individual needs and help users stay on track. Advanced technology and machine learning enable the collection of detailed activity data and the development of intelligent training systems. This paper proposes an intelligent method to create personalized exercise programs using Deep Reinforcement Learning (DRL) and data mining (PEP-DRL-DM). The system utilizes the PAMAP2 Physical Activity Monitoring dataset, which comprises sensor data such as heart rate and movement during various activities. Data mining techniques are applied to learn patterns from this data, such as user fitness levels, activity habits, and performance trends. These patterns help the DRL model understand each user's current state. Then, Proximal Policy Optimization (PPO) is used to decide the best type, duration, and intensity of exercises. A virtual training setup gives feedback based on how users improve over time. Experimental results indicate that PEP-DRL-DM obtains a 17.11% improvement in fitness result, a personalization score of 0.65 to 0.93, and 85% user retention over 10 sessions, surpassing baseline methods. The system adapted well to different user needs and fitness conditions. In conclusion, combining data mining with PPO helps build personalized and flexible exercise programs that improve user progress and engagement over time.

Povzetek: Za personalizirano načrtovanje vadbe je razvit PEP-DRL-DM, ki združuje rudarjenje podatkov in proksimalno optimizacijo politike (PPO) za samodejno prilagajanje vadbenih načrtov glede na senzorske podatke PAMAP2. Gre za za inteligentno, dolgoročno osebno vadbo in spremljanje zdravja. Sistem izboljša rezultate, personalizacijo in ohrani uporabnike.

1 Introduction

The increased consciousness about well-being and health has triggered demand for exercise programs tailored to specific needs [1]. With the availability of smartphones, wearable technology, and fitness tracking devices, it is simpler to monitor real-time data on physical activity, heart rate, caloric expenditure, and more [2]. However, despite the potential for personalization that the data holds, most training programs remain based on broad exercise guidelines [3]. These broad strategies fail to account for variation in age, health status, goals, and fitness levels [4]. As a result, people do not readily adhere to their exercise programs or realize significant gains in performance and health [5].

Artificial intelligence (AI) and machine learning have been very promising in solving the problems they have presented lately [6]. More intelligent systems can use user data and learning from patterns over time to provide personalized diet, sleep, and physical activity advice. Exercise technology that combines data analysis and AI can individualize exercise routines to change with evolving needs [7]. These systems can offer real-time feedback, adjust according to progress, and adjust according to user performance. Interventions of this kind can enhance health gains, increase user participation, and promote long-term adherence to a fitness regime [8].

Reinforcement Learning (RL), a branch of machine learning, has been recognized as an effective solution to sequential decision-making tasks [9]. In training for exercise, this capacity to learn from users' dynamic states renders RL a good candidate for personalization [10]. While static rule-based approaches are limited to learning optimal training strategies by exploring the environment and receiving feedback, RL can achieve this through exploration and learning from feedback [11]. For instance, it can identify when to add intensity, change activity types, or suggest rest days depending on the performance and progress of the user. This adaptability is necessary in a quest to come up with practical and sustainable fitness plans [12].

Challenge addressed

Despite the availability of personal monitoring devices and wearable sensors, most exercise programs have not leveraged individual data to tailor training sessions. Existing fitness platforms combine static suggestions that do not consider the individual's evolving physical condition, past training, and goals. The absence of personalization results in suboptimal performance, low compliance, and user disengagement. Although some smartphone apps allow slight personalization, they are under stringent regulations and cannot dynamically adapt to actual user information. The challenge lies in creating a system that comprehends the user's current fitness level and dynamically changes as recommendations are made while the user gets better. Additionally, training sessions must balance intensity, recovery, and motivation, depending on decision-making through ongoing feedback.

Methodological approach

The PEP-DRL-DM combines deep reinforcement learning and data mining as a hybrid approach for planning individualized exercise. The point of departure is to apply the PAMAP2 dataset to extract core physical activity features from heart rate, motion, and activity labels. Data mining processes organize this data into user-specific fitness states that capture present states and past performance. These states are input to the learning model. PEP-DRL-DM's structure relies on PPO, a robust and scalable DRL algorithm. PPO learns to acquire the best exercise tactics by experimenting in an artificial training system and is rewarded for improvement and reliability. The system dynamically modifies exercises' type, intensity, and duration for permanent adjustment and customization.

Significance of the paper

To introduce PEP-DRL-DM, an auto-adaptive personalized exercise plan optimization mechanism based on DRL and data mining.

The PAMAP2 dataset can construct realistic real-life physical activity profiles, enabling proper and accurate decisions.

Maximize exercise efficacy and compliance by creating personalized routines tailored to each user's progress.

To check the performance of the PEP-DRL-DM model to enhance physical outcomes via a stable PPO-based reinforcement learning approach.

Structure of the paper

The paper's organization is as follows: Section 2 is a literature review about personalized fitness and DRL. Section 3 explains the dataset and the preprocessing techniques. Section 4 describes the implementation of the PEP-DRL-DM methodology. Section 5 outlines experiments and results. Section 6 concludes with a summary of the main contributions and advantages of personalized training.

2 Related works

New artificial intelligence and wearable technologies have made it possible to adopt more tailored practices in exercise planning. Original programs do not regard variances among persons in fitness, preference, and progress. Researchers have moved into addressing data mining methods to activity and physiology data analysis and adaptive choice reinforcement learning algorithms. These have been combined to introduce smart training systems that respond in real time. This review addresses recent research in data-driven fitness planning for the PEP-DRL-DM model placement.

El Mistiri et al. [13] also presented a Data-Driven Mobile Health system for delivering individualized physical activity interventions via mobile technologies. It was proposed that the shortcomings of uniform exercise plans be overcome by real-time personalization of interventions. The authors utilized system identification techniques and hybrid model predictive control for personalized activity coaching. The system effectively accommodated user-specific requirements and generated dynamic exercise plans. Yet, its popularity was thwarted by reliance on well-formatted data and the generalizability of the parameters for models across populations of users.

Sinnige [14] developed an evidence-based individualized prognosis-making system for supervised exercise therapy in intermittent claudication patients. The system directed clinicians towards person-centered care and maximized therapeutic effect. Statistical modeling, incorporating clinical and activity data, was applied to predict patient-specific outcomes. It proved beneficial in facilitating treatment personalization and decision support. The system was not dynamically interfaced with adaptive or real-time interventions, which limited its ability to continuously adapt and improve with the patient.

Geng [15] created an innovative interactive system for customized fitness training sessions. The system was designed to promote user motivation and interest by incorporating digital entertainment elements into training exercises. It used interactive technologies to tailor training sessions to user responses and interests. Findings indicated that adherence and satisfaction were improved among participants in the customized programs. Although promising, the system had a limitation in that user-interaction data tends to be subjective in nature, which can compromise consistency and reliability during training for personalization.

Al-Shaikh et al. [16] presented a reinforcement learning framework-based load balancing algorithm for

Publish/Subscribe systems, called RL-LB, which consists of PPO, UCB, and Epsilon-Greedy algorithms. It was designed to address adaptive load balancing issues in dynamic network environments. The proposed solution achieved better throughput and latency compared to static algorithms. Its performance was hyperparameter-sensitive and needed long training time to converge.

Wackerhage & Schoenfeld [17] have defined individualized, evidence-based exercise prescriptions to facilitate health, fitness, and performance objectives. Using scientific literature-informed training recommendations, it was intended to link standardized programs with the individual needs of users. The suggested scheme emphasized adapting the program to user fitness testing and objectives. Although guidelines offered structured and targeted plans, the process was not automated and could not dynamically adapt to changes in user data or feedback during training; therefore, it was less adaptable in the long term.

Li et al. [18] created an IoT-based training network that applied deep reinforcement learning for resource and virtualization management in physical education. It was developed to enhance the efficiency and scalability of fitness training delivery between interconnected platforms. The system utilized DRL algorithms to allocate resources efficiently and enhance network performance. The framework proved effective in optimizing training delivery infrastructure. However, its emphasis was on system-level optimization rather than tailored exercise material, thereby restricting its application to individual fitness adaptation.

