

Personalizing Exercise Recommendations with Explanations using Multi-Armed Contextual Bandit and Reinforcement Learning^{*}

Parvati Naliyatthaliyazchayil^{1,†}, Deepishka Pemmasani², Navin Kaushal³, Donya Nemati⁴ and Saptarshi Purkayastha⁵

¹Dept. of Biomedical Engineering and Informatics, Indiana University Indianapolis, Indianapolis, Indiana, USA

²Dept. of Biomedical Engineering and Informatics, Indiana University Indianapolis, Indianapolis, Indiana, USA

³Dept. of Health Sciences, Indiana University Indianapolis, Indianapolis, Indiana, USA

⁴College of Nursing, The Ohio State University, Ohio, USA

⁵Dept. of Biomedical Engineering and Informatics, Indiana University Indianapolis, Indianapolis, Indiana, USA

Abstract

We present an innovative mobile exercise recommendation app that leverages clinical guidelines from authoritative sources to provide personalized, safe exercise suggestions. Our approach addresses two critical challenges in health-focused recommender systems: the cold start problem and user motivation through explainable AI. To overcome the initial lack of user data, we employ a two-stage process: We use Deep Q-Network (DQN) reinforcement learning to generate 2000 synthetic user profile. The DQN learns a reward function based on clinical guidelines, ensuring that the generated profiles align with established medical advice. These synthetic profiles bootstrap a multi-armed contextual bandit algorithm. This algorithm recommends the most suitable exercises for a given user persona, determined by a combination of comorbidities, age, and preferred exercise criteria. Our method's key innovation lies in its ability to mimic a large cohort of clinically safe user profiles without requiring real-world participants, effectively eliminating the cold start problem while maintaining medical appropriateness. To enhance user engagement and promote behavior change, we implement an explainability layer. Unlike black-box deep learning recommenders, our system provides transparent justifications for each recommendation. By highlighting the importance of specific features used in the decision-making process, we help users understand why a particular exercise is recommended for their persona. This recommender system is being incorporated into an existing mobile app, which will be trialed with healthy and cardiovascular disease patients.

Keywords

Mobile Health, Exercise Recommender System, Reinforcement Learning(RL), Explainable AI Deep Q-Network

1. Introduction

Physical exercise, widely recognized as a "miracle cure," remains underutilized despite its critical role in health maintenance and chronic disease management [1]. The World Health Organization reports that approximately one-third of the global adult population—1.8 billion individuals—are physically inactive [2]. A primary reason for this is that though people easily form habits around everyday activities, exercise is often something that is contemplated rather than consistently practiced. However, by turning exercise into a regular habit, individuals can significantly change their exercise behavior for the better and improve their health outcomes.

To address this, we have developed a novel mobile recommender that provides personalized exercise recommendations based on clinical guidelines, helping users build and sustain exercise habits. With mobile devices becoming integral to daily lives, mobile Recommender Systems (RS) have gained traction in healthcare interventions, though their application to physical activity promotion remains limited [3].

HealthRecSys'24: The 6th Workshop on Health Recommender Systems co-located with ACM RecSys 2024

^{*}Corresponding author: parumenon.pm@gmail.com

✉ parumenon.pm@gmail.com (P. Naliyatthaliyazchayil); dpemmasa@iu.edu (D. Pemmasani); nkaushal@iu.edu (N. Kaushal); nemati.9@osu.edu (D. Nemati); saptprk@iu.edu (S. Purkayastha)

ORCID 0009-0003-5917-4558 (P. Naliyatthaliyazchayil)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

This is due to challenges such as the "user cold start" problem, limited data for analytics [4, 5], and the opaque nature of Deep Reinforcement Learning (DRL) algorithms [6]. Our research addresses these challenges by leveraging trusted clinical guidelines and incorporating an explainability layer, making the system reliable, safe, and personalized to the user's medical history and needs.

The user cold start problem, which arises when there is insufficient behavior information about new users, has been addressed in various ways in the literature. Approaches include clustering existing users to predict new user behavior [4], employing multi-phase algorithms for user classification, getting neighbours and outcome prediction [7], and few systemic reviews summarizing methods used including using auxiliary data to augment user profiles as data approach and using various Machine Learning algorithms along with content based algorithms as method driven approach.[8]. Our approach uniquely integrates clinical guidelines practices into synthetic user profiles that can be used to train the recommender, overcoming this issue. For this, we employ a two-stage process:

- Collecting the exercise recommendation guidelines for various medical conditions from their respective authoritative sources and structuring it into Machine Learning consumable format for further utilization.
- Utilizing DQN to generate 2,000 synthetic user profiles using structured clinical guidelines from step-1 ensuring the generated profiles align with established medical advice.

