International Journal of Innovation and Sustainable Development

ISSN online: 1740-8830 - ISSN print: 1740-8822
https://www.inderscience.com/ijisd

Online mobile learning resource recommendation method based on deep reinforcement learning

Pingyang Li, Juan Zhang

DOI: 10.1504/IJISD.2023.10056670

Article History:
Received: 05 August 2022
Last revised: 06 September 2022
Accepted: 31 October 2022
Published online: 02 December 2024

Online mobile learning resource recommendation method based on deep reinforcement learning

Pingyang Li* and Juan Zhang

Changchun College of Electronic Technology,
Changchun, 130000, China
Email: pingyli@mls.sinanet.com
Email: 6531138@qq.com
*Corresponding author

Abstract: In order to improve the recommendation effect of learning resources, this paper designs an online mobile learning resource recommendation method based on deep reinforcement learning. Firstly, the similarity between learners and learning resources, and the similarity between learners' search preference results and learning resources are calculated. Secondly, based on the results of similarity calculation, a multi-agent deep reinforcement learning network is designed, which includes a recommendation agent and a classification agent. Finally, according to the learners' interest preferences (states) of different learning resources, the online mobile learning resources (execution actions) are recommended to the learners, and the final recommendation scheme is obtained through the recommendation agent. According to the experimental results, the maximum recommendation result hit rate of this method is 95.5%, and the highest average ranking degree is 0.926, indicating that the recommendation effect of this method is better.

Keywords: online mobile learning; learning resource; similarity; recommendation agent; classification agent; recommended scheme.

Reference to this paper should be made as follows: Li, P. and Zhang, J. (2025) 'Online mobile learning resource recommendation method based on deep reinforcement learning', *Int. J. Innovation and Sustainable Development*, Vol. 19, No. 1, pp.1–12.

Biographical notes: Pingyang Li received her Bachelor degree of Chinese Language and Literature from Anshan Normal University in 2008. She is an Assistant Professor in Changchun College of Electronic Technology. Her research direction is teaching management.

Juan Zhang received her Master degree of Foreign Linguistics and Applied Linguistics from Jilin University in 2013. She is a lecturer in Changchun College of Electronic Technology. Her research interests include English linguistics and applied linguistics.

1 Introduction

The development of Internet technology has promoted the optimisation of educational environment, accelerated the reform process of traditional learning model, and provided strong technical support for the development of online autonomous learning and mixed teaching model. At present, online mobile learning is becoming more and more popular, which provides an effective way for more learners to learn (Chen and Jing, 2020).

The establishment of online mobile learning resources with rich content is the main driving force for the development of education industry informatisation. At present, the establishment of various online mobile learning platforms provides learners with sufficient and rich learning materials, so that learning is no longer restricted by geographical and time factors (Hu and Zhou, 2021). However, due to the increasing number of learning resources, it is increasingly difficult for learners to choose appropriate learning resources according to their own interests and preferences. Therefore, it has become an urgent priority for online mobile learning platforms to reasonably recommend learning resources that meet the actual needs of learners (Yang et al., 2020).

In Nie (2020), a learning resource recommendation method based on behaviour analysis is designed. Based on the establishment of the learner-learning resource score matrix, the method collects the browsing and downloading behaviour data of learners in real time. After real-time normalised and fused processing of the data, the recommendation list is established through the collaborative filtering process. Although this method is not limited by the number of learning resources, it is easy to confuse data categories in the process of normalised fusion processing, resulting in poor recommendation effect. In Dong (2021), a learning resource recommendation method based on knowledge graph is proposed. After the subject knowledge graph is established, mathematical models are established respectively for learning resources and learners' interests, and the correlation degree of knowledge points covered by learners' interests, knowledge base and learning resources is taken as the optimisation objective. Multi-objective particle swarm optimisation algorithm is used to solve the model and output the recommendation sequence. However, in practice, it is found that this method has the problem of low recommendation result hit rate. In Fan et al. (2020), multi-layer perceptron is used to combine user interest features with matrix vectors of learning resource features, and then an improved deep neural network structure is designed to output the final recommendation result. Although this method has a good effect on the solution of cold start, its recommendation results rank low, resulting in low acceptance of learners.