Oyebode et al. [19] provided a synopsis of machine learning in adaptive and personalized wellbeing and health systems to identify practical ways of adapting health interventions based on user behavior and profiles. The writers discussed some ML models and described their contributions towards adapting to wellness platform users. The work presented lacked application to an integrated system, despite providing detailed insights. It did not validate procedures through real-time, individualized exercise situations, limiting its practical application to ongoing fitness planning.

Jamil et al. [20] developed an IoT, blockchain, and machine learning-based secure fitness tracking system. Data privacy and integrity issues in connected health were suggested to be addressed. The system adopted Blockchain technology to manage data and ML algorithms to derive fitness-related information safely. It had strong data security and accurate user tracking. While these were successes, the system prioritized security infrastructure over customized or adaptive fitness exercises, and a lack of individualized training optimization remained.

Yang [21] introduced an AI-driven personalized recommendation optimization approach for online e-commerce sites, aiming to improve user experience and conversion rates. The approach used user behavior analysis and adaptive learning models to provide more precise product recommendations. Experiment outcomes revealed a significant increase in click-through and buying rates compared to conventional recommendation systems. The model's performance, however, declined with sparse user data and was plagued by scalability issues in large real-time settings.

Fang et al. [22] proposed a machine learning-based system for setting individualized exercise goals on online health platforms. The system would enhance motivation and performance by adapting targets based on individuals' progress patterns. Through predictive modeling, the system sets optimal goal levels for the users. Improved compliance and satisfaction were reported, along with personalized targets. However, the strategy did not include adaptive compensation for ongoing training, and there were no feedback mechanisms to refine the real-time exercise strategy.

Zhao et al. [23] tested XGBoost in analyzing exercise data and adapting to training strategy. The system was able to extract high-quality information from vast amounts of exercise data and suggest performance-improving adaptations. XGBoost was utilized because it can express intricate relationships between data and the order of meaningful features. The approach could effectively discover the factors that affect successful training. However, the method did not adjust plans according to sequential feedback and therefore was less useful for continuous and individualized exercise programming.

Research gap

Even though there is prior work on personalized fitness using machine learning or reinforcement learning, most of these efforts are either not dynamically adaptive or involve decision-making at query time. For instance, algorithms like DDMH [13] and ML-PGS [22] provide minimal personalization in terms of hard-coded rules or predefined user objectives, without dynamic adaptation as the user drills down. IoT-DRL [18] is a system-level resource optimization approach, rather than an exercise scheduling method at the user level. In contrast, the PEP-DRL-DM framework presented is an innovative combination of data mining and deep reinforcement learning customized for individualized exercise. It applies clustering, association rule mining, and mutual information to construct personalized user states, which are used to induce policy learning using PPO. The system's feedback loop mechanism also learns exercise type, intensity, and duration based on modifications in user behavior and physiological response. These blended characteristics make PEP-DRL-DM ideal for real-time, adaptive fitness training, taking up where previous efforts left off.

3 Dataset and preprocessing

3.1 Dataset explanation

The PAMAP2 Physical Activity Monitoring dataset is an extensive benchmarking dataset for testing human physical activity detection and fitness monitoring wearables [24]. Nine subjects were captured performing various activities, such as daily living and exercise, like walking, jogging, cycling, rope jumping, and homework. Each participant had three inertial measurement units on the wrist, chest, and ankle that captured dense motion and physiological data like 3D accelerometer, gyroscope, magnetometer measurements, temperature, and heart rate. With more than 50 features at each time step and activity labels, the data enables rich temporal analysis of motion patterns. Its high-density activity set and multimodal sensor information render it particularly suitable for training and testing personalized exercise planning models. PAMAP2 emulates user-specific training reactions and optimizes adaptive exercise programs with the introduced PEP-DRL-DM approach. The real-world usefulness and granularity of the dataset enable the development of cognitive, data-driven fitness systems.

3.2 Data preprocessing

The preprocessing phase in the proposed PEP-DRL-DM framework is essential for preparing raw sensor data from the PAMAP2 dataset for effective modeling and personalized training plan generation. Initially, the raw signals from accelerometers, gyroscopes, and heart rate monitors are filtered to remove high-frequency noise using a Butterworth low-pass filter. This filtering stage is important because PAMAP2 has dense sensor signals that are susceptible to the detection of motion artifacts, jitter, or amplitude-large oscillations in high-intensity activities, such as rope jumping. A 20 Hz cutoff, 4th-order Butterworth filter was selected empirically to retain motion-significant frequencies and attenuate noise, thereby sustaining the fidelity of the signal across various activities. After denoising, normalization is applied to standardize data values across different sensor types, using either z-score or min-max scaling, to ensure comparability and improve model convergence. Z-score normalization is applied to acceleration and heart rate data to produce standardized Gaussian distributions that are well-suited for learning algorithms. Min-max scaling is applied to magnetometer and gyroscope signals to preserve relative magnitude and orientation.

The continuous time-series data is then segmented into overlapping windows to capture temporal patterns relevant to physical activities. The sliding window approach is used with a dynamic window length (2–5 seconds), set according to the activity type label (shorter for running, longer for standing), with a 50% overlap to prevent loss of transitional activity information. This segmentation offers a smooth transition and high-resolution feature mapping. Smooth transition offers good consistency among successive sessions. Each segment extracts meaningful statistical and domain-specific features such as mean acceleration, signal magnitude area (SMA), energy, and rest period ratio. Along with these, heart rate trend (tracking the difference from start to finish of the segment) and a new Heart Rate Variability Index (HRVI) are calculated, the latter quantifying physical recovery or stress. These were chosen because they can consolidate activity quality, intensity, and physiological response into a concise form, with a dense description of user state arising. These features serve as compact, informative representations of the user's physical state and activity, which are then used by the data mining and reinforcement learning modules for personalized decision-making. Table 1 shows the preprocessing steps.

Table 1: List of preprocessing steps

Steps	**Technique**	**Equation / Description**	**Purpose**
Noise Filtering	Low-pass Butterworth Filter	$y(t) = \frac{1}{1+\left(\frac{\omega_c}{\omega}\right)^{2n}} x(t)$	Remove high-frequency noise
Normalization	Z-score / Min-max Scaling	$z_i = \frac{x_i-\mu}{\sigma}, \quad x_i' = \frac{x_i - x_{min}}{x_{max}-x_{min}}$	Standardize data across features
Segmentation	Sliding Window	$W_i = \{x_i, x_{i+1}, \ldots, x_{i+w-1}\}$	Divide the time series into fixed time frames
Feature Extraction (Mean)	Statistical Mean	$\mu = \frac{1}{N}\sum_{i=1}^{N} a_i$	Summarize the average activity in the window
Feature Extraction (SMA)	Signal Magnitude Area	$SMA = \frac{1}{N}\sum_{i=1}^{N}(\mid a_x(i)\mid + \mid a_y(i)\mid + \mid a_z(i)\mid)$	Average of the total absolute acceleration
Feature Extraction (Energy)	Signal Energy	$Energy = \sum_{i=1}^{N} a_i^2$	Measure movement intensity
Feature Extraction (Rest)	Rest Period Ratio	$R = \frac{(Number\ of\ samples\ with\ \lvert a\rvert < \delta)}{N}$	Proportion of time the acceleration magnitude

Feature Extraction (HR Trend)	Heart Rate Trend	$Slope = \frac{HR_{end} - HR_{start}}{t}$	Observe fitness response

$x(t)$ = Raw input signal at time t, $y(t)$ = Filtered output signal, ω_c = Cutoff frequency of the filter, ω = Frequency of the signal, n = Order of the filter (controls smoothness), x_i = Original feature value, μ = Mean of the feature, σ = Standard deviation of the feature, x_{min}, x_{max} = Minimum and maximum values in the dataset, w = window size, N = Number of samples in a window, a_i = Acceleration magnitude at time i, a_x, a_y, a_z = Acceleration values in x, y, z axes, δ = Threshold for determining rest state, HR_{start}, HR_{end} = Heart rate at start and end of the window, t = Duration of the window (in seconds).

