Hindawi
Scientific Programming
Volume 2022, Article ID 7089057, 8 pages
https://doi.org/10.1155/2022/7089057

Research Article

Research on Sports Dance Video Recommendation Method Based on Style

Jiangtao Sun[1] and Haiying Tang[2]

[1]*Wenhua College, Wuhan 430074, China*
[2]*Wuhan Institute of Physical Education, Wuhan 430079, China*

Correspondence should be addressed to Haiying Tang; 2004068@whsu.edu.cn

Received 23 March 2022; Revised 4 April 2022; Accepted 16 April 2022; Published 6 May 2022

Academic Editor: Jie Liu

At present, sports dance teaching still tends to "demonstration" training. Students have limited time and space for autonomous learning, and their enthusiasm for participation is not high, which leads to a decline in classroom learning efficiency. In view of this, video teaching has become popular in sports dance classrooms, providing a new model for sports dance teaching. Video recommendation is particularly important for the improvement of teaching quality. A sports dance video recommendation method based on style is proposed. The factorization machine model is used to combine features and process high-dimensional sparse features, the deep neural network model is adopted as the value function network of the deep Q-learning algorithm, and the deep Q-learning algorithm is used as the decision function to solve the recommendation accuracy and diversity question. Through the application experiment of sports dance video recommendation, it is resulted that the recommendation accuracy of the proposed model is slightly higher than that of traditional recommendation algorithm and the recommendation diversity is obviously better than that of traditional recommendation algorithm. The advantages and feasibility of the proposed model are verified.

1. Introduction

Sports dance is a way of sports, which needs refining, organization, arrangement, and technical processing. Its main mean of the expression is to show the flexible footwork and beautiful dance posture of various rhythms of the human body, and express feelings and reflect social life through this artistic form [1–3]. It is mainly composed of 10 dance types in two series: modern dance and Latin dance. With the progress of society, there has been an upsurge of national fitness, and there has been an upsurge of popularizing and popularizing sports dance in the society. Due to the accelerated pace of life today, many people are busy with work and hope to enrich their spare time activities. However, because dance learning is usually taught by participating in training courses, many people do not have much time to learn. Therefore, it has become a trend to teach students sports dance by recording videos and video recommendations.

The traditional manual teaching process of sports dance is mainly divided into two parts: teachers' explanation of theoretical knowledge and teaching purpose, learning, and guidance of practical courses [4–7]. The second part is the focus of learning. It includes teachers' dance demonstration performance and explanation, correct requirements and practice of dance posture, decomposition practice of turning step and flower step, mutual cooperation between male and female partners, dance practice, dance appreciation, harmonious practice of music, etc.

The sports dance video recommendation system actually uses the existing computer technology to present the process of sports dance teaching to students in the form of video recommendation and applies the recorded sports dance teaching video course to real teaching [5, 7]. The design and development of sports dance teaching video recommendation system make the teacher's teaching content displayed in the form of video. Through the real-time recording of the

teacher's dance scene, students can learn more clearly and quickly. At the same time, it also plays a positive role for students to watch the teaching process repeatedly and take care of students at different levels. In addition, through computer technology, the video teaching of sports dance can also be presented to the students in the form of forums or discussion groups so that the students can no longer learn at a single point, but can learn and communicate with other students in time through the network, learn from each other's strengths, and make up for their weaknesses, to promote communication among students, improve students' learning interest, and stimulate students' learning enthusiasm [7].

It can be seen that sports dance video recommendation is of great significance to improve the quality of sports dance teaching. The research on sports dance video recommendation method not only solves the learning constraints of time and space in the process of sports dance teaching but also makes the teaching process reproducible and decomposed, and provides support for the development of video teaching.

2. Related Works

Recommendation systems [8, 9] have gradually produced a variety of solutions and become an independent discipline now. Many achievements have been achieved in industry applications. The recommendation system itself analyzes and studies the behavioral preferences of users through data information mining and establishes a user-specific interest model, thereby recommending information that may satisfy their interests. Traditional recommendation system algorithms can be roughly divided into three types: collaborative filtering recommendation, content-based recommendation, and hybrid recommendation.

Although traditional recommendation algorithms can solve most information filtering problems, they cannot solve the problems of data sparseness, cold start, and repeated recommendation problems. In recent years, many companies have used deep learning, multi-arm gambling machines, and other algorithms to improve and have obtained good recommendation results in response to the above issues. YouTube [10] used deep learning for video recommendation prediction for the first time in the recommendation system. It successfully filtered and extracted the video content users were interested in from the large-scale data volume and recommended it. Acar et al. [11] proposed an offline evaluation method and controlled experiment based on streaming data. Karatzoglou et al. [12] systematically proposed to apply deep learning to traditional recommendation systems, adding deep learning to the conventional content recommendation and collaborative filtering recommendation methods to deal with recommendation prediction of large-scale data volume. Therefore, deep learning has become a hot spot in current recommendation system research.