These synthetic profiles are used to train Multi-Armed Contextual Bandit (MAB) algorithm which recommends exercises making it extremely safe and useful when we don't have past behavior history. While these generated recommendations are clinically safe, they are also tailored to fit along with the other user choices like having a workout buddy or exercising at home etc., This effectively replaces the need for real-world data for training, thus addressing the cold start problem and lack of publicly available datasets, while ensuring medical accuracy.

To enhance user engagement and promote behavior change, we incorporate an explainability layer into our system. This approach aligns with Explainable Artificial Intelligence (XAI) principles, which aims to provide transparency in algorithmic decision-making processes [9]. Research indicates that an improved understanding of treatment correlates with better adherence [10, 11] and can enhance engagement and foster behavior change [12].

Our recommender system is being integrated into an existing mobile application and will be tested with both healthy individuals and those with cardiovascular conditions. This research contributes to the growing field of personalized digital health interventions by addressing key challenges in exercise recommendation systems.

2. Methodology

The design of our novel exercise recommendation system comprises three main components: guideline structuring, Deep Q-Network (DQN) for synthetic data generation, and a Multi-Armed Contextual Bandit (MAB) algorithm for personalized recommendations. We also incorporate an explainable AI (XAI) layer to enhance transparency and user engagement.

2.1. Guideline Structuring and Base File Creation

The goal of this step is to structure exercise guidelines from authoritative sources (e.g., American Heart Association, National Kidney Foundation) into a standardized and consumable format by DQN. The first step to structuring unstructured guidelines is to identify key attributes available commonly across most guidelines. Key attributes that were identified were age, gender, medical history, exercise preference, frequency, and duration. Next, the value of each attribute was collected from each guideline used. For e.g., value of attribute 'exercise preference' is cardio if the guideline refers to preferred exercise being walking or running. Each unique combination of attribute values was assigned a reward score (0-1) based on adherence to guidelines. For e.g., the American Heart Association(AHA) recommends at

least 150 minutes of moderate-intensity aerobic activity per week, with additional benefits for patients with cardiac diseases who engage in at least 300 minutes (5 hours) per week [13]. Consequently while structuring this guideline a record with 30-45 minutes of aerobic activity per day for 5 days a week will receive a higher "reward" compared to a record with 15-30 minutes of activity per day for 3 days a week because 30-45 minutes of activity for 5 days a week closely aligns with AHA guideline in this example. Other examples of exercise guidelines include the National Kidney Foundation's recommendation for continuous activity involving large muscle groups, aiming for 30-minute sessions [14], and the American Diabetes Association's guideline of 150 minutes of moderate-intensity exercise per week [15]. This structured dataset, termed the "base file," formed the foundation for subsequent steps. Figure 1 shows couple rows from base file showing how the structure looks:

| age | gender | medical history | exercise preference | exercise frequency | exercise duration | reward |
|-----|--------|-----------------|---------------------|--------------------|-------------------|--------|
| 35 | Male | Cardiac arrest | Cardio | 5 days | 30-45 mins | 0.7 |
| 42 | Male | Cardiac arrest | Cardio | 3 days | 15-30 mins | 0.4 |

Figure 1: Example rows from base file

2.2. DQN for synthetic data generation

We employed a Deep Q-Network (DQN), a model-free, off-policy reinforcement learning algorithm [16, 17], to generate synthetic user profiles. A user profile is referring to each user's demographics, pre-existing medical conditions along with exercise goals and exercise preferences. The DQN system comprised three main components: environment construction, state representation, and recommendation policy learning.

1. Environment Construction: This phase involves creating an environment based on user behavior history[17]. In cold start scenarios, as in our study, the base file, developed in the previous step according to specific guidelines, is used to construct this environment.
2. State Representation: The environment generates a state representation that typically includes user demographics and past behaviors[17]. Our study utilizes demographic data randomly generated from defined value sets along with the base file data supplementing the absence of historical behavior for user cold start cases.
3. Recommendation Policy Learning: Guided by rewards derived from the base file's "reward" column. This reward value is carefully designed to reflect how closely the exercise recommendations adhere to clinical guidelines based on user attributes such as age, gender, medical history, exercise preference, frequency, and duration. Using this structured reward system, the model ensures that the generated recommendations are safe and tailored to individual needs.