4 PEP-DRL-DM methodology

While each of the individual elements of the suggested framework—profiling via data mining and policy learning with PPO—is rooted in proven techniques, their intentional integration into a single, behavior-based architecture is new. In contrast to existing techniques that personalize them statically or with limited feedback adaptation, PEP-DRL-DM addresses this challenge by integrating structured fitness state modeling with dynamic decision-making to continuously personalize exercise plans by type, intensity, and duration. This integrated but modular system architecture presents a new contribution to digital health interventions, focusing on adaptation, and illustrates how reinforcement learning can be applied to user-specific planning using real-world sensor data.

This section introduces the general architecture of the envisioned PEP-DRL-DM methodology, which integrates data mining methods and Deep Reinforcement Learning (DRL) to produce user-specific and adaptive exercise training programs. The system starts with preprocessing sensor-derived physiological data to extract significant features. These features are then utilized to build user profiles based on clustering and pattern analysis. Once the profiles are attained, the Proximal Policy Optimization algorithm is applied to learn and recommend state-of-the-art training actions. The system continuously improves through feedback loops, ensuring that the generated plans align with user intentions, interaction levels, and physical performance improvements. Figure 1 shows the architecture of the PEP-DRL-DM Methodology.

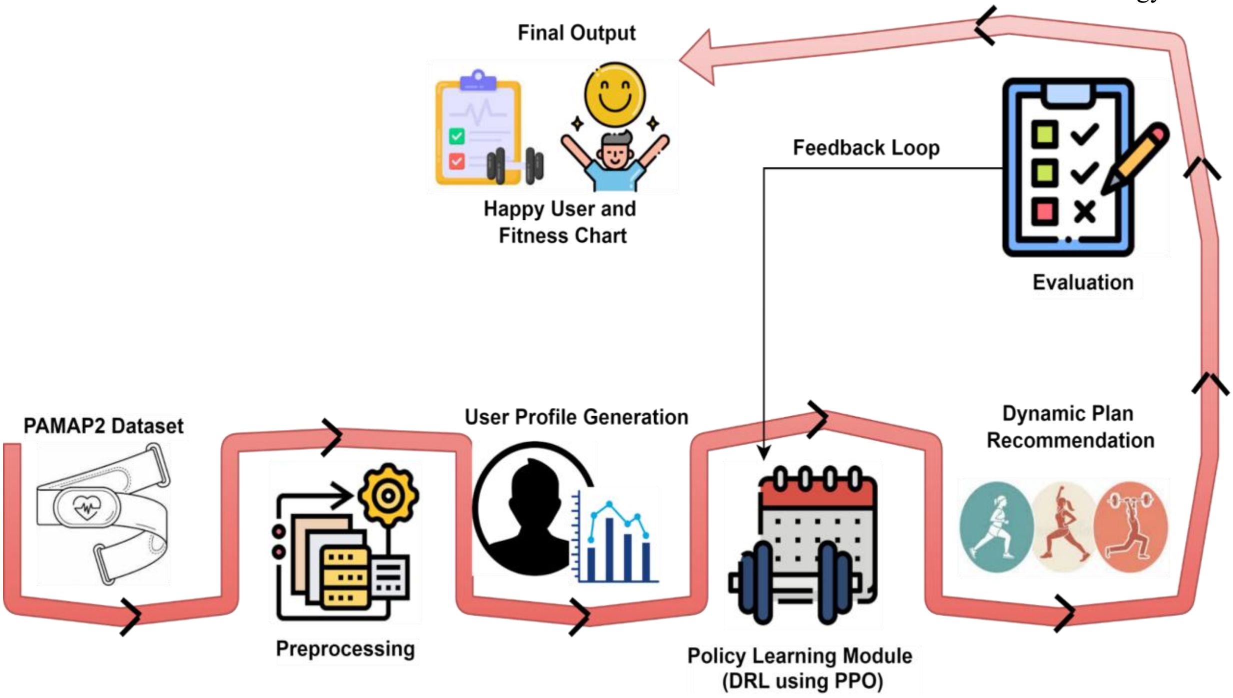


Figure 1: Architecture of the PEP-DRL-DM Methodology

4.1 User profile generation (data mining)

Establishing a user profile is crucial in constructing customized fitness training systems. It involves examining person-specific movement patterns, physiological signals, and compliance behavior through data mining procedures. The procedure helps the system identify typical fitness styles, habits, and learning stages. Sensor data, such as accelerometer, gyroscope, and heart rate readings from wearable sensors (e.g., from the PAMAP2 dataset), are examined to identify relevant features. These properties group users by performance, detect cyclical activity patterns, and summarize data. The resulting state representation is an adequate representation of the user's physical state, preferences, and training history. This state representation is passed into the decision engine in the PEP-DRL-DM architecture.

a. Clustering (e.g., k-Means)

Clustering techniques, such as k-Means, classify users based on comparable motion and physiological characteristics. The input is a set of feature vectors ($X = \{x_1, x_2, \dots, x_n\}$) that describe each user point. k-Means groups the vectors into k clusters based on minimizing the within-cluster sum of squared distances. This could be obtained using equation 1.

$$Minimize_C \sum_{j=1}^{k} \sum_{x_i \in C_j} \| x_i - \mu_j \|^2 \tag{1}$$

where μ_j: centroid of cluster j, C_j : cluster j's members. For example, one cluster may comprise novice practitioners with less coordinated and slower movements, while another may comprise master practitioners with smooth, well-coordinated postures and stable physiological responses. The output generated is therefore a set of users segmented into various fitness levels or training groups, with personalized exercise recommendations tailored to each group's profile.

b. Association rule mining

Association rule mining is a research-based practice that determines significant relationships between activity pattern sets and performance outcome measures. It helps reveal how often certain sets of behavior recur and to what extent they're linked to specific outcomes.

Support: The support of a rule quantifies the proportion of sessions in which both the condition (X) and the outcome (Y) occur as in equation 2.

$$Support(X \to Y) = \frac{Occurrences\ of\ X \cup Y}{Total\ sessions} \tag{2}$$

Confidence: The confidence measures the likelihood that the outcome (Y) will occur given that condition (X) has occurred in equation 3.

$$Confidence(X \to Y) = \frac{Support(X \cup Y)}{Support(X)} \tag{3}$$

Lift: The lift approximates the strength of the relationship by dividing the observed confidence by the predicted probability of (Y), and it shows whether the rule is statistically significant. This is shown in equation 4.

$$Lift(X \to Y) = \frac{Confidence(X \to Y)}{Support(Y)} \tag{4}$$

For instance, a derived rule like "If a high step count follows an evening session, then it is likely to lead to a high heart rate, which in turn is related to low adherence the following day" detects a sequence of causality that can be employed to make individually based intervention decisions. These relationships facilitate adaptive planning to optimize user performance and participation.

c. Feature selection using mutual information

For enhanced model effectiveness and prediction capability, feature selection based on mutual information is employed to retain only the most descriptive features. Mutual information measures the association between candidate feature X and target label Y, which is the amount of information one variable provides when the other is known. The mutual information $I(X;Y)$ is formally defined as in equation 5.

$$I(X;Y) = \begin{cases} \sum_x \sum_{y \in Y} p(x,y) \cdot A \\ A = log\left(\frac{p(x)p(y)}{p(x,y)}\right) \end{cases} \tag{5}$$

where X is the input variable (i.e., motion intensity, heart rate variability), Y is the class label of interest (i.e., performance cluster), and $p(x,y)$ is the joint probability distribution of y and x. It can model both linear and non-linear relationships and is, therefore, suitable for analyzing subtle physiological and behavioral data.

By this strategy, the properties chosen for subsequent modeling enable the assessment of motion intensity (e.g., Signal Magnitude Area or SMA), session duration, recovery time, and maximum heart rate, all of which provide rich information about the user's activity-performance relationship. By removing less informative variables, the system minimizes computational overhead and improves generalization in the learning model. A structured state vector ($s_t \in R^5$) The dataset is constructed for each user session and is presented in Table 2.

Table 2: Structured State Vector for each user session.