At present, most of the recommendation algorithms are based on the static recommendation process and generate a fixed recommendation strategy by collecting and processing a large amount of data information, such as multicriteria decision-making method [13–22], which has a significant improvement in solving the diversification of information recommendation. The problem of cold start is also unable to adapt to the short-term interest changes of users and make effective information recommendations. Therefore, many scholars began to try to use reinforcement learning [23] to solve the problems in the recommendation system. Reinforcement learning is a learning algorithm based on the interaction of the environment. It has developed independently from the two fields of animal behavior research and optimal control. It has been abstracted and formalized as a Markov decision process problem. Later, through the study of many scientists, a relatively complete system, approximate dynamic programming was formed. Reinforcement learning [23] is a dynamic interactive learning strategy algorithm. Therefore, reinforcement learning is used to solve cold start problems and the inability to adapt to users' short-term interest recommendation in recommendation algorithms. Taghipour et al. [24] first proposed to use the Q-learning algorithm combined with web page information to solve the problem of web page recommendation. However, the Q-learning learning algorithm cannot effectively solve the recommendation task of the Markov decision process with large state space and action space. Choi [25] proposed a biclustering learning algorithm to alleviate the above problems, but the effect was not expected. The deep Q-learning deep reinforcement learning algorithm [26] used the value function estimation method, and it solves the problems existing in the Markov decision-making process [27] by iterating the Bellman equation to achieve convergence to the optimal value function. The proposal of policy gradient solves the problem that the value is difficult to calculate, and this method can directly learn the policy.

3. Markov Decision Process

Markov decision process is a mathematical model of sequential decision. It is used to simulate the randomness strategy and return that can be realized by agents in the environment where the system state has Markov nature.

Markov decision process is built based on a set of interactive objects, that is, agents and environment. The elements include state, action, strategy, and reward. In the simulation of Markov decision process, the intelligent experience perceives the current system state and acts on the environment according to the strategy, so as to change the state of the environment and get rewards. The accumulation of rewards over time is called reward.

The theoretical basis of Markov decision process is Markov chain, so it is also regarded as a Markov model considering action. The Markov decision process established in discrete time is called "discrete-time Markov decision process"; on the contrary, it is called "continuous time Markov decision process." In addition, Markov decision process has some variants, including partially observable Markov decision process, constrained Markov decision process, and fuzzy Markov decision process [27].

The factorization machine algorithm is used to combine features and deal with high-dimensional sparse features,

effectively learn the cross-hidden relationship between features, and then use the deep Q-learning algorithm to solve the optimal value of the recommendation decision. First, the core point of sports dance video recommendation is to simulate the recommendation process as a Markov decision process. The initial state of the agent is s_0, and then, an action a_0 is selected from the action set to execute. After execution, the agent will follow the action a_0. The agent s_0 changes to the next state s_1 according to the reward function of action a_0, and then, the action a_1 is selected, and the above steps are continuously looped until a strategy chain reaches the reward accumulation value, which is selected.

Combined with sports dance video recommendation, the Markov solution process can be more refined. The user and the recommendation system can be simulated as two dynamic interactive objects, as shown in Figure 1. In a time slice t, the recommendation system obtains the user's viewing record s_t, trains the action set A through the reward function of s_t, and then selects the sports dance video with the highest reward value in A to recommend to the user, obtain the user's rating for the video, and put it into the experience pool to continue training the recommendation strategy.

In general, the Markov decision process is a Markov reward process with decision making, which means that all states have Markov properties, that is, when a random process is given a current state and all past states. In the case of the conditional distribution probability of its future state depend only on the current state, the Markov property can be expressed in mathematical form as a state s_t has Markov property if and only if it satisfies

$$P[S_{t-1}|S_t] = P[S_{t+1}|S_1, \ldots, S_t]. \tag{1}$$

A Markov decision process can consist of quintuples $\langle S, A, P, R, \gamma \rangle$.

S is the set of all environmental states, and $s_t \in S$ represents the current Agent's state s_t at time t. The evaluation of sports dance video is used for testing, and s is defined as $s_t = \{\text{movie}_t^1, \ldots, \text{movie}_t^n\}$, here denoted as the top n videos watched and rated by the user.

A is the set of limited executable actions of the agent, and A is the set of all videos recommended to the user. The text $a_t \in A$ is represented as the recommended action obtained by the agent through the reward function training at time t, that is, the sports dance video advised by the recommendation system agent to the user through s_t at the current time.

