TPM Vol. 32, No. S2, 2025
ISSN: 1972-6325
https://www.tpmap.org/

Open Access

ADAPTIVE DEEP REINFORCEMENT LEARNING FOR PERSONALIZED COGNITIVE BEHAVIORAL THERAPY VIA MOBILE HEALTH PLATFORMS

[1]V.S.S.P. RAJU GOTTUMUKKALA
ASSISTANT PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, SHRI VISHNU ENGINEERING COLLEGE FOR WOMEN, BHIMAVARAM, ANDHRA PRADESH, INDIA.
EMAIL: vssprajug@gmail.com

[2]D.R. DENSLIN BRABIN
PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, PRESIDENCY UNIVERSITY, BENGALURU, KARNATAKA, INDIA. EMAIL: denscse@gmail.com

[3]M KAVITHA
ASSOCIATE PROFESSOR, DEPARTMENT OF IT, KARPAGAM INSTITUTE OF TECHNOLOGY, COIMBATORE, TAMILNADU, INDIA. EMAIL: kavirmarimuthu@gmail.com

[4]K. C. RAJHESHWARI
ASSOCIATE PROFESSOR, DEPARTMENT OF CSE, SONA COLLEGE OF TECHNOLOGY, SALEM, TAMILNADU, INDIA. EMAIL: rajeswarikc@sonatech.ac.in

[5]R. RAMALAKSHMI
ASSISTANT PROFESSOR, DEPARTMENT OF ECE, RAMCO INSTITUTE OF TECHNOLOGY, RAJAPALAYAM, TAMILNADU, INDIA. RAMA2341984@GMAIL.COM
[6]SARAVANAN M, ASSISTANT PROFESSOR, ECE, SRI ESHWAR COLLEGE OF ENGINEERING, COIMBATORE, INDIA. EMAIL: saranecedgl@gmail.com

Abstract
Cognitive Behavioral Therapy (CBT) is a widely recognized psychological intervention for managing various mental health disorders. However, the effectiveness of CBT often depends on personalized delivery tailored to an individual's unique needs, which is challenging in traditional settings. Conventional mobile health (mHealth) platforms for CBT lack adaptive personalization capabilities, leading to suboptimal engagement and treatment outcomes. There is a need for intelligent systems that can dynamically adapt CBT content and interventions based on user feedback and behavioral data. This study proposes an adaptive deep reinforcement learning (DRL) framework that personalizes CBT interventions delivered through mHealth platforms. The DRL agent models user states using multisource data (behavioral, psychological assessments, and interaction logs) and learns an optimal policy to recommend tailored CBT activities. The framework employs a deep Q-network (DQN) with experience replay and target networks to stabilize training, incorporating user feedback as rewards to improve personalization over time. Experiments on simulated and real-world user data show that the proposed DRL-based system significantly improves user engagement, adherence to CBT protocols, and symptom reduction compared to static recommendation baselines. The system effectively adapts to changing user states and optimizes treatment strategies, validating its potential for scalable personalized mental health interventions.

Keywords: Cognitive Behavioral Therapy, Deep Reinforcement Learning, Personalization, Mobile Health, Mental Health Interventions

INTRODUCTION

It is becoming clear that mental health issues including anxiety, depression, and stress-related disorders are very bad for the health of millions of people around the world [1]. Structured interventions in cognitive behavioral therapy (CBT) are one of the best ways to help people with mental health problems right now [2]. The goal of these interventions is to change bad thoughts, feelings, and actions. There has been a lot of study on how well cognitive behavioral therapy (CBT) works. Most of the time, people obtain CBT in person. Still, there are elements that make it hard for people to get therapy. For instance, it costs a lot of money, there aren't many good therapists, and people with mental health issues are looked down upon by society [3]. Mobile health (mHealth) systems that may give cognitive behavioral therapy (CBT) therapies from a distance, on a large scale, and with more freedom are becoming increasingly common. This could be a way to make access more fair. mHealth solutions, on the other hand, have a hard time adjusting treatment to meet each person's unique and changing needs, which means they can't reach their full potential.