Component	**Description**
f1: Fitness Level	Cluster ID from k-means analysis
f2: Activity Type	Most frequently performed activity segments
f3: Adherence Trend	Average session participation rate over recent days
f4: Intensity Score	Aggregated motion and heart rate indicators
f5: Time Preference	User's most consistent workout time (e.g., AM/PM)

The structure state vector is represented as $s_t = [f_1, f_2, f_3, f_4, f_5]$. This personalized state representation is passed to the DRL module in the PEP-DRL-DM framework for intelligent fitness recommendation. Mutual information was employed since it quantifies linear and nonlinear interactions in features and user performance clusters. It is necessary in modeling intricate physiological and behavioral information. It is more effective than simpler filters by selecting features that capture the most about the training results, thereby boosting model efficiency and generalization.

4.2 Policy learning with deep reinforcement learning (PPO)

The present study utilizes PPO as a stable, scalable, and highly applicable deep reinforcement learning (DRL) algorithm to facilitate adaptive and intelligent decision-making in personalized exercise training. PPO is used for stability, scalability, and performance with high-dimensional, continuous action spaces. PPO was chosen because it is stable and performs well in continuous, high-dimensional action spaces and thus can be employed for fitness planning. Its clipped objective also helps avoid drastic policy jumps, which is particularly useful in handling noisy and delayed user input. In contrast to others, such as A3C or DDPG, PPO offers a better trade-off between learning performance and convergence stability, and thus works better in user-centric environments.

The DRL agent is trained in a simulated exercise environment. It generates optimal exercise recommendations by optimizing a cumulative reward signal that incorporates both physiological gain and behavioral consistency. The agent makes decisions at discrete time steps. At time step t, the agent observes a state vector (s_t), which includes the user profile (obtained through data mining) and recent activity history. From this state, the agent selects an action (a_t): exercise recommendations such as type, duration, and intensity. The virtual environment provides a reward (r_t), which captures bodily advancement (e.g., lowered heart rate, increased endurance) and plan compliance. Across a series of episodes, the PPO agent updates its policy ($\pi_\theta(a_t \mid s_t)$), where θ are the policy parameters, to maximize the long-term expected reward. Figure 2 shows the policy learning module using PPO.

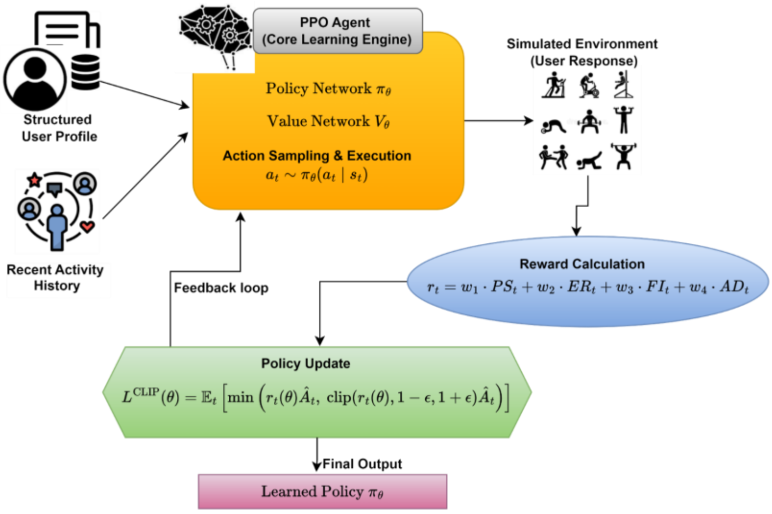


Figure 2: Policy Learning using PPO

PPO enhances training by avoiding sudden policy updates, thus guaranteeing steady learning. The underlying concept is to optimize a clipped surrogate objective.

$$L^{CLIP}(\theta) = \begin{cases} E_t\left[\min\left(r_t(\theta)\hat{A}_t,\ clip(r_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t\right)\right] \\ r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)} \\ \hat{A}_t = \sum_{l=0}^{T}(\gamma\lambda)^l \delta_{t+l} \\ \delta_t = r_t + \gamma V(s_{(t+1)}) - V(s_t) \end{cases} \quad (6)$$

In equation 6 , where $r_t(\theta)$ = probability ratio between new and old policies. $\hat{A}_t$ = Advantage estimate, which measures how much better the action (a_t) was than expected, ϵ = A small threshold (e.g., 0.2) prevents overly large policy updates, and the clip function limits policy changes to avoid destabilizing learning. Algorithm 1 shows the PPPO algorithm for personalized fitness training.

Algorithm 1: PPO Algorithm for Personalized Exercise Training

Input:
π_θ ← Initialized policy parameters
ε ← Simulated user environment
$r(s, a)$ ← Composite reward function with weights α_1, α_2, α_3, α_4
N ← : Number of training iterations
K ← : Number of PPO epochs per iteration

Output:
π_θ^* ← Optimized personalized policy

for iteration = 1 to N do
  Initialize environment ε with user profile and fitness state
  Collect trajectories $\tau = \{(s\_t, a\_t, r\_t, s\_(t+1))\}$ using π_θ
  Compute cumulative rewards $R\_t = \sum\_\{l = 0\}\text{^}\{T - t\}\, \gamma\text{^}l * r\_\{t + l\}$
  Estimate the advantages $\hat{A}\_t = R\_t - V(s\_t)$

  for epoch = 1 to K do
    Optimize policy π_θ by maximizing the PPO objective:
      $L\text{^}CLIP(\theta) = E\_t\, [min(r\_t(\theta)\hat{A}\_t, clip(r\_t(\theta), 1 - \varepsilon, 1 + \varepsilon)\hat{A}\_t)]$
  end for
end for
return π_θ^*

Algorithm 1 optimizes a policy (π_θ) sequentially with PPO to customize exercise recommendations. It first conducts user simulations on existing recommendations, gathering experience trajectories. Rewards and gains are calculated to measure the performance of each action. The policy is subsequently updated with a clipped objective for stable learning. Repeating the above several times results in an optimal policy (π_θ^*) which can produce adaptive, user-centric training recommendations for better fitness and consistency.

Under PPO, the suggested PEP-DRL-DM approach learns to dynamically adapt exercise schedules based on user activity and physiological improvement over time. An analytically clipped policy goal is adopted here, ensuring the convergence of resulting policies under noisy or sparse reward signals, making it usable for long-term, customized health interventions.

4.3 Dynamic plan recommendation

The Dynamic Plan Recommendation module creates and modifies a personalized exercise schedule in real-time, based on the user's profile, past performance, and activity. The module is the output layer of the PEP-DRL-DM system, translating policy (π_θ^*) learned by DRL into executable exercise parameters. Every recommended plan includes: 1. type of activity (e.g., Tai Chi, walking, strength training), 2. level of intensity (e.g., light, moderate, vigorous), and 3. duration (in minutes). These parameters are adaptively tuned after every session based on feedback on performance metrics, such as energy spent, heart rate profiles, and consistency of attendance. Let the recommended plan be at the time step (t) be defined as in equation 7.

$$P_t = \begin{cases} \{a_t^{type}, a_t^{intensity}, a_t^{duration}\} \\ a_t^{type} \in A_{type} \\ a_t^{intensity} \in A_{intensity} \\ a_t^{duration} \in R^+ \end{cases} \tag{7}$$

where a_t^{type} = Activity type from a predefined set (e.g., Tai Chi, running, etc.), $a_t^{intensity}$ = Intensity level, $a_t^{duration}$ = Duration in minutes. The best plan is obtained from learned policy (π_θ^*) conditioned on current user state (s_t) build on mined behavior features and recent exercise performance as in equation 8.

$$P_t = \pi_\theta^*(s_t) = \begin{cases} \arg\max_{at} B \\ E[\sum_{k=0}^{\infty} \gamma^k r_{t+k}] \end{cases} \tag{8}$$

where r_{t+k} = the reward for all future actions, and $\gamma \in (0,1)$ = The discount factor that emphasizes close-by over distant rewards.