P is the state transition probability matrix, and the mathematical formula is as follows:

$$p_{ss'}^{a_t} = P\left[S_{t+1} = s\prime | S_t = s, A_t = a_t\right], \tag{2}$$

where S_t represents the state at time t, S_{t+1} represents the state at time $t+1$, and A_t represents the actions of different videos recommended by the recommendation system agent to the user at time t. When the time t ends, it turns to time $t+1$, and then, the recommendation system agent will update the state S_{t+1} at time $t+1$ to $s_{t+1} = \{\text{movie}_t^1, \ldots, \text{movie}_t^n, a_t\}$.

R is the reward function of the current Markov reward process [10]. At time t, state s_t is the reward expectation obtained by A entering state s_{t+1}. The mathematical formula is defined as follows:

$$R_s = E[R_{t+1}|S_t = s]. \tag{3}$$

For the videos recommended by the recommendation system agent, users have different scores of movie ratings to generate various movie feedback, so at time t, the recommendation system agent will obtain instant rewards according to varying feedback as $R_t = (s_t, a_t)$. γ is the discount factor, and its value range is generally specified as $\gamma \in (0, 1)$. The discount factor is used to adjust the impact of future rewards on the current accumulated rewards. If $\gamma = 0$, it means that the recommendation system agent pays more attention to the earned reward in time; when $\gamma = 1$, it means that the recommendation system agent pays more attention to the long-term accumulated reward.

From the above definition, it can be concluded that the task of the recommendation system agent is to achieve the learning process of maximizing the reward function through the optimal recommendation strategy [11]. At time t, the mathematical definition of reward G_t is as follows:

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}. \tag{4}$$

4. Algorithm

4.1. Concepts. One-hot coding combined with the FM model is used to preprocess the data, and then, the deep Q-learning algorithm in deep reinforcement learning is used as the sports dance video recommendation algorithm. The deep Q-learning algorithm is based on the approximate iteration of the value function. It uses a deep neural network as a Q-value network to extract complex features. One-hot encoding can represent the discrete features in the dataset with numbers. Still, one-hot encoding will introduce the problem of sparse features, so the factorization machine algorithm is used to solve the problem of difficulty in dealing with combined features under the condition of light features.

The ultimate goal of using the deep Q-learning algorithm for a sports dance video recommendation is to obtain the optimal recommendation strategy by maximizing the long-term cumulative reward. The first step is to perform feature extraction on the original data and use the user's rating for the movie as the reward value. Figure 2 shows the corresponding original data structure. Due to the existence of discrete features in the original data set, the deep Q-learning algorithm cannot directly use it as the input data of the value function network.

Therefore, one-hot coding encodes the discrete features and expands the discrete features into the Euclidean space. The values of discrete features can correspond one-to-one with points in Euclidean space. From the aspect of model training, not only can the distance between different features be calculated more reasonably, but also can the nonlinear ability of the model be improved. The principle of one-hot

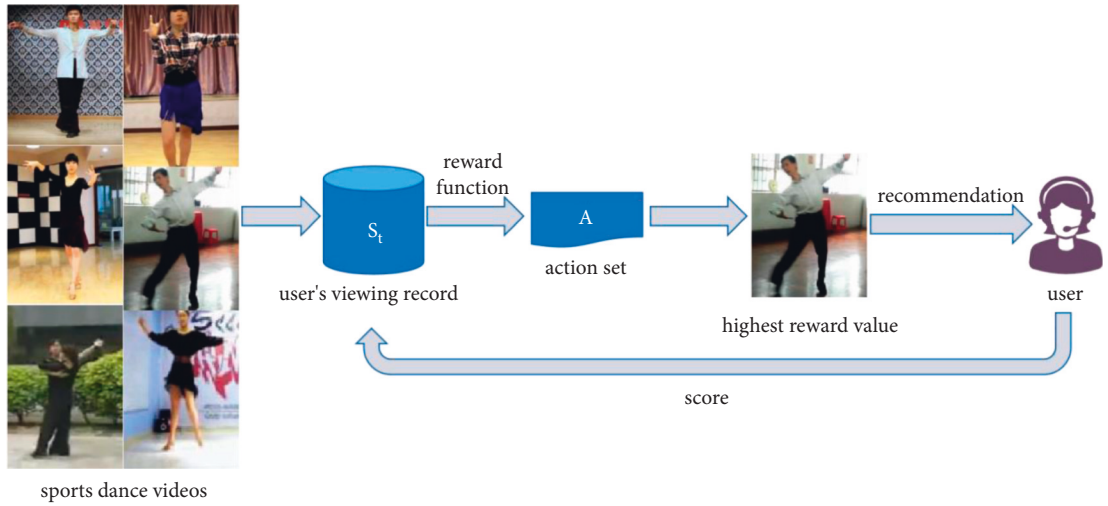

Figure 1: Reinforcement learning basic architecture.