In view of the above problems, this study designed a new online mobile learning resource recommendation method based on deep reinforcement learning. The design ideas are as follows:

Firstly, the similarity of preference features between different learners, the similarity between different learning resources, and the similarity between learners' search preference results and learning resources are calculated respectively.

Then, a multi-agent deep reinforcement learning network is designed, which includes a recommendation agent and a classification agent. In the classification agent, the fusion results of learner preference features are obtained by inputting learner preference feature test samples. In the recommendation agent, the learner's interests, preferences and download records observed by the recommendation agent are input, and the learning resources to be recommended are output.

Finally, according to the learners' interest preferences (states) of different learning resources, the online mobile learning resources (execution actions) are recommended to the learners, and the final recommendation scheme is obtained through the recommendation agent.

2 Similarity calculation

2.1 Similarity calculation between learners

For online mobile learning resources unknown to users, cosine similarity should be used to measure the similarity between two neighbourhood learners (Zhang et al., 2019a, 2019b). It is known that there are two learners u and v, whose historical scores for an online mobile learning resource are g_u and g_v, respectively. Then the formula for calculating the cosine similarity $sim(u,v)$ between learners is as follows:

$$sim(u,v) = \frac{g_u \times g_v}{\|g_u\|_2 \times \|g_v\|_2} = \frac{\sum_1^{I_c} g_{u,I} \times g_{v,I}}{\sqrt{\sum_1^{I_u} g_{u,I}{}^2} \times \sqrt{\sum_1^{I_v} g_{v,I}{}^2}} \tag{1}$$

In equation (1), for online mobile learning resource I, the scores of learner u and learner v are $g_{u,I}$ and $g_{v,I}$, respectively, the set of learning resources scored by each learner is I_u and I_v respectively, and the intersection of learning resources scored by the two learners is I_c.

In order to more accurately obtain the similarity between learners relative to the learning resource I_c that has been scored by both learners, the following similarity expression can be used:

$$sim(u,v)_{I_c} = \frac{\sum_1^{c} g_{u,c} \times g_{v,c}}{k \times \sqrt{\sum_1^{I_u} g_{u,c}{}^2} \times \sqrt{\sum_1^{I_v} g_{v,c}{}^2}} \tag{2}$$

In equation (2), the two learner scores about learning resource c are $g_{u,c}$ and $g_{v,c}$ respectively, and k represents the superposition coefficient.

2.2 Computing the similarity between learning resources and learning resources

Given online mobile learning resources I_a and I_b, document fragments x_1 and x_2 are obtained by XML transformation of their contents. The comprehensive similarity between the two documents and the two types of online mobile learning resources is the same, and the expressions are as follows:

$$sim(I_a, I_b) = sim(x_1, x_2) = \delta_1 \times semsim(x_1, x_2) + \delta_2 \times strusim(x_1, x_2) \tag{3}$$

In equation (3), the similarity between different learning resources includes two contents, namely, content similarity and structure similarity. δ_1 represents the weight of content similarity, δ_2 represents the weight of structure similarity, and there exists $\delta_1 + \delta_2 = 1$.

When solving the similarity between learning resources and learning resources, the influence degree of content and structure on the document information of learning resources is taken as the basis of weight matching (Liu et al., 2021; Wang and Guo, 2018). Among them, the similarity of the content and structure of the two learning resource documents is $semsim(x_1, x_2)$ and $strusim(x_1, x_2)$, and the calculation process of the two is shown as follows:

$$semsim(x_1, x_2) = \frac{\mu_1 \times x_1 + \mu_2 \times x_2}{\delta_1 \times \delta_2} \tag{4}$$

In equation (4), μ_1 and μ_2 are the attribute membership functions of two learning resource documents, respectively.

$$strusim(x_1, x_2) = \left| \frac{\left(x_1 - \overline{x_1}\right) \times \left(x_2 - \overline{x_2}\right)}{l\left(I_a, I_b\right)} \right| \tag{5}$$

In equation (5), $l\left(I_a, I_b\right)$ represents the absolute distance between two types of online mobile learning resources in the multi-dimensional space.