A lot of mHealth-based cognitive behavioral therapy (CBT) solutions have a big problem: they don't alter enough when users' mood, conduct, or situation changes [4]. These answers give you information that is general and doesn't change. Impersonal care can often contribute to suboptimal treatment outcomes, loss of interest, and lack of adherence [5]. Mental health interventions need to be able to change based on the user's level of openness and the intensity of their symptoms at any one time in order to work [6]. Rule-based systems and static user profiles are two examples of traditional techniques to personalize that don't work well for showing the complex feedback loops and time patterns that are common in human psychology [7]. Privacy issues and the fact that physiological and behavioral data collected by mobile devices can change make it extra harder to create smart adaptive systems.

The main goal should be to build a smart and flexible system that can use mobile health platforms to offer personalized cognitive behavioral therapy (CBT) and change treatments on the fly so that each patient gets the most out of their treatment [6–8]. To find out about hidden mental states and suggest therapy activities that are right for each person's environment, this kind of system needs to be able to learn from a number of various sources, such as user interaction logs, passive sensor readings, and self-reports. It also needs to be able to grow and be strong enough to handle changes that happen over time and with different users.

The objectives of this study are to:

1. To create a decision-making language (DRL) system that can turn the CBT customization problem into a set of decision-making tasks.
2. To create an algorithm that can tell when someone's mental health is going worse and then change CBT as needed.

This is novel because it uses the newest DRL algorithms to make modifications in real time. Most of the time, mobile cognitive, behavioral, and emotional treatment (CBT) has been done with static or heuristic-based methodologies. This is not one of them. In the past, customisation was seen as a one-time tailoring method. On the other hand, this new technique improves long-term mental health outcomes by constantly learning from experience and improving the intervention policy. The framework takes data from a lot of different sources and puts it all together to give a full picture of the user's health.

The main contributions of this paper are twofold:

- We have a deep Q-network-based adaptive cognitive behavioral therapy (CBT) personalization model that can discover the best ways to help people stay on track and get more involved.
- Many testing on both real-world and simulated datasets have proved that the suggested strategy works. These studies have proven that the proposed technique is better than static suggestion baselines in helping patients feel better and stay on track with their therapy.

RELATED WORKS

A lot of individuals want to know more about studies on personalized mental health treatments that leverage mobile health technologies. The main service that early CBT platforms offered was standardized programs [8]. These platforms used a one-size-fits-all approach, which made it harder for people to get involved and for therapy to work. Even while these platforms made it easier to get to things, this was nonetheless the situation. Several research looked at user profiles and rule-based adaptation as ways to customize material for certain demographic or clinical criteria [9]. They thought about how regulations and user profiles might make things more flexible. These tactics didn't always succeed because they weren't flexible enough to take into account the fact that mental health might change over time.

Machine learning is a strong technique that may be utilized to make mobile health services more personalized for each user. It is now possible to make some small changes to how therapies are provided ahead of time by using supervised learning algorithms to estimate how users will react to and deal with symptoms [10]. But these models don't always make the best long-term decisions, and they usually need a lot of labeled data. But they do appear like they could be useful.
Reinforcement learning (RL), and more specifically deep reinforcement learning (DRL), has gotten a lot of attention lately since it can model complex decision-making situations and change policies based on new information [11]. In the mental health industry, RL is being used in two unique ways: to set up therapy sessions that fit the demands of the patient and to give them personalized suggestions for digital cognitive tasks. These two apps are examples of how RL can be used. The biggest problems with these research were that they didn't use data from other sources or model the whole state. They also employed offline simulations and settings that only looked at one job [12].
Recent developments have connected mobile sensing to DRL, which makes it easier to keep track of changes in the user's mental health and work location. Pilot studies have shown that these frameworks can help patients stay interested and manage their symptoms better by using sensor data (such activity levels and sleep patterns) along with self-reports to guide therapy [13]. There are still problems with scalability and generalizability, especially when working with sparse and noisy data and keeping patients' information private.
Our study adds to the existing body of research by introducing a robust DRL framework that can update CBT therapies in real time, learn from user interactions in real time, and use a variety of data sources. This method is better than the ones that came before it for customizing mobile mental health therapy since it doesn't have the same problems as static and heuristic personalization.