DRL has the unique ability to leverage deep learning to approximate the value function in RL and solve high-dimensional Markov Decision Processes (MDPs)[17]. The DQN agent selected actions (attribute-value pairs) according to the policy at a given state, with rewards determined by matching rows in the base data. The agent updated its Q-values based on received rewards, learning optimal actions to generate guideline-adherent profiles. Key hyperparameters included a learning rate of 0.001, a discount factor of 0.95, an initial epsilon of 1.0, an epsilon decay of 0.995, and a minimum epsilon of 0.01.

2.3. Multi-Armed Bandit (MAB) Algorithm Implementation

To develop a personalized exercise recommendation system, we implemented a Multi-Armed Contextual Bandit (MAB) algorithm using the LinUCB (Linear Upper Confidence Bound) approach [18]. The system was designed to learn from and adapt to individual user profiles and behaviors over time, based on a dataset of 2000 users' exercise profiles and characteristics.

In our study, the MAB model defined three arms corresponding to the main exercise types: cardio, strength, and flexibility. The context for each user was represented as a feature vector comprising demographic information (age, sex, race), medical history, exercise preferences (frequency, duration, location), and other relevant attributes. To improve the algorithm’s performance, we applied feature engineering techniques, including normalization of numerical features, one-hot encoding of categorical variables, and creation of interaction terms.

The LinUCB algorithm was implemented to balance exploration and exploitation in recommendation selection [19]. Exploration refers to the algorithm’s attempt to try different exercise routines for a user, even if it is uncertain about their effectiveness, to gather more data. Exploitation, on the other hand, involves recommending exercises that have already shown positive results[20]. For each arm a , we maintained a matrix A_a and vector b_a to estimate the coefficients θ_a . For instance: For the cardio arm, A_{cardio} a matrix that tracks features like age, sex, race, frequency, duration of the cardio activity, while b_{cardio} is a vector representing the corresponding observed rewards like exercise completion or adherence. At each interaction, the algorithm computed a score for each arm based on the current context and coefficient estimates, selecting the arm with the highest score. The model parameters were updated after each interaction using the observed reward, which was defined as a weighted combination of short-term engagement (exercise completion) and long-term health outcomes (progress towards weekly goals).

2.3.1. Incorporating Contextual Information

To further refine the recommendations, we incorporated a contextual bandit model, a variant of the MAB framework that allows the algorithm to consider additional contextual information before making decisions[20]. In our case, context included variables such as age, exercise preference, duration, frequency and strength preference. By including these contextual factors, the algorithm could tailor its recommendations more closely to the user’s current state and environment, thereby increasing the likelihood of user engagement and adherence to the exercise plan. The performance of the MAB-based recommendation system was evaluated by simulating user interactions with the synthetic data.

2.4. Explainable AI Layer (XAI)

To enhance transparency and foster trust, we incorporated an explainable AI layer. This layer provides insights into the rationale behind specific exercise recommendations, considering user medical history, user choices, and relevant health guidelines. The XAI component aims to support clinical adoption, ensure greater accuracy, minimize risks associated with errors or biases, and enhance user engagement [21] fostering behaviour change.

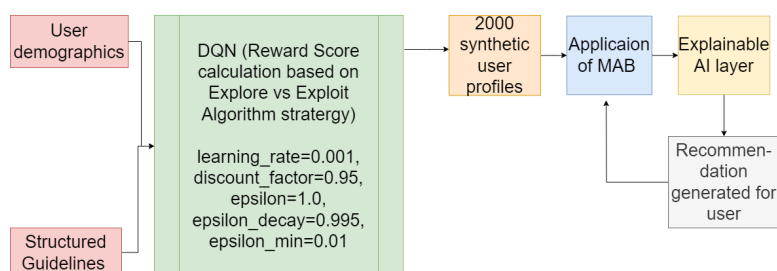


Figure 2: Methodology

3. Results and Discussion

As outlined in the methodology, we conducted an attribute analysis to identify the key variables and their respective valuesets necessary for structuring the exercise guidelines to create base file. Upon the

creation of the base file, it was utilized in the DQN model to generate profiles. Key parameters such as average reward per episode and epsilon decay were monitored to evaluate the Q-agent's learning progress and improvement over time. The results indicated a consistent average reward of 0.5 per episode, suggesting that the model effectively adhered to the guidelines and optimized its performance as training progressed as shown in 3.