2.3 The similarity calculation between learners' search preference results and learning resources

In the recommendation process of online mobile learning resources, the historical download rate of learning resources is a very effective and important characteristic factor (Zhao et al., 2021). One of the most important factors affecting the download rate of learning resources is the direct match between learners' search results and their interest preferences (Wang et al., 2018). Assuming that the label of learners' interest preference is p and the label of online mobile learning resources is q, the cosine similarity between them is calculated, and the similarity between them within the same category is as follows:

$$sim(p, q) = \frac{p \times q}{g_p \times g_q} \tag{6}$$

In equation (equation (6), g_p represents the label score of learners, and g_q represents the label score of learning resources. On this basis, a preset threshold e is set to make $e < sim(p, q)$.

However, in real work, learners' preference for a learning resource is time-dependent, and the degree of interest changes dynamically at different times. Therefore, when calculating the similarity between learners' search preference results and learning resources, the characteristics of learners' interest offset must be considered (Song et al., 2022; Li et al., 2021).

In the online mobile learning network space, the weight of each node is the main parameter reflecting the degree of learners' interest in learning resources, and the interest migration model can continuously update the weight of nodes according to the search time and frequency of learners, so as to reflect the change of learners' interest. When the learner's search is closer to the current time, the search frequency of the same learning resources is higher, the connection weight value between nodes is higher, and the learner's interest preference for this node is stronger. Otherwise, smaller weights will be allocated between nodes. The weight between learner u and learning resource I can be described as follows:

$$\omega_{uI} = \sum_{1}^{P} \left(\frac{P}{1 + e^{(t - t_h) + t_0}} + 1 \right) \tag{7}$$

In equation (7), t represents the current time, P represents the frequency of searching for the same content, t_h represents feedback time for learning resources, and t_0 represents the time coefficient of learner's interest offset. Then, the similarity between learners' search preference results and learning resources is calculated. The process is as follows:

$$sim(u, I) = \omega_{uI} \times g_{u,I} \tag{8}$$

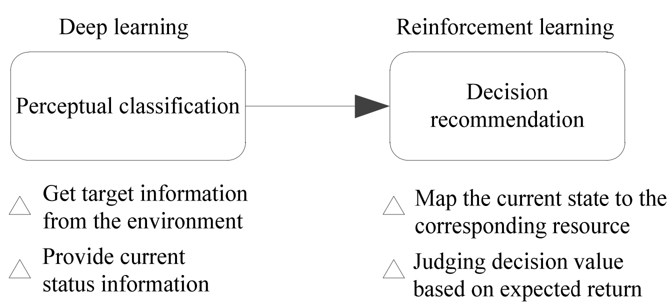
In equation (8), $g_{u,I}$ represents the scoring of learning resources by learners.

According to the calculation results of the above similarity, similar learning resources can be quickly provided for learners with high similarity, which can not only improve the recommendation efficiency of online mobile learning resources, but also improve the accuracy of resource recommendation. After the similarity is calculated, online mobile learning resources are recommended based on the deep reinforcement learning theory.

3 Recommend online mobile learning resources based on deep reinforcement learning process

Deep reinforcement learning is used to implement the recommendation processing of online mobile learning resources. Deep reinforcement learning integrates the perceptual classification results of deep learning and the decision results of reinforcement learning, which can obtain better analysis and processing results. Its structure is shown in Figure 1.

Figure 1 Basic structure of deep reinforcement learning



In the process of recommending online mobile learning resources based on deep reinforcement learning, online mobile learning resources (actions) are recommended to learners according to their interest preferences (states) for different learning resources, and their interest preferences are updated according to the feedback information (reward values) given by learners (Wang et al., 2021). Therefore, this study designs a multi-agent deep reinforcement learning network, which includes a recommendation agent and a classification agent.

(A) *Recommendation agent*

The process of learning resource recommendation is realised through deep reinforcement learning. Its main task is to let the recommendation agent learn an optimal recommendation scheme, which is a mapping process from 'state' to 'action', and its input information is the learner's state (learner's interest, preference and download record) observed by the recommendation agent. The output information is the action to be performed by the recommendation agent (the learning resource to be recommended). The learner will give feedback after interacting with the learning resources recommended by the recommendation agent. The recommendation agent calculates the reward value of the recommendation based on the feedback information. The higher the cumulative reward value, the better the recommended scheme. The optimal recommendation scheme shows that the recommendation agent can maximise the reward value after recommending learning resource a_t to learners according to their feedback status j_t. Among them, the calculation formula of cumulative reward value and optimal recommendation scheme is as follows:

$$Q = \sum_{t=1}^{T} \gamma^t \times r_t \times \frac{a_t}{j_t} \tag{9}$$

$$C = argmaxQ \times \frac{a_t}{j_t} \tag{10}$$

In equations (9) and (10), t represents the recommendation times, γ^t represents the discount factor, r_t represents the reward value of the *t*th recommendation scheme, Q represents the cumulative reward value, and C represents the optimal recommendation scheme.