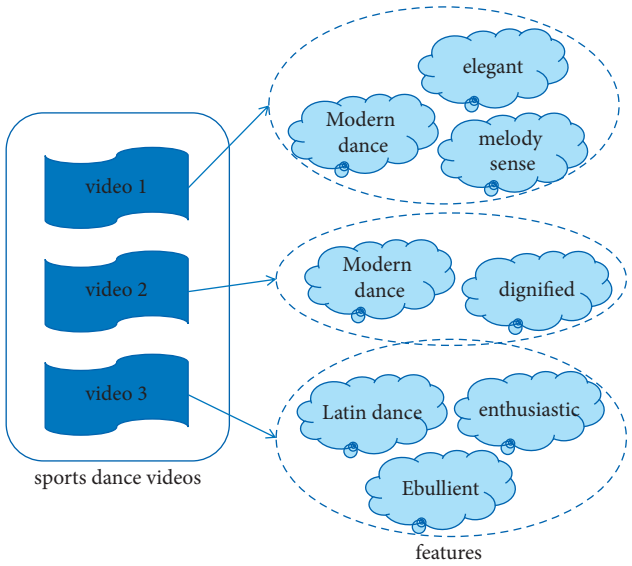

Figure 2: Sports dance video and its features.

encoding is to use the N-bit status register to encode n states with different codes, each attribute has its independent register bit, and only one bit indicates whether it is valid or not at any time. As shown in Table 1, the feature set is processed by using one-hot encoding.

Since one-hot encoding will introduce the problem of feature sparseness, which will cause the dimensional disaster of the neural network, the factorization machine algorithm is used to perform further feature processing on the feature set. A second-order polynomial model is used. A combination of features x_i and x_j is used, where $x_i x_j$ represents the combined feature, x_i is the value of the i^{th} feature, and n represents the number of features of the sample. w_0, w_i, and w_{ij} are model parameters, respectively. The second-order polynomial model is as follows:

$$y(x) = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w_{ij} x_i x_j. \tag{5}$$

Matrix decomposition is used to solve w_{ij}. It is known that in model-based collaborative filtering, the user matrix and the sports dance video matrix can form a unique rating matrix. For each user and video, a hidden vector can be used to represent it. Different users and videos are represented as different two-dimensional vectors. The dot product of the user vector and the movie vector is the user's rating of the movie in the matrix.

It can be seen from the definition that for any $N \times N$ real symmetric matrix, this real symmetric matrix has N linearly independent and can be orthogonalized to get a set of eigenvectors, which are orthogonal and have a module of 1. So, the real symmetric matrix A can be decomposed into

$$A = \varsigma \wedge \varsigma^T, \tag{6}$$

where $\wedge$ is defined as a real diagonal matrix, and Q is defined as an orthogonal matrix.

Similarly, suppose there is asymmetric matrix W consisting of all the current quadratic parameters w_{ij}; in that case, this matrix can be decomposed in the form of $W = V \wedge V^T$, where the jth column of V is defined as the latent vector of the jth dimension feature. The inner product of the latent vector corresponding to xi and the latent vector corresponding to x_j is equal to the cross-term coefficient of the feature components x_i and x_j, so each parameter of the symmetric matrix can be defined as $w_{ij} = \langle v_i, v_i \rangle$.

The deep Q-learning algorithm uses a deep neural network with a weight parameter of θ as the network model of this deep neural network's action-value function. The weights and biases are represented by θ and γ, respectively. The loss function of the deep neural network model is as follows:

$$L_i(\theta_i) = E\left[\left(r + \gamma \max Q\left(s\prime, a\prime, \theta_i\right) - Q(s, a, \theta_i)\right)\right]. \tag{7}$$

The structure diagram of the deep neural network is shown in Figure 3.

It can be seen from the structure diagram that the deep neural network consists of an embedding layer and three

Table 1: Sports dance video feature set based on one-hot encoding.