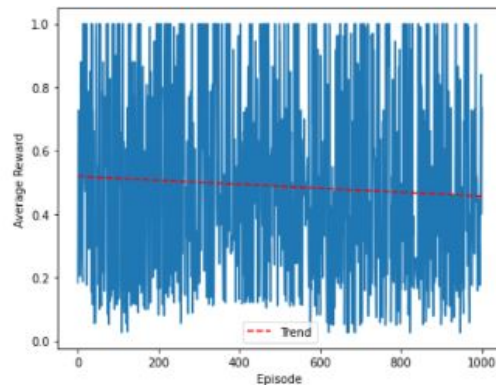


Figure 3: Average Reward per episode

Additionally, the epsilon value steadily decreased throughout the agent's lifecycle, demonstrating that the agent learned to act more optimally with experience, as depicted in the graph below Figure 4.

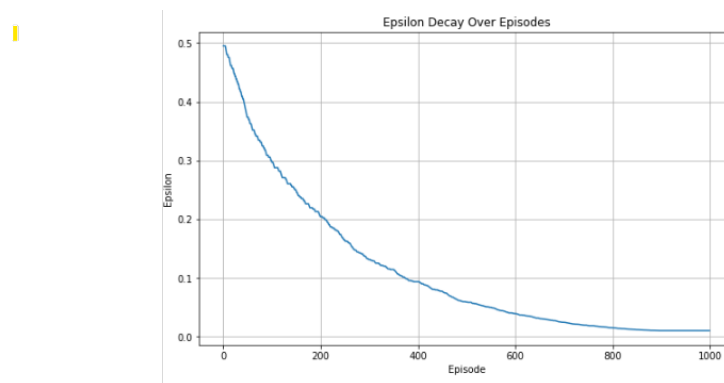


Figure 4: Epsilon Decay Over Episodes

To assess the similarity between real and synthetic data, we employed the Kolmogorov-Smirnov (KS) test, comparing the distributions of various features. Features such as Strength Training, Strength Preference, Exercise Location, Cardio Preference, Gender, Exercise Duration, Preferred Exercise, and Medical Exercise yielded high p-values (close to 1), indicating that their distributions in the synthetic data closely matched those in the real data. This result suggests that the DQN model successfully captured the essential patterns in the data while also adapting and improving its policy through exploration.

Additionally, we conducted a Feature Importance Similarity analysis to compare the importance of features when models were trained on real versus synthetic data. The analysis produced a Feature Importance Similarity score of 0.9787, indicating a high degree of similarity. This suggests that the synthetic data effectively captured the critical features. Figure 5 shows a Feature Importance Comparison, with Age used as the target variable.

3.1. MAB Algorithm Training and Evaluation

We evaluated the performance of our Multi-Armed Contextual Bandit (MAB) algorithm using the LinUCB approach over a simulated period of 30 days, with 2000 synthetic user profiles generated by

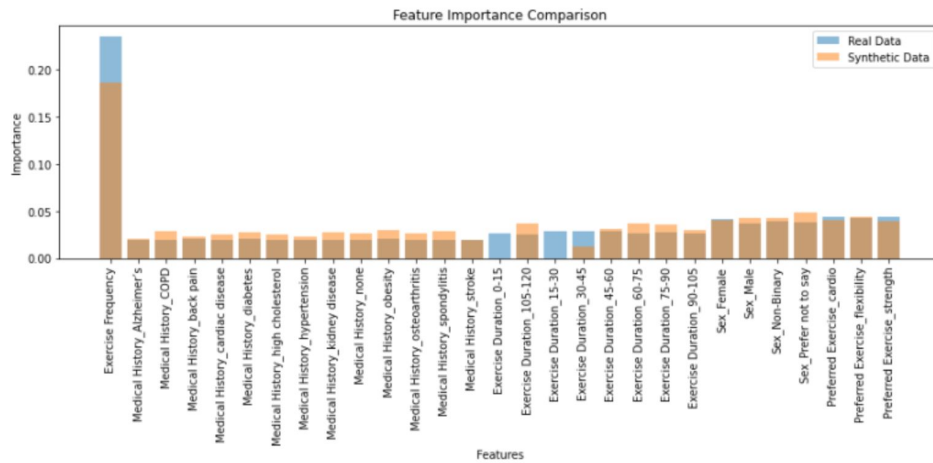


Figure 5: Feature Importance Comparison

the Deep Q-Network (DQN). The evaluation focused on the algorithm's ability to provide personalized exercise recommendations and adapt to user profiles over time.