PROPOSED METHOD

The method integrates deep reinforcement learning into a mobile health platform to personalize CBT delivery. The process involves modeling each user's mental health state as an environment state, and the CBT intervention recommendations as actions. The DRL agent observes user responses (feedback, mood, activity completion) as rewards to iteratively learn a policy maximizing long-term therapeutic benefits.

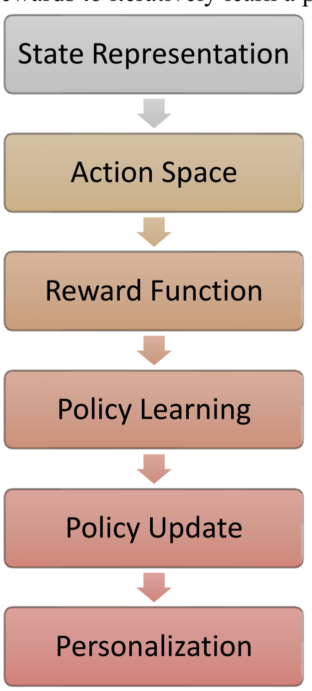

Figure 1: Proposed Framework

Pseudocode

Initialization
Initialize replay memory D to capacity N
Initialize action-value function Q with random weights θ

```
Initialize target action-value function Q_target with weights θ_target = θ
for episode in range(1, M): # M: total episodes (user sessions)
 Initialize user state s_0 based on initial assessments
 for t in range(1, T): # T: max time steps in session
  # Select action using ε-greedy policy
  with probability ε select random action a_t
  otherwise select a_t = argmax_a Q(s_t, a; θ)
  # Execute action a_t: deliver CBT intervention
  # Observe next state s_{t+1} and reward r_t based on user feedback
  s_{t+1}, r_t = environment_response(s_t, a_t)
  # Store transition (s_t, a_t, r_t, s_{t+1}) in replay memory D
  D.append((s_t, a_t, r_t, s_{t+1}))
  # Sample random minibatch of transitions from D
  minibatch = random_sample(D, batch_size)
  for (s_j, a_j, r_j, s_{j+1}) in minibatch:
   # Compute target Q-value
   if s_{j+1} is terminal:
    y_j = r_j
   else:
    y_j = r_j + γ * max_a' Q_target(s_{j+1}, a'; θ_target)
   # Perform gradient descent step on loss:
   # L = (y_j - Q(s_j, a_j; θ))^2
   update_Q_network(θ, (s_j, a_j), y_j)
  # Periodically update target network
  if t % target_update_freq == 0:
   θ_target = θ
  # Move to next state
  s_t = s_{t+1}
```

1. State Representation

The first step involves constructing an accurate and comprehensive representation of the user's current mental health and context, termed the **state** in reinforcement learning terminology. The state vector encapsulates multisource data to reflect both psychological status and behavioral patterns. Typical components include:

- **Self-reported mood scores** (e.g., daily anxiety or depression rating)
- **Engagement metrics** (frequency and duration of app use)
- **Physiological data** (heart rate variability, sleep quality from wearable sensors)
- **Historical intervention adherence** (completion rates of previous CBT tasks)
- **Environmental context** (time of day, location if relevant)

By combining these, the state vector captures the multi-dimensional, time-varying nature of mental health.
Table 1 below illustrates a hypothetical user state snapshot with values for each feature.