Sports dance videos	Features						
	Modern dance	Latin dance	Elegant	Melody sense	Dignified	Enthusiastic	Ebullient
Video 1	1	0	1	1	0	0	0
Video 2	1	0	0	0	1	0	0
Video 3	0	1	0	0	0	1	1

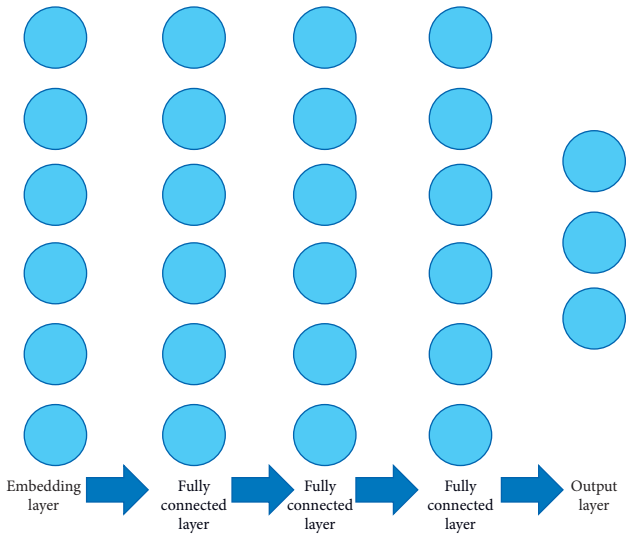

Figure 3: Deep neural network structure.

fully connected layers, and each neuron will use the activation function for calculation before outputting the result. Sigmoid function is used as the activation function of the deep neural network to meet the requirements of effectively learning sparse features.

To facilitate the calculation, the number of neuron nodes in the embedding layer and the fully connected layer is set to the exponential power of 2. The depth of the network is increased in multiples of 2 in turn. Assuming that the number of sports dance videos in the recommendation system is M, the number of output nodes is M, and the output node will output the predicted reward value of each video after it is recommended.

4.2. Deep Q-Learning Algorithm for Sports Dance Video Recommendation. Due to using the same network to generate the following target Q and estimate the current Q, it can lead to oscillations and even divergence. Therefore, deep Q-learning uses experience replay and target network methods to solve this problem.

Experience playback means that during the interaction between the agent and the environment, the experience is stored in the experience pool D. Each training will randomly sample a small batch of data from D for training to eliminate the correlation between samples. Its function is to destroy the correlation between the series and solve the correlation between the Q-value and the target Q-value. The target network does not interact with the environment, nor is it updated at every step, only at certain stages. Each update assigns the current network parameters directly to it.

The process of the deep Q-learning movie recommendation algorithm is as follows:

(1) Initialize experience pool D with capacity N, which is used for historical experience recovery. Use the deep neural network as the network model of the current predicted action-value function Q value, and initialize the weight parameter θ of the network model. Set the number of rounds of model training as M, the maximum number of training times the agent can perform. Initialize the input of the Q-value network model, information the scoring matrix processed by the FM algorithm, and calculate $\varphi 1 = \varphi(s1)$.

(2) Repeat the single empirical trajectory time step from $t = 1$ to T.

(3) Repeat the sports dance video recommendation training for each user from $u = 1$ to U.

(4) Take the user's initial rating as the initial movie recommendation state S, and select a random movie plan with probability ε for recommendation.

(5) If the recommended movie is a movie that the user likes, update the current movie recommendation state as

$$S_t + 1 = S_t \cup a. \tag{8}$$

(i) Then set the reward to $r = 1$, and calculate the input sequence as follows:

(6) If the recommended movie is a movie that the user likes, update the current movie recommendation status to $S_{t+1} = S_t$, set the reward to $r = 0$, and add $(\varphi_t, a, r, \varphi_{t+1})$ to the experience pool D, and compute the input sequence $\varphi_{t+1} = \varphi(S_{t+1})$ for the next time step.

(7) Randomly sample a small batch of stored samples $(\varphi_t, a, r, \varphi_{t+1})$ from the experience pool D.

(8) If the current state is an end state, set yi = ri, if the current state is a nonend state, then set

$$y_i = r_i + \gamma \max_{a'} Q\left(\phi_{j+1}, a', \theta\right). \tag{10}$$

(9) Calculate loss function using gradient descent algorithm

$$L_i(\theta_i) = E\left[\left(r + \gamma \max_{a'} Q\left(\phi_{j+1}, a', \theta\right) - Q(s, a, \theta_i)\right)\right]. \tag{11}$$

(10) Output value function network.

The flow chart of the algorithm is shown in Figure 4.