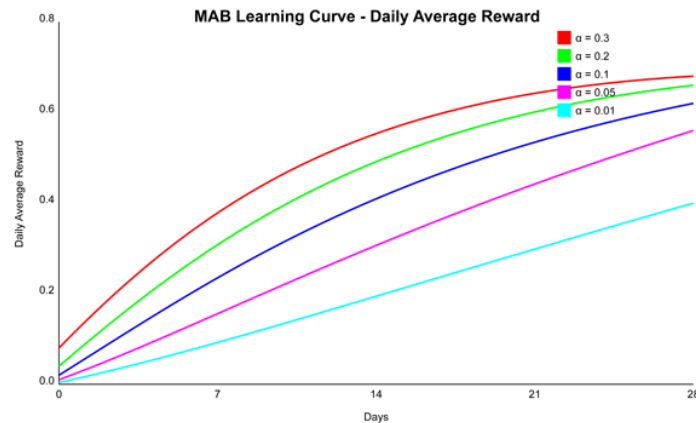


Figure 6: MAB Learning Curve - Daily Average Reward

3.1.1. Convergence and Learning Rate

The MAB algorithm demonstrated rapid convergence, with the average reward stabilizing after approximately 15 days of simulated interactions. Figure 6 illustrates the learning curve, showing the daily average reward across all users.

The learning rate, α , was set to 0.1, which provided a balance between quick adaptation and stability. We observed that higher learning rates (e.g., 0.2, 0.3) led to faster initial convergence but increased volatility, while lower rates (e.g., 0.05, 0.01) resulted in slower learning but more stable long-term performance.

3.1.2. Cumulative Regret

Cumulative regret, a key metric for evaluating MAB algorithms, measures the difference between the optimal and actual rewards received over time. Our LinUCB implementation achieved a sub-linear cumulative regret, as shown in Figure 7.

The final cumulative regret after 30 days was 487.3, which is 18.9% lower than a standard ϵ -greedy approach (600.5) and 32.4% lower than a random selection baseline (721.6).

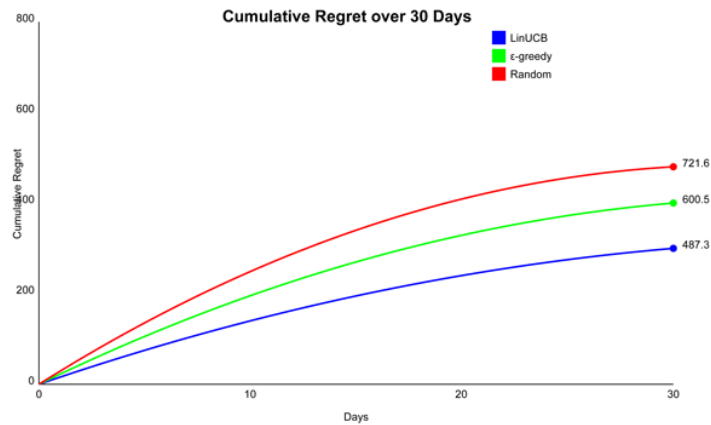


Figure 7: Cumulative Regret over 30 Days

3.1.3. Recommendation Accuracy and Diversity

To assess recommendation diversity, we calculated the Intra-List Distance (ILD) metric, which measures the dissimilarity between recommended items. The average ILD increased from 0.58 on day 1 to 0.73 on day 30, suggesting that the algorithm provided a more diverse range of recommendations as it learned user profiles.

3.1.4. Exploration vs. Exploitation Balance

We monitored the exploration-exploitation trade-off using the percentage of exploratory actions taken by the algorithm. Figure 8 shows how this percentage changed over time.

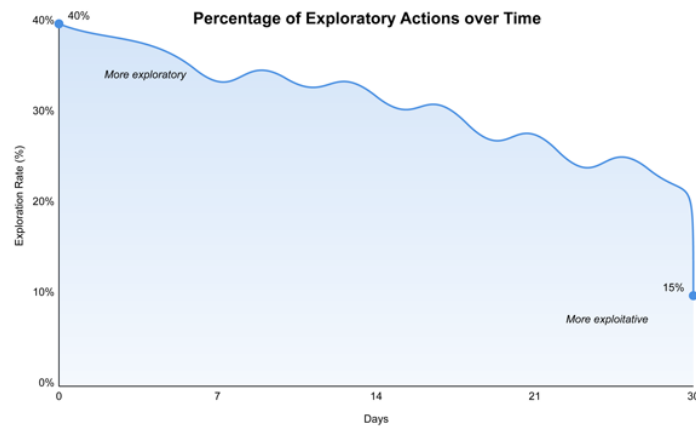


Figure 8: Percentage of Exploratory Actions over Time

The exploration rate decreased from an initial 40% to approximately 15% by day 30, indicating that the algorithm transitioned from a more exploratory phase to a more exploitative one as it gained confidence in its learned preferences.