(B) *Classification agent*

In the classification agent structure of deep reinforcement learning network, multiple feature classifiers are constructed to implement weighted summation and fusion processing of learner preference features through score information, which are described as follows:

Input: learner preference feature test sample set.

Output: learner preference feature fusion result.

The constructed classifier is composed of two modules, namely, inception-ResNet module and residual module. Among them, the inception-ResNet module can search for the optimal sparse structural units and rapidly expand the nodes between activation functions, so as to form a new neural network structure, which can quickly classify features of different scales. The residual module can slow down the degradation of deep network and optimise the performance of training process. At the same time, the residual module can superimpose the output results of the previous layer network with the historical data, and take the superimposed results as the input information of the next layer network, so as to reduce the training error.

Assuming that b is the input information of the deep reinforcement learning network, the output result of the residual module can be expressed through the following equation:

$$R = h(b) \otimes \mu\omega - \epsilon \tag{11}$$

In equation (equation (11), $h(\cdot)$ represents the fitting function, $\otimes$ represents the convolution calculation, ω represents the weight parameter within the deep reinforcement learning network, ϵ represents the loss amount, and μ represents the average classification threshold. If $h(\cdot)$ belongs to the identity mapping, the network output R_1 of the first-layer residual structure can be calculated by the following formula:

$$R_1 = h(b-1) \times \omega - \epsilon \tag{12}$$

On this basis, a sum algorithm is used to describe all the outputs of the deep reinforcement learning network. If the deep reinforcement learning network has a total of n residual modules, the following formula can be used to calculate the overall output of the deep reinforcement learning network:

$$R_n = h(b_n) + G\sum_{1}^{n} R \times \omega - \epsilon + b \tag{13}$$

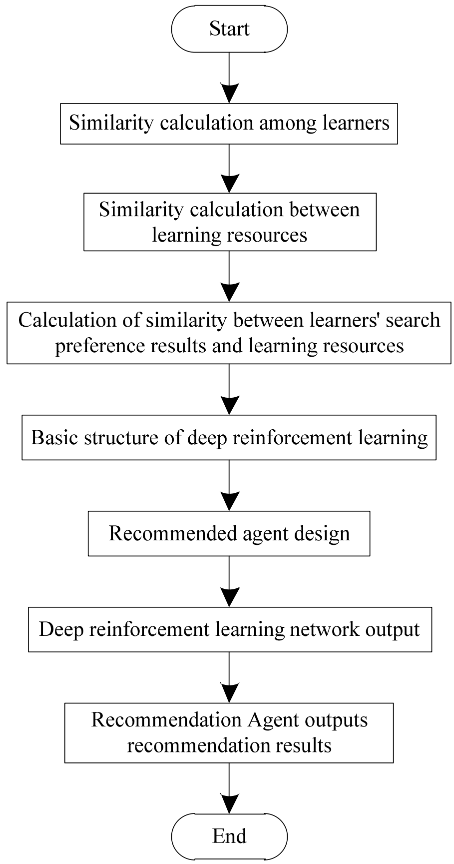
In equation (13), G represents the activation function, b represents the bias quantity, and its value is constant. Through the operation process of recommendation agent and classification agent, combined with the calculation results of three kinds of similarity in Part 2, and based on the classification processing of online mobile learning resources, the final recommendation scheme is output by the recommendation agent. The process is as follows:

$$R_n = CQ \times \frac{R_n \times sim(u, I)}{\sigma \times sim(u, v) + \lambda \times \eta \times sim(I_a, I_b)} \tag{14}$$

In equation (14), η represents the topic-resource feature distribution parameter, λ represents the result of resource feature distribution, and σ represents the affective polarity of learners.