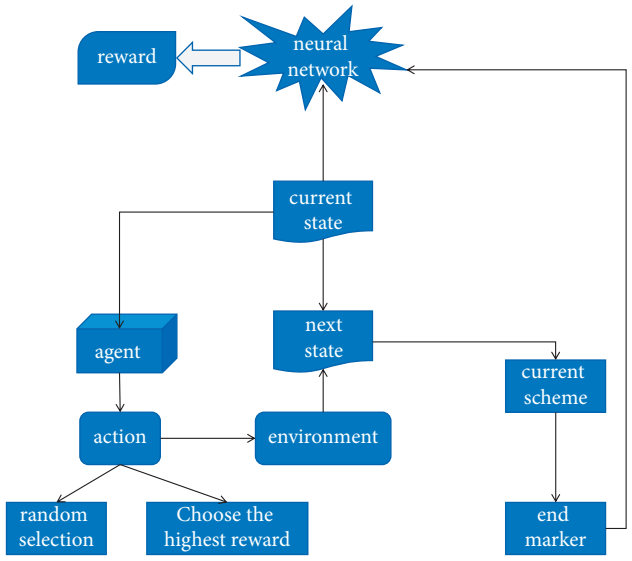

Figure 4: Sports dance video recommendation algorithm.

5. Case Study

Experimental hardware environment is one PC, 8 GB memory, and 256 GB hard disk. The operating system of the testing platform is Windows 10, the programming language is python language, and the programming is carried out on the Jupyter Notebook. The experiment uses the dataset of sports dance video evaluation, including 10000 rating records for 100 sports dance videos by 400 users (teachers and students).

To verify that the video recommendation algorithm based on deep reinforcement learning is better than the traditional collaborative filtering algorithm, we will compare the deep Q-learning algorithm (DQLA) and the collaborative filtering algorithm (CFA) on recommendation accuracy and video recommendation diversity at the same dataset and then analyze the two algorithms.

The recommendation accuracy in the experiment is defined as the proportion of users' scores for sports dance videos, and the diversity of videos recommendations in the experiment is defined as the proportion of the types of recommended videos in the total types.

According to the experimental design, the accuracy of the two algorithms is compared. The comparison results are shown in Table 2.

A comparison of recommendation accuracy of DQLA and CFA is shown in Figure 5.

It can be seen from the comparison chart that when the recommendation number of sports dance videos is less than 45, the accuracy of DQLA is slightly lower than that of the traditional CFA. Still, its recommendation effect shows an upward trend. When the recommendation number of sports dance videos is more than 45, the recommendation accuracy of DQLA is obviously better than that of CFA; when the peak of the optimal number of recommendations is reached, the recommendation accuracy gradually decreases, but it is also higher than the recommendation accuracy of CFA under the same conditions. This shows that under the premise of increasing the number of video recommendations, the recommendation accuracy of DQLA is significantly higher than that of CFA. Therefore, it is proved that using the deep Q-learning algorithm of deep reinforcement learning for sports dance video recommendation is beneficial to improve the accuracy of recommendation.

Table 2: Recommendation accuracy of DQLA and CFA.

Number of sports dance videos	Recommendation accuracy	
	DQLA	CFA
15	0.1609	0.2122
25	0.1768	0.2036
35	0.1877	0.2001
45	0.1956	0.1944
55	0.2066	0.1912
65	0.2102	0.1867
75	0.1978	0.1841
85	0.1861	0.1812

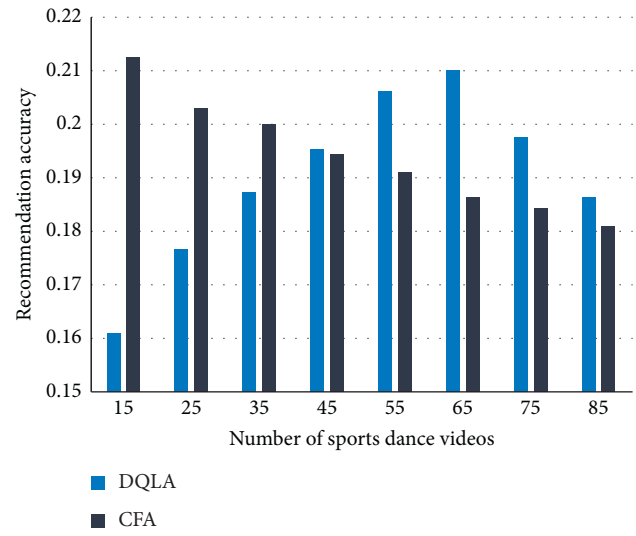

Figure 5: A comparison of recommendation accuracy of DQLA and CFA.

According to the experimental design, the recommendation diversity of the two algorithms is compared. The comparison results are shown in Table 3.

A comparison of recommendation diversity of DQLA and CFA is shown in Figure 6.

It can be seen from the comparison chart that DQLA always has a higher diversity ratio of sports dance video recommendation than CFA in the sample interval, and its diversity recommendation effect has been on the rise. This shows that when the deep Q-learning algorithm of deep reinforcement learning is used for sports dance video recommendation, the diversity effect of video recommendation has been significantly improved.