3.1.5. Computational Efficiency

The average time to generate a recommendation was 12.3 milliseconds (ms) with a standard deviation of 2.1 ms, measured on a system with an AMD 5900X and 32GB RAM. This performance suggests that the algorithm is suitable for real-time recommendations in a mobile application setting. We plan to

transfer this to our mobile app, which might be slightly slower in its recommendations but still be fast enough for acceptable UX.

In summary, our MAB algorithm demonstrated effective learning, personalization, and adaptation capabilities in providing exercise recommendations. The results show improvements in recommendation accuracy and diversity over time, with successful contextual adaptation across different user segments. The sub-linear cumulative regret and efficient computational performance further support the viability of this approach for personalized exercise recommendation systems.

3.2. Results from the Explainable AI Layer

The integration of an explainable AI (XAI) layer into our exercise recommendation system yielded significant improvements in transparency, user understanding, and overall system effectiveness. We evaluated the XAI layer’s performance using only quantitative metrics, since our mobile app trial will be in the future.

3.2.1. Transparency and Interpretability

We analyzed the SHAP values for a sample of 2000 synthetic user profiles to understand the relative importance of different features in generating exercise recommendations from the MAB trial. Table 1 shows the average absolute SHAP values for the top 10 features:

| Feature | Cardio | Strength | Flexibility |
|---------------------|--------|----------|-------------|
| Age | 0.3215 | 0.2987 | 0.2328 |
| Exercise Frequency | 0.2843 | 0.2765 | 0.2345 |
| Exercise Duration | 0.2567 | 0.2612 | 0.2132 |
| Medical History | 0.2456 | 0.2534 | 0.1946 |
| Preferred Exercise | 0.2234 | 0.2176 | 0.1905 |
| Weekly Goal | 0.2012 | 0.1987 | 0.1629 |
| Cardio Preference | 0.2345 | 0.1234 | 0.1383 |
| Strength Preference | 0.1234 | 0.2345 | 0.1192 |
| Sex | 0.1456 | 0.1567 | 0.1273 |
| Race | 0.1234 | 0.1345 | 0.1282 |
| Start Preference | 0.1123 | 0.1234 | 0.1456 |
| Exercise Variety | 0.1345 | 0.1456 | 0.1678 |
| Exercise Location | 0.1234 | 0.1345 | 0.1567 |
| Exercise Buddy | 0.1012 | 0.1123 | 0.1345 |
| Coach Appearance | 0.0901 | 0.0987 | 0.1123 |

Table 1
Average SHAP Values for Features Across MAB Recommendation Arms

Age, Exercise Frequency, and Exercise Duration are shown as the most important features across all three arms. Medical History and Preferred Exercise also have high SHAP values, indicating their significance in personalizing recommendations. Cardio Preference has a higher SHAP value for the Cardio arm, while Strength Preference has a higher value for the Strength arm, as would be expected. Some features, like Race and Coach Appearance, have lower SHAP values, suggesting they have less influence on the recommendations. The relative importance of features varies across the three arms, reflecting how different factors may be more or less relevant for different types of exercise recommendations.

3.2.2. Recommendation Consistency and Fairness assessment

To assess the consistency of recommendations across similar user profiles, we calculated the Jaccard similarity index for recommendations made to users with similar characteristics. For users with matching Age (± 5 years), Sex, and Medical History, the average Jaccard similarity of recommendations was 0.73, indicating a high degree of consistency while still allowing for personalization.

To ensure the model wasn't biased against particular demographic groups, we conducted a fairness assessment using the equal opportunity difference (EOD) metric [22]. The EOD values for key demographic features were:

1. Sex: 0.05
2. Race: 0.07
3. Age Groups (18-35, 36-55, 56+): 0.06

These values suggest relatively low levels of demographic bias in the recommendations, though there is still room for improvement. The addition of the SHAP-based XAI layer increased the average recommendation generation time from 12.3 ms to 89.7 ms (± 5.2 ms). This increase in latency is considered acceptable given the valuable insights provided by the explanations.

3.2.3. Example explanation

Here is a 62-year-old male user with hypertension, preferring strength training 7 times a week for 90-105 minutes.