Through the above calculation, complete the recommendation of online mobile learning resources. The degree of similarity between learners and learning resources, and learners search preferences results and learning resources, based on the calculation of similarity between the depth of the design of a multi-agent reinforcement learning network, according to the learner's interest in different learning resource preference (state), recommend to learners online mobile learning resources (action), The final recommendation scheme was obtained through the recommendation agent. The online learning resource recommendation process based on deep reinforcement learning is shown in Figure 2.

Figure 2 Process of recommending online learning resources



4 Experiment and result analysis

In order to verify the practical application performance of the online mobile learning resource recommendation method based on deep reinforcement learning designed above, the following experimental test process is designed.

4.1 Experimental environment design

The experiment took an online mobile network resource platform as the data object, randomly selected some learning resources from the million song open data source of the million song dataset site as the experimental object, and established the learning resource training dataset. The number of learners in the dataset is 1570, and the number of learning materials files is 300TB.

It is recommended that the test process be controlled by a PC with the following configurations: CPUIntel(R) Xeon(R) E5507; Python3.3; Eclipse integrated environment.

The parameters of the deep neural network are shown in Table 1.

Table 1 Parameter setting of deep neural network

Project	*Parameter*
Iteration step	400
Reward control factor	0.01
Memory cache unit size	10000
Online learning rate	0.003
Future bonus discounts	0.7
Target value network update step	50
Training data size	32

4.2 Compared with the design

In order to avoid the uniformity of experimental results, the method of Nie (2020) and the method of Dong (2021) are used as comparison methods to complete the performance verification together with the method in this paper.

In the experiment, the hit rate of recommendation results and the average ranking degree are respectively used as indexes to achieve the comparative evaluation of the recommendation effect of different methods. Among them:

The hit ratio of recommendation results is used to reflect the proportion of correct recommendations in the recommendation results. The larger the value, the larger the proportion of effective recommendations, and the recommendation effect is proportional to it.

The average ranking degree is used to reflect the ranking of recommendation results. For personalised recommended online mobile learning resources, learners are usually more likely to accept the top content in the recommendation list. Therefore, the higher the value of the average ranking degree, the higher the ranking of the recommendation results.

4.3 Results and analysis

Firstly, the recommendation result hit ratio of different methods is verified, and the result is shown in Figure 3.

By analysing the results shown in Figure 3, it can be seen that with the continuous increase of the number of learning resources, the hit rate of recommendation results of different methods also changes, but this change has no obvious regularity. The recommended result hit rate of the method in Nie (2020) reaches the maximum when the number of learning resources is 240tb, which is 87%. The recommended result hit rate of the method in Dong (2021) reaches the maximum when the number of learning resources is 180tb, which is 81%. The recommended result hit rate of this method reaches the maximum when the number of learning resources is 300tb, which is 95.5%, and the minimum recommended result hit rate can also reach 87%.

Then, the average ranking degree is used as the index to verify the performance of different methods, and the results are shown in Table 2.

Figure 3 Comparison of recommended results hit ratio of different methods (see online version for colours)

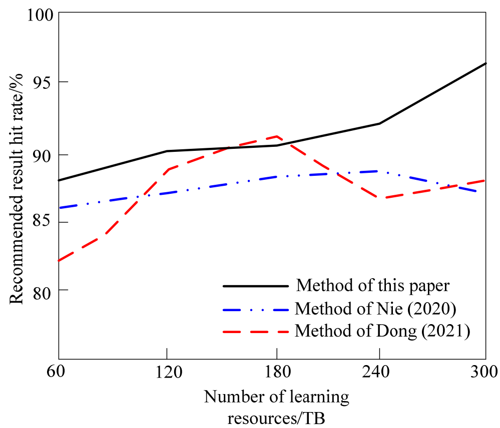


Table 2 Comparison of average ranking degree of recommendation results by different methods

Number of learning resources/TB	*Method of this paper*	*Method of Nie (2020)*	*Method of Dong (2021)*
60	0.875	0.774	0.734
120	0.879	0.801	0.763
180	0.880	0.843	0.778
240	0.892	0.861	0.792
300	0.926	0.887	0.830

By analysing the results shown in Table 1, it can be seen that with the increase of the number of learning resources, the average ranking degree of the three recommended methods gradually increases. Although the average ranking degree of the algorithm in this paper shows a weak downward trend, compared with the two traditional methods, although the average ranking degree of the recommended results of this method grows slowly, its value is large, and the highest average ranking degree can reach 0.926, which shows that the recommended results of this method are ranked higher and the recommended effect is better.