Table 3: Recommendation diversity of DQLA and CFA.

Number of sports dance videos	Recommendation diversity	
	DQLA	CFA
15	0.4711	0.4322
25	0.4801	0.4388
35	0.4878	0.4431
45	0.4892	0.4492
55	0.4913	0.4551
65	0.4934	0.4580
75	0.4966	0.4652
85	0.4981	0.4698

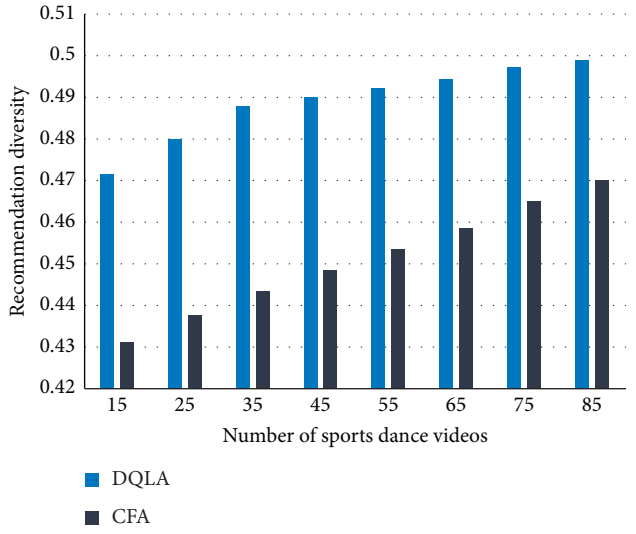

Figure 6: A comparison of recommendation diversity of DQLA and CFA.

6. Conclusion

By using the deep reinforcement learning method, the deep Q-learning algorithm is used for sports dance video recommendation training and compared with the collaboration filtering algorithm in the accuracy and the diversity of sports dance video recommendation. The comparison results show that the accuracy of sports dance video recommendation is better than the traditional collaboration filtering algorithm when the number of sports dance video recommendation reaches a specific number, and the diversity of sports dance video recommendation is obviously better than that of the collaboration filtering algorithm. It can be proved that the application of deep reinforcement learning to sports dance video recommendation can effectively solve the problems of inaccurate recommendation of traditional recommendation algorithms and single recommendation content. At the same time, the deep reinforcement learning algorithm can better learn the user's interest characteristics to provide a better video recommendation solution for the user. However, the deep Q-learning algorithm cannot solve the cold-start video recommendation problem. Therefore, next we will be to study how the recommendation system can still accurately locate the user's video of interest when there is no user's video viewing history data.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the 2020 Provincial Teaching Research Project of Hubei "Practice and research on 'online and offline' mixed teaching of sports dance" (No. 2020736).

References

[1] D. Sekulic, R. Kostic, and D. Miletic, "Substance use in dance sport," *Medical Problems of Performing Artists*, vol. 23, no. 2, pp. 66–71, 2008.

[2] N. Barker-Ruchti, "Sport, dance and embodied identities," *Sport, Education and Society*, vol. 11, no. 2, pp. 195–197, 2006.

[3] E. Wanke, T. Fischer, H. G. Pieper, and D. Groneberg, "Dance sport: injury profile in Latin American formation dancing," *Sportverletzung - Sportschaden: Organ der Gesellschaft fur Orthopadisch-Traumatologische Sportmedizin*, vol. 28, no. 03, pp. 132–138, 2014.

[4] P. Markula, "The intersections of dance and sport," *Sociology of Sport Journal*, vol. 35, no. 2, pp. 159–167, 2018.

[5] A. Y. Chu and C.-H. Wang, "Differences in level of sport commitment among college dance sport competitors," *Social Behavior and Personality: An International Journal*, vol. 40, no. 5, pp. 755–766, 2012.

[6] M. Skwiot, Z. Śliwiński, and G. E. Śliwiński, "Perfectionism and burnout in sport and dance," *Physikalische Medizin, Rehabilitationsmedizin, Kurortmedizin*, vol. 59, no. 03, pp. 135–140, 2020.

[7] A. F. Zhao, "Sports dance teaching based on virtual environmen," *Basic and Clinical Pharmacology and Toxicology*, vol. 127, p. 244, 2020.

[8] M. Robillard, R. Walker, and T. Zimmermann, "Recommendation systems for software engineering," *IEEE Software*, vol. 27, no. 4, pp. 80–86, 2010.

[9] R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins, "Recommendation systems: a probabilistic analysis," *Journal of Computer and System Sciences*, vol. 63, no. 1, pp. 42–61, 2001.