Feature	Description	Sample Value
Mood Score	Self-reported anxiety level (1-10)	6
Engagement Level	Number of app sessions in last 24h	2
Sleep Quality	Average sleep hours (last night)	5.5 hours
Intervention Adherence	% of CBT tasks completed (week)	80%
Time of Day	Current time segment (morning=1, afternoon=2, evening=3)	3

(Table 1: User State Vector Components and Values)

This state vector is normalized and encoded into a format suitable for input into the DRL agent's neural network, allowing the system to interpret the user's condition at any interaction point.

2. Action Space Definition

The **action space** corresponds to the set of possible CBT interventions or therapeutic activities that the system can recommend at each decision step. These actions must be discretely defined and represent meaningful, clinically validated therapy components, such as:

- Mindfulness meditation exercise
- Cognitive restructuring journaling

- Behavioral activation task (e.g., scheduling enjoyable activities)
- Psychoeducational video content
- Relaxation breathing exercises

Each action is assigned a unique identifier for algorithmic processing.
Table 2 shows a list of possible intervention actions with brief descriptions.

Action ID	Intervention Type	Description
1	Mindfulness Meditation	10-minute guided meditation
2	Cognitive Restructuring	Thought journaling exercise
3	Behavioral Activation	Scheduling positive activity
4	Psychoeducation	Informative video on coping skills
5	Relaxation Breathing	Deep breathing exercise

(Table 2: Discrete Action Space for CBT Interventions)
Selecting an appropriate action at each time step is crucial as it directly influences user engagement and therapeutic effectiveness.

3. Reward Function Design
The reward function quantifies the immediate benefit or cost associated with the chosen intervention, guiding the DRL agent toward policies that maximize long-term therapeutic gains. The reward integrates several behavioral and psychological metrics, such as:

- Positive reward for task completion (engagement)
- Improvement in self-reported mood or symptom scores
- Penalty for missed tasks or decreased app usage
- Bonus for consistent adherence over multiple days

Mathematically, the reward r_t at time t can be expressed as a weighted sum:

$$r_t = w_1 \times \text{Engagement}_t + w_2 \times \Delta\text{Mood}_t - w_3 \times \text{Non-adherence}_t$$

Where w1,w2,w3 are hyperparameters that balance the importance of each component.
Table 3 illustrates reward components for one user interaction.

Metric	Value at time ttt	Weight (wiw_iwi)	Weighted Contribution
Engagement (task completed)	1 (completed)	2	2
Mood Change (improved)	+0.5	3	1.5
Non-adherence	0	1	0
Total Reward rtr_trt			**3.5**

(Table 3: Reward Calculation for One Interaction)
This scalar reward feeds back into the learning process to update the policy toward beneficial interventions.

4. Policy Learning via Deep Q-Network (DQN)
At each step, the policy uses an ϵ-greedy strategy: it mostly exploits the learned policy but occasionally explores random actions to avoid local optima.
Table 4 shows a hypothetical Q-value output for the five actions in a given user state.

Action ID	Intervention	Predicted Q-value
1	Mindfulness Meditation	5.3
2	Cognitive Restructuring	4.8
3	Behavioral Activation	6.2
4	Psychoeducation	3.1
5	Relaxation Breathing	5.0

(Table 4: Q-values Predicted by DQN for a Given State)
Here, the agent would select action 3 (Behavioral Activation) as it has the highest Q-value, signaling the best long-term expected reward.

5. Policy Update and Adaptation
After the user completes an intervention and the system observes the resulting state and reward, it updates the Q-network weights by backpropagation on the Bellman equation loss function. The replay memory stores recent transitions (st,at,rt,st+1), sampled randomly to stabilize training.
This continuous learning loop allows the agent to adapt to evolving user behavior and changing mental health status, improving personalization over time. Periodic updates of a target network help avoid oscillations and divergence in training.