5 Conclusion

In order to improve the hit rate of learning resource recommendation results and make the recommendation results more satisfying to the needs of learners, this paper designs an online mobile learning resource recommendation method based on deep reinforcement learning, and verifies the performance of the method from both theoretical and experimental aspects. In the process of online mobile learning resource recommendation, this method has high recommendation result hit rate and average ranking degree. Specifically, compared with the recommendation method based on behaviour analysis, the recommendation result hit rate of Method of this paper is higher, with the highest reaching 95.5%. Compared with the recommendation method based on knowledge graph, the average ranking degree of the recommendation results of Method of this paper is significantly improved, and the highest average ranking degree can reach 0.926. Therefore, it is fully demonstrated that the proposed recommendation method based on deep reinforcement learning can better meet the requirements of online learning resource recommendation.

References

Chen, L. and Jing, Y. (2020) 'Design of mobile learning platform based on deep learning', *Modern Electronics Technique*, Vol. 43, No. 14, pp.177–179.

Dong, M. (2021) 'Multi-objective learning resource recommendation algorithm based on knowledge graph', *Computer Systems and Applications*, Vol. 30, No. 04, pp.139–145.

Fan, H., Shi, S., Zhang, B., Zhang, Y., Lin, Q. and Sun, H. (2020) 'Improved learning resource recommendation algorithm of deep neural network based on MLP', *Application Research of Computers*, Vol. 37, No. 09, pp.2629–2633.

Hu, W. and Zhou, X. (2021) 'Design and implementation of mobile collaborative learning APP based on connective learning theory', *Microcomputer Applications*, Vol. 37, No. 04, pp.42–44.

Li, X., Liang, H., Feng, J., Xiao, J. and Peng, W. (2021) 'Design of personalized learning resource recommendation system for online education platform', *Computer Technology and Development*, Vol. 31, No. 02, pp.143–149.

Liu, F., Tian, F., Li, X. and Lin, L. (2021) 'A collaborative filtering recommendation method for online learning resources incorporating the learner model', *CAAI Transactions on Intelligent Systems*, Vol. 16, No. 06, pp.1117–1125.

Nie, L. (2020) 'Personalized recommendation of learning resources based on behavior analysis', *Computer Technology and Development*, Vol. 30, No. 07, pp.34–37+41.

Song, F., Sui, D. and Zhou, X. (2022) 'Intelligence learning resource recommendation algorithm based on deep learning', *Journal of Nanjing University of Science and Technology*, Vol. 46, No. 2, pp.185–191.

Wang, G. and Guo, X. (2018) 'Personalized recommendation service based on user behavior analysis in social network environment', *Information Studies: Theory and Application*, Vol. 41, No. 08, pp.102–107.

Wang, G., Yuan, H., Huang, X. and Min, L. (2021) 'Recommendation algorithm for E-learning resources based on improved collaborative filtering', *Journal of Chinese Computer Systems*, Vol. 42, No. 05, pp.940–945.

Wang, X., Zhao, J. and Xu, Z. (2018) 'Research on user interest mining and personalized recommendation based on socialized annotation', *Modern Information*, Vol. 38, No. 07, pp.67–73+80.

Yang, J., Hu, G. and Wang, M. (2020) 'Research on configuration effect of mobile learning adoption motivation', *Library Tribune*, Vol. 40, No. 02, pp.64–73.

Zhang, M., Hua, Q., Jia, W., Chen, R. and Ji, X. (2019a) 'A personalized recommendation system based on deep neural network', *Journal of Southwest University(Natural Science)*, Vol. 41, No. 11, pp.104–109.

Zhang, P., Wang, Y. and Zhang, Z. (2019b) 'Research on personalized recommendation algorithm based on label and multi-information', *Computer Engineering and Applications*, Vol. 55, No. 05, pp.159–165.

Zhao, J., Sun, S., Guo, J., Zhong, Y., Wang, M. and Qin, Y. (2021) 'Recommendation algorithm design of network learning resource based on ontology', *Microelectronics and Computer*, Vol. 38, No. 1, pp.64–69.