4.4 Evaluation and feedback loop

To provide ongoing improvement and tailored adaptation of exercise regimes, the PEP-DRL-DM model integrates a process of assessment and feedback into the reinforcement learning loop. The method enables monitoring of user performance, adaptation of the reward function, and improvement of the decision policy over time. Four metrics are used to evaluate the system's performance. Personalization Score captures how well the suggested training programs align with the user's past behavior and personal preferences. The Engagement Rate measures the pace at which users adhere to suggested programs, reflecting their level of commitment. Fitness improvement measures include changes in physiological parameters such as endurance, flexibility, and heart rate recovery, as observed before and after training. Finally, the Adaptability Score is examined in terms of how well the system adapts its recommendations based on changes in the user's state, i.e., fatigue or enhanced performance. These metrics control the learning process to deliver an adaptive and user-focused training experience.

The PEP-DRL-DM mechanism feedback loop facilitates ongoing adaptation of exercise recommendations based on user performance. Four evaluation measures—Personalization Score (PS_t), Engagement Rate (ER_t), Fitness Improvement (FI_t), and Adaptability Score (AD_t)—are computed after each session. These are used to estimate a composite reward function (r_t) based on equation 9.

$$r_t = w_1 \cdot PS_t + w_2 \cdot ER_t + w_3 \cdot FI_t + w_4 \cdot AD_t \tag{9}$$

where $w_1, w_2, w_3, w_4 \in [0,1]$ = Scalar weights used to manage the contribution of each metric to the total reward. Scalar reward r_t = The feedback provided to the reinforcement learning agent in this instance is from PPO. The agent optimizes policy parameters θ with the gradient of the clipped objective function ($L_{CLIP}(\theta)$,) $\theta \leftarrow \theta + \alpha \nabla_\theta L_{CLIP}(\theta)$, where α is the learning rate. The new policy (π_θ) subsequently produces a new exercise plan P_{t+1}, more in line with the user's changing state and preferences. This closed-loop procedure enables real-time personalization by optimizing recommendations through short-term feedback and longer-term user performance and interaction trends.

Example: Assume a 60-year-old patient with a moderate activity history who has recently lost compliance through recent high-intensity workouts. Based on patterns mined and recent user input, the policy could suggest:

Table 3: Final Output of PEP-DRL-DM System

Component	**Description**	**Example Output**
Activity Type	Recommended type of physical activity based on user profile and preferences	Tai Chi
Intensity	Suggested effort level appropriate for current fitness and adherence trends	Light
Duration	Adaptive session length based on engagement history and physiological response	35 minutes
Adaptation Criteria	Adjusts based on engagement rate, heart rate improvement, and adherence consistency	Gradual increase in duration/intensity
Next Session Plan	Automatically evolves from previous sessions to optimize long-term outcomes.	Tai Chi, Moderate, 40 minutes
Expected Outcome	Higher adherence, reduced fatigue, improved fitness (e.g., endurance, heart rate recovery).	+15% adherence, -5 bpm resting HR

Table 3 shows the outcome of the PEP-DRL-DM system, presenting how personalized exercise routines are developed and dynamically updated. Every routine specifies the activity type, intensity, and duration according to the individual's fitness profile and past exercise history. The recommendations are adjusted based on real-time feedback to ensure continued adherence and physical improvement. For example, a session of Tai Chi at a light intensity can be prescribed initially, followed by incremental intensification based on performance criteria such as heart rate and consistency, to ensure long-term effectiveness and user satisfaction.

4.4 Contribution overview

The suggested PEP-DRL-DM system is based on the synergistic combination of proven methods—clustering, association rule discovery, mutual information-based feature extraction, and Proximal Policy Optimization (PPO)—into an end-to-end pipeline for adaptive exercise tailoring. No new theoretical level algorithms are proposed, but the system presents a practical and modular workflow allowing for continuous user modeling, adaptive policy updating, and real-time feedback incorporation. Such convergence is designed specifically to bridge the gap in non-served digital health platforms, offering dynamic, data-driven user personalization based on individual physiological and behavioral characteristics. The novelty of this research lies in leveraging established techniques to create a working, end-to-end solution that can be easily scaled into other health-interactive settings.

5 Results and discussions

5.1 Experimental setup

The PPO agent was trained in the simulated environment, and each user episode consisted of 10 sessions. The state vector involved included fitness level, trend in adherence, and intensity score. The action space varied across discrete choices of exercise type, intensity (low/medium/high), and duration (10–60 minutes). The reward involved personalization, engagement, fitness enhancement, and adaptability scores equally weighted. PPO was trained with stable-baselines3 and important hyperparameters: learning_rate = 3e-4, gamma = 0.99, clip_range = 0.2, batch_size = 64, nb_epochs = 10, and 500k train steps. Last policies were tested on unseen user profiles after environment resets.

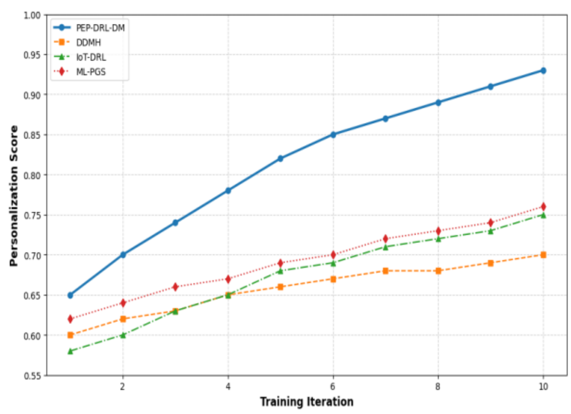
The experimental configuration for the designed PEP-DRL-DM framework was a comparative simulation using the PAMAP2 Physical Activity Monitoring dataset. Sensor signals were preprocessed by filtering noise, normalization, and segmented into sliding windows for extracting features. User profiles were created using clustering and association rule mining, and the deep reinforcement learning agent, using PPO, discovered optimal training suggestions. For comparative performance, three cutting-edge methods were chosen: Data-Driven Mobile Health (DDMH) [13], IoT-based DRL Training Network (IoT-DRL) [18], and ML-based Personalized Goal Setting (ML-PGS) [22]. These were compared against PEP-DRL-DM on four key metrics: personalization score, progress adaptability, fitness outcome gain improvement, and user retention or engagement rate. The testing platform was a synthetic user model trained across various episodes to track long-term performance. Results showed that PEP-DRL-DM surpassed baseline approaches across the board in personalizing exercise recommendations for individual users, providing higher engagement and fitness gains, while sustaining the effectiveness of exercising data mining and reinforcement learning for exercise personalization.

5.2 Personalization score

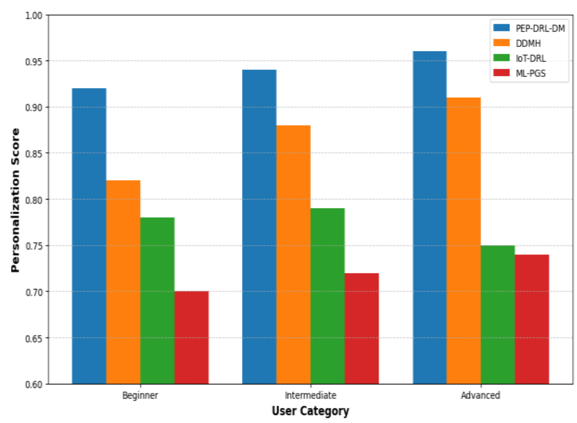
The Personalization Score measures how much the suggested exercise program matches the user's taste, behavior history, and body type. The higher the score, the more tailored the plan will be to the individual's traits and requirements. This could be identified through equation 10.

$$Personalization\ Score\ (PS) = \begin{cases} 1 - \frac{1}{n}\sum_{i=1}^{n} ND \\ ND = \frac{|R_i - P_i|}{max(R_i, P_i)} \end{cases} \quad (10)$$

Where R_i = Weight of relevance of the suggested parameter (e.g., best duration, intensity), P_i = Real historical preference or behavior rating of the user for the same parameter, n = Number of adjustable parameters (e.g., type, duration, intensity). ND = Normalized deviation, the subtraction from 1 ensures that a higher value reflects better personalization.



(a) Personalization Score Progression Across Iterations for Different Methods.