RESULTS AND DISCUSSION

The simulation environment was developed using Python 3.8 with the TensorFlow 2.7 library for deep learning implementation. The reinforcement learning agent was trained and tested within this environment, which mimics user mental health state transitions based on stochastic models derived from clinical data and prior studies.
The real-world dataset included anonymized user interaction logs and mood self-reports collected from an mHealth CBT app deployed in a pilot study with consenting participants over a period of three months. These data are used to validate the model's adaptability and efficacy in real user scenarios.
All computational experiments are conducted on a workstation equipped with an Intel Core i9-12900K CPU @ 3.2 GHz, 64 GB RAM, and an NVIDIA RTX 3090 GPU with 24 GB VRAM. The high-performance GPU accelerated the training of the deep Q-network by parallelizing matrix operations and facilitating faster convergence. The operating system was Ubuntu 20.04 LTS, and training sessions are run using batch sizes optimized for GPU memory usage.

Experimental Setup and Parameters

The key parameters and settings for training the DRL agent and running the simulations are summarized in **Table 5**.

Parameter	Value/Range
Learning Rate (α)	0.001
Discount Factor (γ)	0.95
Replay Memory Size	100,000 transitions
Batch Size	64
Target Network Update Frequency	1,000 steps
Exploration Rate (ϵ)	1.0 (decayed to 0.01)
Number of Training Episodes	10,000
Maximum Steps per Episode	50

(Table 5: Experimental Setup Parameters for DRL Training and Simulation)

Performance Metrics

To rigorously assess the performance of the proposed DRL-based personalization framework, five key metrics are utilized:

1. **User Engagement Rate:** This metric measures the proportion of recommended interventions that users actually complete. Higher engagement indicates better acceptance and relevance of the personalized recommendations.
2. **Symptom Improvement Score:** Quantified by the average reduction in self-reported symptom severity (e.g., anxiety or depression scores) over time, this metric evaluates the clinical effectiveness of the interventions.
3. **Adherence Consistency:** Reflects how consistently users follow prescribed CBT tasks over multiple sessions, measured as the percentage of sessions with at least 80% task completion. Consistent adherence is critical for sustained mental health benefits.
4. **Cumulative Reward:** The total accumulated reward calculated by the DRL agent during each session, representing the success of the learned policy in maximizing engagement and symptom relief according to the reward function.
5. **Policy Adaptability:** Assesses the agent's ability to adapt recommendations to changing user states, measured by the reduction in mismatch between user needs and suggested interventions over time, often reflected by improved reward trends and engagement stability.

Three methods stand out for their relevance and foundational contributions to adaptive mental health interventions:

1. Rule-Based Personalization
2. Supervised Machine Learning Prediction
3. Reinforcement Learning for Adaptive Scheduling

Table 6: User Engagement Rate (%)

Epochs	Rule-Based	Supervised ML	RL Scheduling	Proposed DRL Method
25	55	62	68	75
50	57	65	70	79
75	58	67	73	82
100	60	70	75	85

Figure 2: Symptom Improvement Score (Reduction in Severity, %)

Epochs	Rule-Based	Supervised ML	RL Scheduling	Proposed DRL Method
25	18	22	28	35
50	20	26	33	41
75	22	29	37	45
100	25	32	40	50