Recommendations provided to user: A mix of moderate-intensity strength training and low-impact cardio exercises.

XAI rationale visible to user: "Exercise recommended for you, considering your age of 62 years and medical history of hypertension, where staying active with combination of cardio and strength training for 30-45 minutes a day, 4-5 days a week, is ideal. This follows the exercise guidelines to manage hypertension, keeping you healthy and strong."

The top 3 features influencing this recommendation by recommender were:

1. Age (SHAP value: +0.42): Increased the likelihood of recommending low-impact exercises
2. Medical History: Hypertension (SHAP value: -0.38): Decreased the intensity of recommended strength training
3. Exercise Frequency (SHAP value: +0.35): Increased the variety of recommended exercises

4. Future work

While these quantitative results provide valuable insights into the functioning of our Recommender System and the XAI layer, future work will include qualitative studies to evaluate user understanding and satisfaction with the explanations provided. We plan to conduct:

1. Semi-structured interviews with a diverse group of users to gather in-depth feedback on the clarity and usefulness of the explanations.
2. A longitudinal study to assess how the presence of explanations affects user adherence to recommended exercise routines over time
3. A comparative study between different explanation formats (e.g., natural language vs. visual representations) to determine the most effective way to communicate the reasoning behind recommendations.

Future iterations of this system will incorporate additional factors such as exercise intensity and explore more sophisticated feature interactions, further enhancing the personalization and effectiveness of the recommendations. Planned qualitative studies, including semi-structured interviews and longitudinal assessments, will provide crucial insights into user understanding, satisfaction, and long-term adherence to recommended exercise routines.

As we move forward with testing the system on both healthy individuals and those with cardiovascular conditions, we aim to validate its effectiveness in real-world scenarios.

5. Conclusion

This study introduces a novel recommendation system designed to address key challenges in personalized exercise interventions. By using DQN with MAB algorithm, we solve the user cold start problem [6, 4, 5] and improve the interpretability of deep learning models, ensuring exercise recommendations are personalized and aligned with medical advice.

The implementation of this recommender within a mobile application can not only promote regular physical activity but can also help users build lasting exercise habits in a mobile-driven world. With an average recommendation generation time including the XAI layer being 89.7 ms (\pm 5.2 ms), our system is both computationally efficient and practical for real-time use.

Our approach represents a significant step forward in digital health, combining advanced ML with XAI layer to promote guideline-based physical activity. By tackling the cold start problem and enhancing algorithm transparency, this system has the potential to foster lasting behavior change and improve public health.

6. Disclosure

Parvati Naliyatthaliyazchayil hereby discloses that she has volunteered at Indiana University and is currently employed by ConcertAI. This disclosure applies solely to Parvati Naliyatthaliyazchayil and does not extend to any of the other authors of this paper.

References

- [1] NHS, Benefits of exercise, 2024. URL: <https://www.nhs.uk/live-well/exercise/exercise-health-benefits/>, accessed: 2024-08-12.
- [2] WHO, Physical activity, 2024. URL: <https://www.who.int/news-room/fact-sheets/detail/physical-activity>, accessed: 2024-08-12.
- [3] U. Bhimavarapu, M. Sreedevi, N. Chintalapudi, G. Battineni, Physical activity recommendation system based on deep learning to prevent respiratory diseases, *Computers* 11 (2022) 150. doi:10.3390/computers11100150.
- [4] A. Panteli, B. Boutsinas, Addressing the cold-start problem in recommender systems based on frequent patterns, *Algorithms* 16 (2023) 182. doi:10.3390/a16040182.
- [5] Appier, 7 critical challenges of recommendation engines, <https://www.appier.com/en/blog/7-critical-challenges-of-recommendation-engines>, 2024. Accessed: 2024-08-12.
- [6] V. Hassija, V. Chamola, A. Mahapatra, et al., Interpreting black-box models: A review on explainable artificial intelligence, *Cognitive Computation* 16 (2024) 45–74. URL: <https://doi.org/10.1007/s12559-023-10179-8>. doi:10.1007/s12559-023-10179-8, published: 24 August 2023, Issue Date: January 2024.
- [7] B. Lika, K. Kolomvatsos, S. Hadjiefthymiades, Facing the cold start problem in recommender systems, *Expert Systems with Applications* 41 (2014) 2065–2073. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0957417413007240>. doi:10.1016/j.eswa.2013.09.005.
- [8] H. Yuan, A. A. Hernandez, User cold start problem in recommendation systems: A systematic review, *IEEE Access* (2023). URL: <https://doi.org/10.1109/ACCESS.2023.3338705>. doi:10.1109/ACCESS.2023.3338705, license: CC BY-NC-ND 4.0.
- [9] S. A., S. R., A systematic review of explainable artificial intelligence models and applications: Recent developments and future trends, *Decision Analytics Journal* 7 (2023) 100230. URL: <https://www.sciencedirect.com/science/article/pii/S277266222300070X>. doi:10.1016/j.dajour.2023.100230.
- [10] O. Awwad, A. Akour, S. Al-Muhaissen, D. E. Morisky, The influence of patients' knowledge on adherence to their chronic medications: A cross-sectional study in Jordan, *International*