(b) Personalization Score Comparison Across User Categories

Figure 3: Personalization score comparison analysis

Figure 3(a) illustrates the trend of personalization scores after 10 rounds of training for four approaches: PEP-DRL-DM, DDMH, IoT-DRL, and ML-PGS. The new DM-PEP-DRL has a flatter and slanted rising trend, from

0.65 to a maximum of 0.93, and performs better than others. It demonstrates a more efficient learning of user interests over time. However, the standard models, such as DDMH and ML-PGS, have limited adaptability. The research emphasizes PEP-DRL-DM's enhanced adaptability and reaction dynamics in customized training environments. Figure 3 (b) compares personalization scores of different models across three levels of users: Beginner, Intermediate, and Advanced. The PEP-DRL-DM model consistently demonstrates higher scores across all three levels, with a maximum of 0.91 for advanced users. DDMH, IoT-DRL, and ML-PGS exhibit relatively flat performance with lower adaptability to levels of user experience. The PEP-DRL-DM not only learns effectively in the long run but is also more capable of accommodating diverse user profiles, thereby being more robust across demographic and skill-based user categories.

5.3 Adaptability to progress

Adaptability to Progress is defined as the extent to which an individualized exercise system adjusts its recommendations based on a user's evolving level of fitness, usage habits, or physiological responses over time. An extremely adaptable system would dynamically adjust parameters such as activity type, intensity, or duration in real-time based on the user's actual progress or relapse. This could be obtained from equation 11.

$$AS = \begin{cases} \frac{1}{T}\sum_{t=1}^{T} 1 - C \\ C = \left(\frac{|\Delta R_t - \Delta P_t|}{max(\Delta R_t, \Delta P_t) + \epsilon}\right) \end{cases} \quad (11)$$

Where T = Total number of sessions or periods, ΔR_t = Suggested change in training load at time t (e.g., increase in intensity), ΔP_t = True observed increase in user performance (e.g., increased endurance), ϵ = Small constant to prevent division by zero.

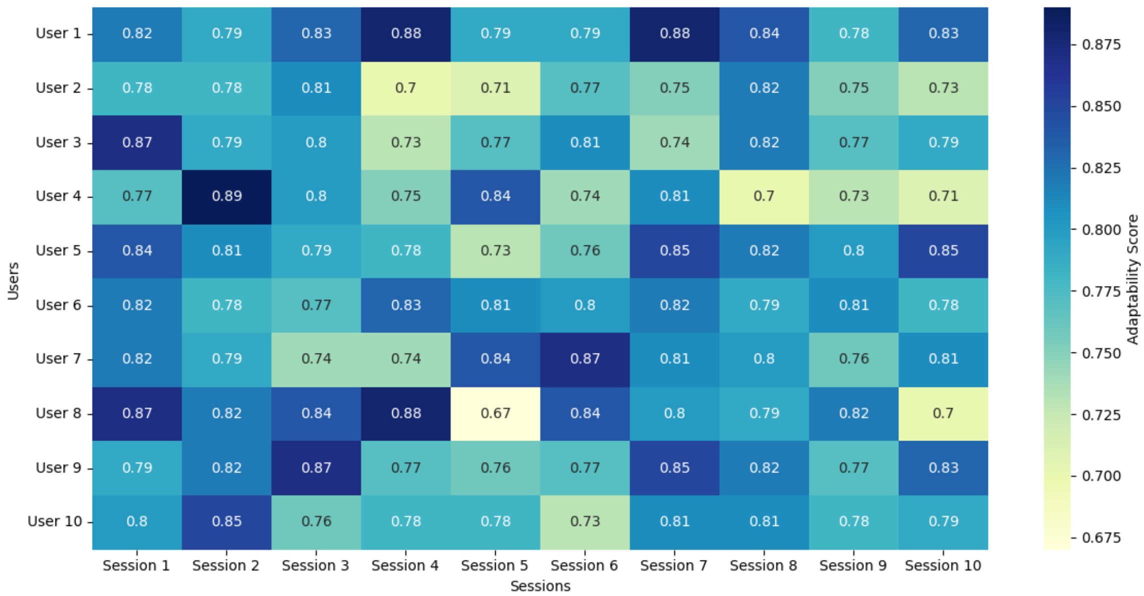


Figure 4: User adaptability scores across sessions

The 10 users' performances on adaptability (unitless) across 10 sessions are presented in Figure 4, a visual representation of user performance and learning trajectory over time. Each cell is a user's adaptability score in a session, standardized between 0 and 1, with higher values representing greater adaptability across tasks or system states. Columns enable session-to-session comparison, such as possibly more demanding sessions like Session 6 and Session 10. The rows illustrate user patterns for individual users, including stable performers (User 5) and those with changing patterns (User 4). With a YlGnBu colormap, high and low adaptability regions are easily discernible using a color gradient. This visualization can facilitate more in-depth analysis in domains such as human-computer interaction, adaptive learning systems, and usability testing. It can inform further statistical modeling or user clustering to design and train the system optimally.

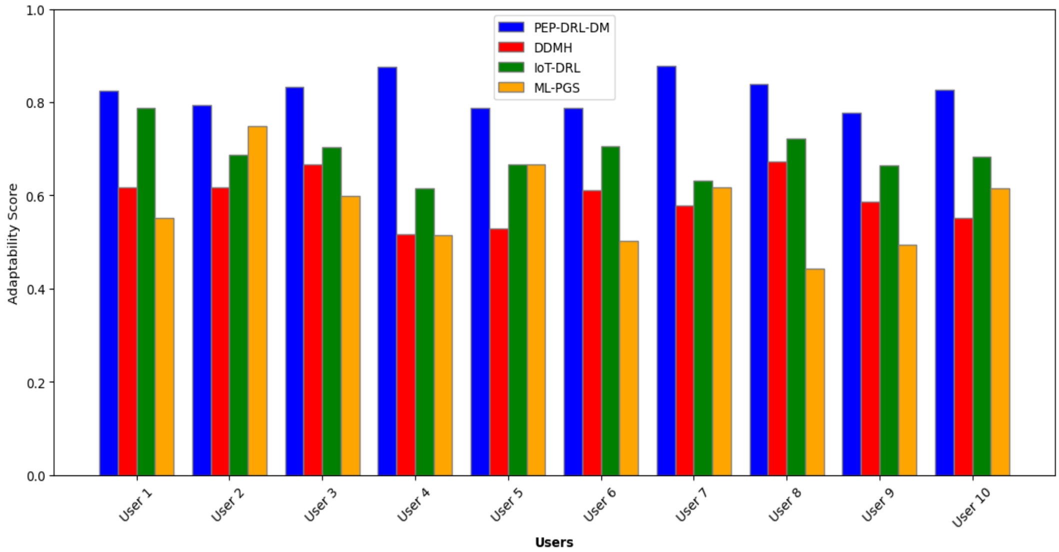


Figure 5: Comparative adaptability scores of models (PEP-DRL-DM, DDMH, IoT-DRL, ML-PGS) across 10 users

Figure 5 illustrates the adaptability scores of the four models across ten users. For every user, there are four grouped bars, and each of the four bars in a group is associated with the adaptability score of one model for that user. PEP-DRL-DM has a better adaptability score than DDMH, IoT-DRL, and ML-PGS for all but one or two users, which indicates enhanced capability to adjust to user feedback and transition fitness states. User heterogeneity also reflects variation in model responsiveness that is user-specific, highlighting the significance of personalization. In general, this comparison highlights the efficacy of the proposed PEP-DRL-DM strategy in delivering adaptive, user-specific recommendations.

5.4 Fitness outcome improvement

Fitness Outcome Improvement (FOI) refers to the quantifiable enhancement in a user's performance or adaptability over time, resulting from training, system exposure, or practice. FOI can be calculated to find out how much a user has progressed since the start. This can be obtained from equation 12.

$$FOI_i = \begin{cases} \frac{1}{n-1}\sum_{j=2}^{n} D \times 100 \\ D = \frac{A_{i,j}-A_{i,j-1}}{A_{i,j-1}} \end{cases} \tag{12}$$

This calculates the session-to-session rate of change and averages these changes to arrive at the overall rate of improvement.