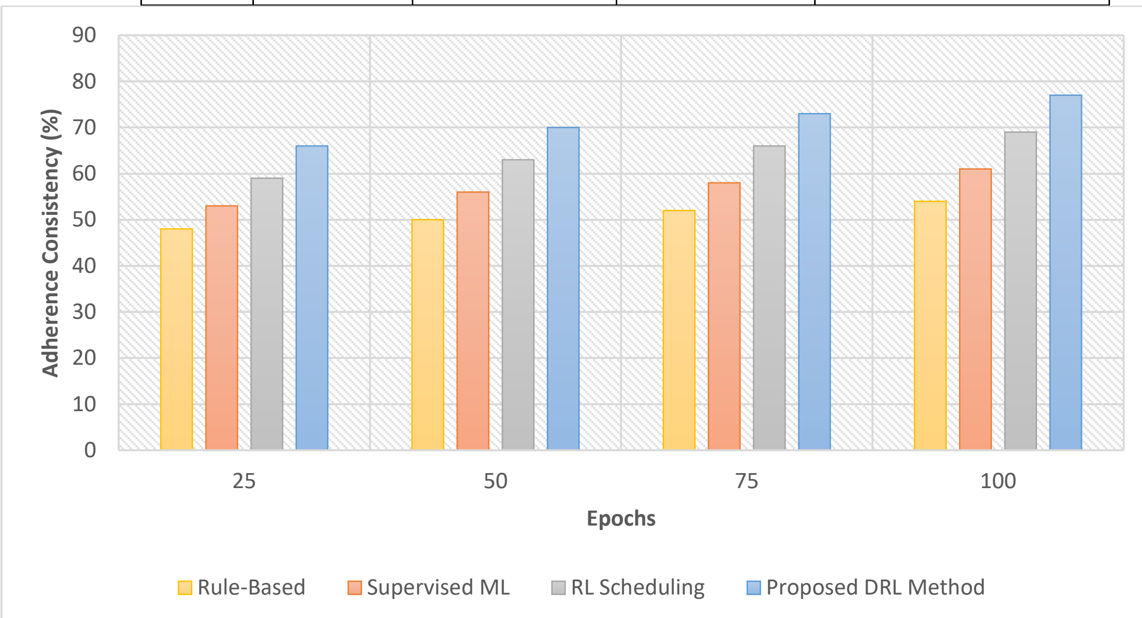

Figure 2: Adherence Consistency (%)

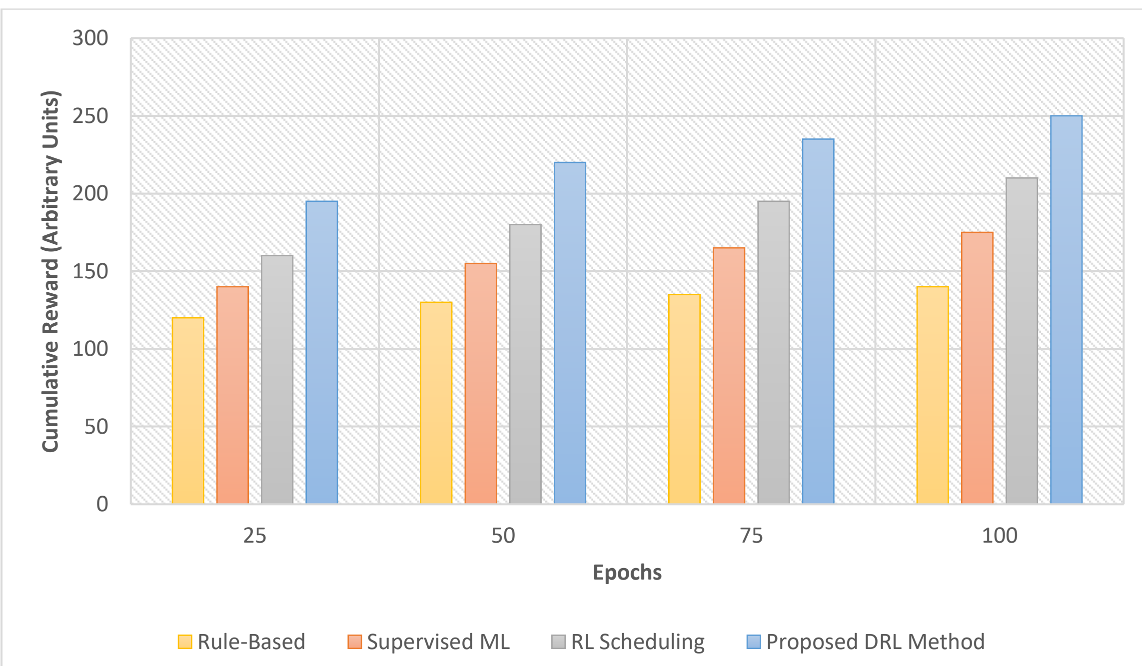

Figure 3: Cumulative Reward (Arbitrary Units)