[10] Y. Gao, S. Chen, and X. Lu, "Research on reforcement learning technology: a review," *Acta Automatica Sinica*, vol. 30, no. 1, pp. 86–100, 2004.

[11] E. Acar, F. Hopfgartner, and S. Albayrak, "Fusion of learned multi-modal representations and dense trajectories for emotional analysis in videos," in *Proceedings of the 2015 13th International Workshop on Content-Based Multimedia Indexing (CBMI)*, Prague, Czech Republic, June 2015.

[12] A. Karatzoglou and B. Hidasi, "Deep learning for recommender systems," in *Proceedings of the 11th ACM Conference on Recommender Systems (RecSys)*, Como, Italy, 2017.

[13] L. Li, C. Mao, H. Sun, Y. Yuan, and B. Lei, "Digital twin driven green performance evaluation methodology of intelligent manufacturing: hybrid model based on fuzzy rough-sets AHP, multistage weight synthesis, and PROMETHEE II," *Complexity*, vol. 2020, Article ID 3853925, 24 pages, 2020.

[14] A. A. Ganin, P. Quach, M. Panwar et al., "Multicriteria decision framework for cybersecurity risk assessment and

management," *Risk Analysis: An Official Publication of the Society for Risk Analysis*, vol. 40, no. 1, pp. 183–199, 2020.

[15] L. Li, J. Hang, H. Sun, and L. Wang, "A conjunctive multiple-criteria decision-making approach for cloud service supplier selection of manufacturing enterprise," *Advances in Mechanical Engineering*, vol. 9, no. 3, Article ID 168781401668626, 2017.

[16] M. Sayan, T. Sandlidag, N. Saltanoglu, and B. Uzen, "The use of multicriteria decision-making method—fuzzy VIKOR in antiretroviral treatment decision in pediatric HIV-infected cases - ScienceDirect," *Applications of Multi-Criteria Decision-Making Theories in Healthcare and Biomedical Engineering*, vol. 18, pp. 239–248, 2021.

[17] L. H. Li, J. C. Hang, Y. Gao, and C. Y. Mu, "Using an integrated group decision method based on SVM, TFN-RS-AHP, and TOPSIS-CD for cloud service supplier selection," *Mathematical Problems in Engineering*, vol. 2017, Article ID 3143502, 14 pages, 2017.

[18] L. Li, T. Qu, Y. Liu et al., "Sustainability assessment of intelligent manufacturing supported by digital twin," *IEEE Access*, vol. 8, Article ID 175008, 2020.

[19] G. Sir and E. Sir, "Pain treatment evaluation in COVID-19 patients with hesitant fuzzy linguistic multicriteria decision-making," *Journal of Healthcare Engineering*, vol. 2021, Article ID 8831114, 11 pages, 2021.

[20] L. Li and C. Mao, "Big data supported PSS evaluation decision in service-oriented manufacturing," *IEEE Access*, vol. 8, no. 99, p. 1, 2020.

[21] Y. Li and L. Li, "Enhancing the optimization of the selection of a product service system scheme: a digital twin-driven framework," *Strojniški vestnik - Journal of Mechanical Engineering*, vol. 66, no. 9, pp. 534–543, 2020.

[22] L. Li, B. Lei, and C. Mao, "Digital twin in smart manufacturing," *Journal of Industrial Information Integration*, vol. 26, no. 9, Article ID 100289, 2022.

[23] B. Xin, H. X. Yu, Y. Qin, Q. Tang, and Z. Zhu, "Exploration entropy for reinforcement learning," *Mathematical Problems in Engineering*, vol. 2020, Article ID 2672537, 12 pages, 2020.

[24] N. Taghipour and A. Kardan, "A hybrid web recommender system based on Q-learning.23rd annual ACM symposium on applied computing," *APPLIED COMPUTING*, vol. 1-3, pp. 1164–1168, 2008.

[25] L. C. Choi, J. S. Lee, and S. C. Park, "Double layered genetic algorithm for document clustering," *Communications in Computer and Information Science*, vol. 257, pp. 212–218, 2011.

[26] S. Ohnishi, E. Uchibe, Y. Yamaguchi, K. Nakanishi, Y. Yasui, and S. Ishii, "Constrained deep Q-learning gradually approaching ordinary Q-learning," *Frontiers in Neurorobotics*, vol. 13, 2019.

[27] O. Alagoz, H. Hsu, A. J. Schaefer, and M. S. Roberts, "Markov decision processes: a tool for sequential decision making under uncertainty," *Medical Decision Making*, vol. 30, no. 4, pp. 474–483, 2010.