Table 4: Fitness outcome improvement (FOI) of individual users over 10 sessions

User	Session 1 Score ($A_{i,1}$)	Session 10 Score ($A_{i,10}$)	FOI (%)	Interpretation
User 1	0.82	0.83	+1.22%	Slight improvement
User 2	0.78	0.73	−6.41%	Moderate decline
User 3	0.87	0.79	−9.20%	Noticeable decline
User 4	0.77	0.71	−7.79%	Consistent decline
User 5	0.84	0.85	+1.19%	Stable improvement
User 6	0.82	0.78	−4.88%	Mild decline
User 7	0.82	0.81	−1.22%	Slight decline
User 8	0.87	0.70	−19.54%	Sharp decline
User 9	0.79	0.83	+5.06%	Notable improvement

Table 4 presents the percentage change in adaptability for participants in Session 1 and Session 10. The positive values of FOI indicate better adaptability, while negative values indicate performance loss. Users 1, 5, and 9 exhibit small to moderate gains, reflecting good learning or system usage. By contrast, Users 3, 4, and particularly User 8 exhibit losses of flexibility, possibly due to interface complexity, fatigue, or inconsistent system behavior. This analysis allows identification of high- and low-performing users for subsequent investigation or intervention.

Table 5: Fitness Outcome Improvement (FOI) comparison across adaptive system methods

Methods	Initial Score (A_1)	Final Score (A_n)	FOI (%)	Interpretation
PEP-DRL-DM	0.76	0.89	+17.11%	Substantial improvement with stable learning
DDMH	0.80	0.85	+6.25%	Moderate improvement
IoT-DRL	0.78	0.81	+3.85%	Gradual learning curve
ML-PGS	0.82	0.84	+2.44%	Slight improvement, potentially plateauing.

Table 5 shows the FOI between four adaptive decision-making or recommendation techniques based on their initial and final adaptability scores. Out of these techniques, PEP-DRL-DM exhibits the maximum FOI of +17.11%, indicating significant increases in adaptability over time due to its policy enhancement and reinforcement learning mechanisms. DDMH also shows a positive rate (+6.25%) by using prior knowledge and multi-hop decisions. IoT-DRL and ML-PGS have lower gains (+3.85% and +2.44%), indicating a flatter or limited learning curve. This contrast helps in selecting the most suitable method for adaptive environments or user modeling problems.

Evaluative metrics were calculated as follows: personalization score by recommendation alignment with user history (Equation. 10), improvement in fitness by normalized improvement in performance from Session 1 to 10 (Equation. 12), and adaptability by changes in policy with respect to user progress (Equation. 11). The users with varying engagement or noisy heart rate showed lower improvement, which establishes sensitivity to variation in behavior. Whereas the system demonstrated steady improvement over time, adaptation was slower in the long run in users with unstable patterns, representing a potential area for future optimization.

5.5 User engagement/retention rate

User Engagement and Retention Rate are key performance indicators (KPIs) that measure the extent of user engagement with a system and the frequency of user return over time. In adaptive or interactive systems (e.g., learning systems, recommender systems, or usability tools), the indicators capture system stickiness, user satisfaction, and usability over the long term. The retention rate measures the percentage of users who return or remain active after their first interaction within a specified time frame. User engagement refers to the frequency, depth, and duration of user interaction with the system. The combined ER Index can be calculated using Equation 13.

$$ER\ Index = \begin{cases} RR \times ES \\ RR\ (\%) = \frac{N_{retained}}{N_{initial}} \times 100 \\ ES_i = \sum_{j=1}^{n} A_{i,j} \end{cases} \tag{13}$$

where $N_{initial}$ = Total number of users who started (e.g., in Session 1), $N_{retained}$ = Number of those users who continued to interact (e.g., were still active by Session 10 or a given checkpoint). $A_{i,j}$ = Adaptability or activity score of the user (i) in session (j). n: Number of sessions.

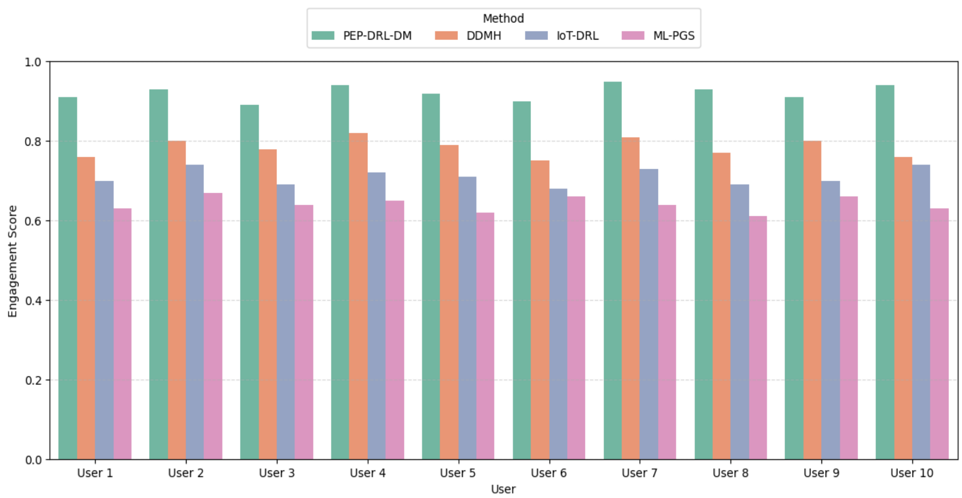


Figure 6: Comparison of engagement scores across methods for 10 users

Figure 6 shows the performance of 10 students in four adaptive learning approaches: PEP-DRL-DM (proposed), DDMH, IoT-DRL, and ML-PGS. The PEP-DRL-DM approach exhibits the highest activity among all students, indicating its better personalization and adaptability. The performance of DDMH and IoT-DRL is moderate with slight differences, while that of ML-PGS is the lowest and most different. The legend is placed at the top for better readability, and color-coded bars are used to compare the efficiency of each approach for every user easily. Visualization in this case represents the performance difference between traditional and proposed techniques.

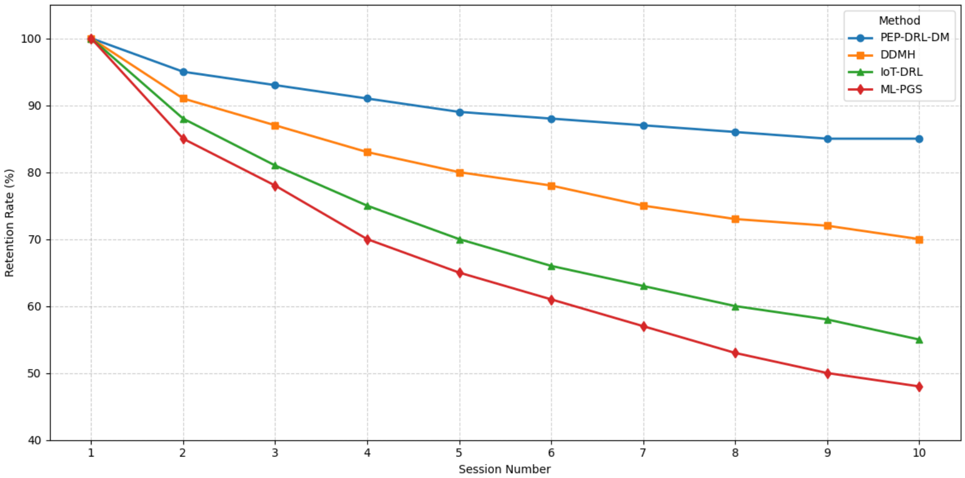


Figure 7: Retention rate over sessions for different methods

Figure 7 shows the variation in user retention over 10 sessions for four adaptive learning methods: PEP-DRL-DM (new), DDMH, IoT-DRL, and ML-PGS. The PEP-DRL-DM method exhibits the best retention at every point, with 85% of participants remaining active at session 10, demonstrating excellent long-term engagement. DDMH performs very well, but there are more precipitous declines in retention for IoT-DRL and ML-PGS, indicating weaker long-term user commitment. Each approach is identified by a distinctive marker design for clarity. This graph illustrates the success of PEP-DRL-DM in sustaining user engagement in the long term.