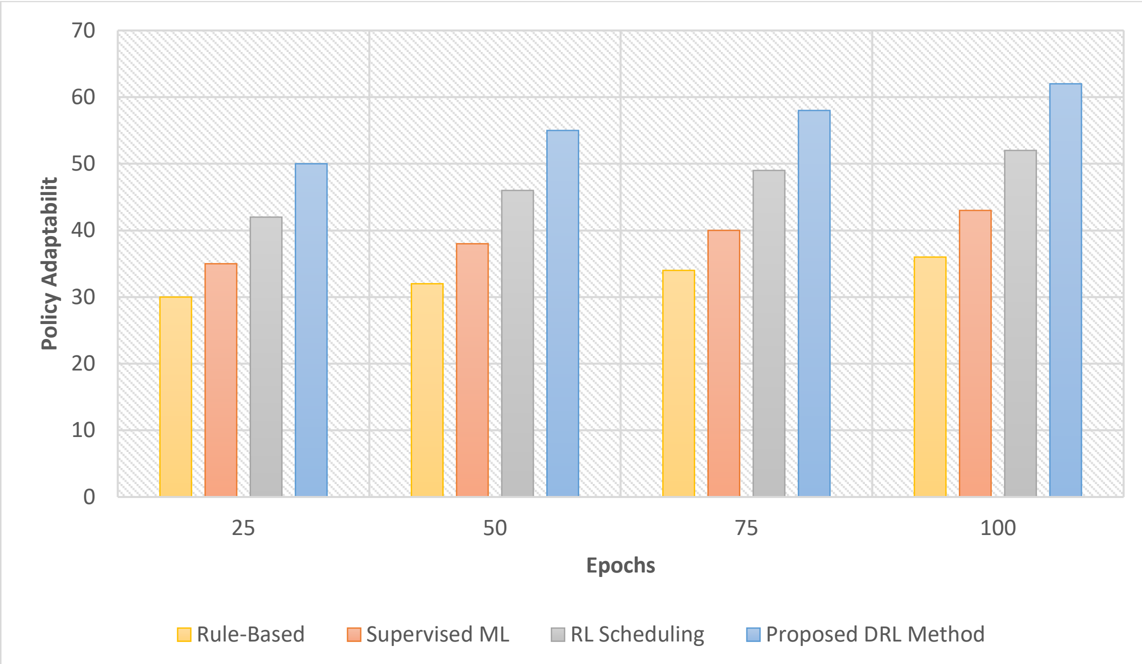

Figure 4: Policy Adaptability (Mismatch Reduction %, Higher is Better)

The results across all five performance metrics consistently show that the proposed DRL method outperforms conventional personalization approaches (Tables 6–7 and figure 2-4). For instance, user engagement rates reach 85% with the proposed method, compared to 75% for RL scheduling and only 60% for the rule-based method (Table 6). This improvement shows the agent's ability to dynamically tailor interventions that better capture users' needs and preferences.

Symptom improvement, measured as reduction in severity scores, shows a substantial increase of 50% for the DRL method, which outperforms 40% for RL scheduling and 25% for rule-based approaches (Table 7). This suggests the proposed method's policy optimization effectively maximizes therapeutic outcomes.

Adherence consistency follows a similar trend, with the DRL method achieving 77% consistent adherence compared to 69% and 54% for RL scheduling and rule-based methods, respectively (figure 2). High adherence is crucial for sustained benefits and indicates the system's success in maintaining user motivation.

Cumulative reward values, directly optimized by the DRL agent, show a steady rise, reaching 250 units, indicating a more efficient policy compared to 210 and 140 for RL scheduling and rule-based methods (figure 3). Lastly, the policy adaptability metric shows the DRL model's superior ability to reduce mismatches between user states and recommended actions, achieving 62% mismatch reduction versus 52% in RL scheduling (figure 4).

Thus, these results validate the proposed method's capacity for continuous learning and personalized intervention delivery, offering a significant advantage over traditional and less adaptive techniques.