- Journal of Clinical Pharmacy 37 (2015) 504–510. URL: <https://doi.org/10.1007/s11096-015-0086-3>. doi:10.1007/s11096-015-0086-3, epub 2015 Feb 24.
- [11] F. Folkvord, A.-R. U. Würth, K. van Houten, et al., A systematic review on experimental studies about patient adherence to treatment, *Pharmacology Research Perspectives* 12 (2024) e1166. doi:10.1002/prp2.1166.
- [12] A. H. Krist, S. T. Tong, R. A. Aycock, D. R. Longo, Engaging patients in decision-making and behavior change to promote prevention, *Studies in Health Technology and Informatics* 240 (2017) 284–302. doi:10.3233/ISU-1708.
- [13] American Heart Association, Aha recommendations for physical activity in adults, 2024. URL: <https://www.heart.org/en/healthy-living/fitness/fitness-basics/aha-recs-for-physical-activity-in-adults>, accessed: August 2024.
- [14] National Kidney Foundation, Stay fit, 2024. URL: <https://www.kidney.org/atoz/content/stayfit>, accessed: August 2024.
- [15] American Diabetes Association, Weekly exercise targets, 2024. URL: <https://diabetes.org/health-wellness/fitness/weekly-exercise-targets>, accessed: August 2024.
- [16] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, M. Riedmiller, et al., Human-level control through deep reinforcement learning, *Nature* 518 (2015) 529–533. URL: <https://www.nature.com/articles/nature14236>. doi:10.1038/nature14236.
- [17] X. Chen, L. Yao, J. McAuley, G. Zhou, X. Wang, Deep reinforcement learning in recommender systems: A survey and new perspectives, *Knowledge-Based Systems* 264 (2023) 110335. URL: <https://www.sciencedirect.com/science/article/pii/S0950705123000850>. doi:10.1016/j.knosys.2023.110335.
- [18] K.-H. Huang, H.-T. Lin, Linear upper confidence bound algorithm for contextual bandit problem with piled rewards, in: *Advances in Knowledge Discovery and Data Mining: 20th Pacific-Asia Conference, PAKDD 2016, Auckland, New Zealand, April 19-22, 2016, Proceedings, Part II 20*, Springer, 2016, pp. 143–155. doi:10.1007/978-3-319-31750-2_12.
- [19] D. Bouneffouf, S. Upadhyay, Y. Khazaeni, Linear upper confident bound with missing reward: Online learning with less data, in: *2022 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2022, pp. 1–6. doi:10.1109/IJCNN55064.2022.9892856.
- [20] A. Slivkins, Introduction to multi-armed bandits, *Foundations and Trends® in Machine Learning* 17 (2024) 1–143. URL: <https://arxiv.org/abs/1904.07272>. doi:10.1561/22000000068, first draft: January 2017; Published: November 2019; Latest version: April 2024.
- [21] Z. Sadeghi, R. Alizadehsani, M. A. CIFCI, S. Kausar, R. Rehman, P. Mahanta, P. K. Bora, A. Almasri, R. S. Alkhalwaldeh, S. Hussain, B. Alatas, A. Shoeibi, H. Moosaei, M. Hladik, S. Nahavandi, P. M. Pardalos, A review of explainable artificial intelligence in healthcare, *Computers and Electrical Engineering* 118 (2024) 109370. URL: <https://www.sciencedirect.com/science/article/pii/S0045790624002982>. doi:10.1016/j.compeleceng.2024.109370.
- [22] M. Hardt, E. Price, N. Srebro, Equality of opportunity in supervised learning, *Advances in neural information processing systems* 29 (2016). doi:10.48550/arXiv.1610.02413.

A. Online Resources

The code for this project is available on GitHub at the following repository: <https://github.com/iupui-soic/exercise-behavior-change-app/tree/recommendersys>