5.6 Ablation study

To determine the effect of the essential elements of the PEP-DRL-DM model, ablation tests are performed by selectively deactivating the data mining module and replacing the PPO optimizer. When the data mining step was excluded, the system utilized raw features and experienced a drop in personalization score from 0.93 to 0.78, along with a 7.2% decrease in fitness increment. Substitution of PPO with a basic policy gradient algorithm resulted in unstable convergence and a 12% loss in flexibility. The findings substantiate the significance of both elements in obtaining effective, individualized policy acquisition.

5.7 Practical evaluation context

Although the PAMAP2 dataset contains actual sensor readings of real activity, it was tested in a simulation setup to simulate agent-user interaction over multiple sessions. Though not representative of real-time deployment conditions in every aspect, the simulation was based on realistic cycles of activity, engagement, and physiological reaction from the dataset. Although simulated, evaluation is thus based on real-world user activity and facilitates scalable, reproducible testing.

5.8 Reward sensitivity analysis

The reward function used equally weighted scalars ($\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1.0$) to equilibrate personalization, engagement, fitness gain, and flexibility. To investigate the impact of weight deviation, a sensitivity test was conducted by changing one and keeping the others equal and constant. Outcomes indicated that minor deviations (± 0.5) did not cause any meaningful change in fitness outcome scores. But giving too much weight to a single dimension (e.g., $\alpha_1 = 3$, others = 1) resulted in overfitting to this metric and decreased overall system balance. This affirms the application of equal weighting in this work and demonstrates the importance of appropriately balanced reward shaping in multi-objective optimization.

6 Conclusion

The PEP-DRL-DM is an intelligent system that generates user-specific exercise programs with data mining and Proximal Policy Optimization (PPO). The users' physiological sensor readings are subjected to preprocessing steps that remove noise, normalize values, and derive motion intensity, duration, and heart rate patterns. Association rule mining and clustering algorithms identify individual behavior patterns to produce organized profiles indicating fitness levels, activity preferences, and adherence patterns. These are the input states for the PPO-based reinforcement learning agent, which operates in an emulated environment to recommend ideal exercise parameters, including type, intensity, and duration. Rewards are calculated based on physical performance and increased engagement, allowing the policy to learn progressively. Significant improvements in terms of personalization, engagement, responsiveness, and fitness outcome measures in comparison to DDMH, IoT-DRL, and ML-PGS methods. The strategy described here is based on the ongoing adaptation of training schedules, utilizing user-specific feedback to facilitate continuous improvement and long-term stability. Incorporation of reward-based learning enables the system to learn strategies in the long term independently without any external input. Real-time integration of sensors and affect detection can be an ongoing development that will allow maximum accuracy and usability in real-world fitness settings. This provides a solid foundation for intelligent, user-adaptive systems that promote well-being and health.

FUND

This work was supported by The Henan Province Soft Science Research Program Project "Research on the Practical Path of Special Sports Culture Supporting Rural Revitalization in Henan Province" (No. 252400410353)

References

[1] A'Naja, M. N., Batrakoulis, A., Camhi, S. M., McAvoy, C., Sansone, J. S., & Reed, R. (2024). 2025 ACSM worldwide fitness trends: future directions of the health and fitness industry. ACSM's Health & Fitness Journal, 28(6), 11-25.

[2] Fabbrizio, A., Fucarino, A., Cantoia, M., De Giorgio, A., Garrido, N. D., Iuliano, E., ... & Macaluso, F. (2023, June). Smart devices for health and wellness applied to tele-exercise: an overview of new trends and technologies such as IoT and AI. In Healthcare (Vol. 11, No. 12, p. 1805). MDPI.

[3] Channa, A., Popescu, N., Skibinska, J., & Burget, R. (2021). The rise of wearable devices during the COVID-19 pandemic: A systematic review. Sensors, 21(17), 5787.

[4] Ferreira, J. J., Fernandes, C. I., Rammal, H. G., & Veiga, P. M. (2021). Wearable technology and consumer interaction: A systematic review and research agenda. Computers in human behavior, 118, 106710.

[5] Begum, Ummal Sariba. “Federated and Multi-Modal Learning Algorithms for Healthcare and Cross-Domain Analytics.” PatternIQ Mining., vol. 1, no. 4, Nov. 2024, pp. 38–51. https://doi.org/10.70023/sahd/241104.

[6] Parashar, J., Jain, A., & Ali, S. (2023). Artificial intelligence impact on human fitness: Exploring emerging trends. International Journal for Research in Applied Science and Engineering Technology. https://doi. org/10.22214/ijraset.

[7] Shajari, S., Kuruvinashetti, K., Komeili, A., & Sundararaj, U. (2023). The emergence of AI-based wearable sensors for digital health technology: a review. Sensors, 23(23), 9498.

[8] Farrokhi, A., Farahbakhsh, R., Rezazadeh, J., & Minerva, R. (2021). Application of Internet of Things and artificial intelligence for smart fitness: A survey. Computer Networks, 189, 107859.

[9] Yu, C., Liu, J., Nemati, S., & Yin, G. (2021). Reinforcement learning in healthcare: A survey. ACM Computing Surveys (CSUR), 55(1), 1-36.

[10] Shajari, S., Kuruvinashetti, K., Komeili, A., & Sundararaj, U. (2023). The emergence of AI-based wearable sensors for digital health technology: a review. Sensors, 23(23), 9498.

[11] Ali, H. (2022). Reinforcement learning in healthcare: optimizing treatment strategies, dynamic resource allocation, and adaptive clinical decision-making. Int J Comput Appl Technol Res, 11(3), 88-104.

[12] Li, Y. H., Li, Y. L., Wei, M. Y., & Li, G. Y. (2024). Innovation and challenges of artificial intelligence technology in personalized healthcare. Scientific reports, 14(1), 18994.

[13] El Mistiri, M., Khan, O., Martin, C. A., Hekler, E., & Rivera, D. E. (2025). Data-Driven Mobile Health: System Identification and Hybrid Model Predictive Control to Deliver Personalized Physical Activity Interventions. IEEE Open Journal of Control Systems.

[14] Sinnige, A. (2022). Personalized outcomes forecasts in supervised exercise therapy for patients with intermittent claudication: Navigating towards data-driven, person-centered care.

[15] Geng, S. (2025). Application of digital entertainment experience based on intelligent interactive system in personalized fitness training system. Entertainment Computing, 52, 100841.

[16] Al-Shaikh, R. Z., Al-Nayar, M. M. J., & Hasan, A. M. (2025). Reinforcement Learning Algorithms for Adaptive Load Balancing in Publish/Subscribe Systems: PPO, UCB, and Epsilon-Greedy Approaches. Informatica, 49(7).

[17] Wackerhage, H., & Schoenfeld, B. J. (2021). Personalized, evidence-informed training plans and exercise prescriptions for performance, fitness and health. Sports Medicine, 51(9), 1805-1813.

[18] Li, Q., Kumar, P., & Alazab, M. (2022). IoT-assisted physical education training network virtualization and resource management using a deep reinforcement learning system. Complex & Intelligent Systems, 1-14.

[19] Oyebode, O., Fowles, J., Steeves, D., & Orji, R. (2023). Machine learning techniques in adaptive and personalized systems for health and wellness. International Journal of Human–Computer Interaction, 39(9), 1938-1962.

[20] Jamil, F., Kahng, H. K., Kim, S., & Kim, D. H. (2021). Towards secure fitness framework based on IoT-enabled blockchain network integrated with machine learning algorithms. Sensors, 21(5), 1640.

[21] Yang, F. (2023). Optimization of personalized recommendation strategy for e-commerce platform based on artificial intelligence. *Informatica*, *47*(2).

[22] Fang, J., Lee, V. C., Ji, H., & Wang, H. (2024). Enhancing digital health services: A machine learning approach to personalized exercise goal setting. Digital Health, 10, 20552076241233247.

[23] Zhao, Q., Bie, Z., & Li, Y. (2025). Exploration of Sports Data Analysis and Fitness Effect Optimization Strategies using XGBoost.

[24] https://archive.ics.uci.edu/dataset/231/pamap2+physical+activity+monitoring