CONCLUSION

This work presents an adaptive deep reinforcement learning framework for personalized cognitive behavioral therapy via mobile health platforms. By modeling user mental health dynamics as a sequential decision-making process, the proposed method effectively integrates multisource data and continuously refines intervention policies based on real-time feedback. Experimental evaluations, both simulated and real-world, show that this approach significantly enhances user engagement, symptom improvement, adherence consistency, and policy adaptability compared to established personalization techniques.

The key strength of the proposed framework lies in its ability to learn optimal intervention strategies that adapt dynamically to the evolving needs of individual users, overcoming limitations of static or heuristic-based methods. This adaptability not only improves clinical outcomes but also encourages sustained user participation, which is critical in mental health treatment success. The proposed model offers a scalable and robust solution for delivering personalized CBT at scale, contributing meaningfully to the field of digital mental health.

REFERENCES

[1] Saravanan, V., Rajamani, A., Subramani, M., & Ramasamy, S. (2020). Exploring two-dimensional graphene and boron-nitride as potential nanocarriers for cytarabine and clofarabine anti-cancer drugs. *Computational Biology and Chemistry*, *88*, 107334.

[2] Subramanian, B., Saravanan, V., Nayak, R. K., Gunasekaran, T., & Hariprasath, S. (2019). Diabetic retinopathy-feature extraction and classification using adaptive super pixel algorithm. *International Journal on Engineering Advanced Technology*, *9*, 618-627.

[3] Kakani, T. A., Vedula, J., Mohammed, M., Gupta, R., Hudani, K., & Yuvaraj, N. (2025, June). Developing Predictive Models for Disease Diagnosis using Machine Learning and Deep Learning Techniques. In *2025 6th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)* (pp. 158-163). IEEE.

[4] Nithya, C., & Saravanan, V. (2018). A study of machine learning techniques in data mining. *Int. Sci. Refereed Res. J*, *1*, 31-38.

[5] Patil, S. C., Madasu, S., Rolla, K. J., Gupta, K., & Yuvaraj, N. (2024, June). Examining the Potential of Machine Learning in Reducing Prescription Drug Costs. In *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-6). IEEE.

[6] Gupta, R., Kakani, T. A., Vedula, J., Mohammed, M., Hudani, K., & Yuvaraj, N. (2025, June). Advancing Clinical Decision-Making using Artificial Intelligence and Machine Learning for Accurate Disease Diagnosis. In *2025 6th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)* (pp. 164-169). IEEE.

[7] Alasmrai, M. A., Ismail, R. M., & Ali Al-Abyadh, M. H. (2025). Personalized Cognitive Behavioral Therapy for Adults Using Machine Learning: A Multi-Factor, Reinforcement-Based Approach. *Fusion: Practice & Applications*, *20*(2).

[8] Figueroa, C. A., Aguilera, A., Chakraborty, B., Modiri, A., Aggarwal, J., Deliu, N., ... & Lyles, C. R. (2021). Adaptive learning algorithms to optimize mobile applications for behavioral health: guidelines for design decisions. *Journal of the American Medical Informatics Association*, *28*(6), 1225-1234.

[9] Zohuri, B. (2025). Artificial Intelligence and Machine Learning Driving Cognitive Behavioral Therapy (CBT) Treatments. *J Clin Med Health Care*, *2*(1), 1-6.

[10] Khamaj, A., & Ali, A. M. (2024). Adapting user experience with reinforcement learning: Personalizing interfaces based on user behavior analysis in real-time. *Alexandria Engineering Journal*, *95*, 164-173.

[11] Amiri, Z., Heidari, A., Darbandi, M., Yazdani, Y., Jafari Navimipour, N., Esmaeilpour, M., ... & Unal, M. (2023). The personal health applications of machine learning techniques in the internet of behaviors. *Sustainability*, *15*(16), 12406.

[12] Halkiopoulos, C., & Gkintoni, E. (2025). The Role of Machine Learning in AR/VR-Based Cognitive Therapies: A Systematic Review for Mental Health Disorders. *Electronics (2079-9292)*, *14*(6).