Research on Personalized Learning Based on Collaborative Filtering Method

Xingguang Hu, Yingchao Wang, QiBo Chen, Qian Liu, Xiaoping Fan*
Hunan University of Finance and Economics, Changsha 410205, China

*2321294504@qq.com

Abstract: With the popularization of the Internet, the rapid development of intelligent education is promoted. They are a new platform for students to learn, and they can quickly help students improve and consolidate knowledge, as well as unlimited time, place, and space. However, due to a large number of learning resources, how to find the resources you need, and suitable from the massive resources is a problem that needs to be solved in this field. The individual characteristics of each learner are different, so when meeting their own needs, they have different requirements for individualization. The personalized learning resource recommendation system was born, according to the different characteristics of students to recommend corresponding learning resources. Since most of the recommendation systems currently recommend a large number of test questions or a large number of books, but each person’s needs are different, this article focuses on the above problems and selects a small piece of content course knowledge points in the learning resources, using course knowledge come as a recommendation point and solve the above problems by designing a personal personalized learning mechanism. The main research contents of this article are as follows: (1) It is proposed to take curriculum knowledge points as the recommended objects. (2) Use collaborative filtering methods and cognitive diagnosis methods to design individual learning mechanisms. (3) Experiment. Through related experiments, the use of collaborative filtering method and cognitive diagnosis method based on curriculum knowledge points recommendation to confirm whether it is feasible.

Keywords: collaborative filtering method; personalized learning recommendation; curriculum knowledge points

1. Introduction
With the rapid development of the Internet, it has promoted the rapid development of intelligent education. However, due to a large number of learning resources, how to find the resources you need, and suitable from the massive resources is a problem that needs to be solved in this field. The individual characteristics of each learner are different, so when they meet their own needs, they have different requirements for individualization. The personalized learning resource recommendation system was born and recommended corresponding learning resources for students according to their different characteristics. Since most recommendation systems currently recommend a large number of test questions or a large number of books, each person's needs are different.

The traditional education model has many limitations, for example, 1. The teaching model remains the same. 2. Restrictions on learning activities. 3. Uneven distribution of educational resources. In order to solve the above-mentioned problems, it is of great significance to make reasonable use of the students' own individual characteristics, to meet the preferences and needs of the learners, and to further improve the academic performance.

2. Related works

Scholars at home and abroad have studied personalized learning models and personalized resource recommendation methods, and have obtained gratifying research results, and found that they have invented more accurate personalized learning models and personalized recommendation algorithms. The application of personalized in different cultural backgrounds was demonstrated by Brom et al. through comparative experiments [1]. Learners get feedback from the three directions of self-personality, self-reflection, and intelligence, making them aware of the importance of the initiative to learning effects. This view is derived from the experience of Kong et al. in the personalized learning center [2]. According to the learner's movement pattern, behavior analysis model, and trajectory prediction, providing learners with personalized recommendation services is derived from the preference service mining method and personalized point of interest recommendation model proposed by Zhu et al. [3]. The personalized adaptive learning model, which is based on learner behavior patterns and learner navigation access data, and then provides personalized services through personalized features, is constructed by Sweta et al. through literature methods [4].

3. Personalized learning mechanism
3.1. Cognitive diagnosis

First of all, we need to know the mastery of students’ knowledge points. The DINA model is established by the data of the R matrix and the Q matrix. The \( \lambda_u \) is estimated by the maximum posterior probability of the EM algorithm, and its value after the probability is 0~1. Continuous value between 1. The vector \( \hat{\lambda}_u \) is redefined and estimated, and the student u's mastery of knowledge point k is calculated by the formula (3-1):

\[
\hat{\lambda}_{uk} = P(\lambda_{uk} = 1 | R_u) = \frac{\sum_{x=1}^{\lambda_u} P(\lambda_x | R_u)}{\sum_{x=1}^{\lambda_u} P(\lambda_x | R_u)} = \frac{\sum_{x=1}^{\lambda_u} \prod_{v=1}^{V} L(R_{uv}, \alpha_x, \hat{\lambda}_v, \hat{g}_v) P(\lambda_x)}{\sum_{x=1}^{\lambda_u} \prod_{v=1}^{V} L(R_{uv}, \alpha_x, \hat{\lambda}_v, \hat{g}_v) P(\lambda_x)}
\]

(3-1)

After obtaining the students’ true mastery of the test questions, due to the difference in the size of the data, the arithmetic average is taken to obtain the students’ average knowledge mastery of the knowledge required by the test \( S_{avguv} \):

\[
S_{avguv} = \frac{1}{\sum_{k=1}^{X} \prod_{x=1}^{K} l_{n_{xk}}} \sum_{k=1}^{X} \prod_{x=1}^{K} l_{n_{xk}}
\]
Under the average knowledge mastery of the knowledge required by the students to answer the questions, the actual answer $g_v$ and $s_v$ of the students' prior parameters are obtained, and the answer of the students' true level is as follows:

$$
R_{uv} = \begin{cases} 
1 & \text{if } (1-s)S_{avguv} \\
0 & \text{if } s_S(S_{avguv} + g_s(1-S_{avguv})) \\
\frac{s_S(S_{avguv} + (1-g_s)(1-S_{avguv}))}{s_S(S_{avguv} + (1-g_s)(1-S_{avguv}))} & \text{else}
\end{cases}
$$

(3-3)

3.2. Student score prediction

After the above two steps, the student’s cognitive level is obtained, and then the content-based collaborative filtering algorithm is used to obtain the corresponding student score prediction. The prior data $b_{uv}$ can be extracted from the student’s true level matrix $A$:

$$
b_{uv} = b_u + b_v, \quad b_u = \frac{1}{V} \times \sum_{i=1}^{V} A_{ui}, \quad b_v = \frac{1}{U} \times \sum_{i=1}^{U} A_{iv}
$$

(3-4)

The student's potential answer to the test question $\eta_{uv}$ is:

$$
\eta_{uv} = \bar{u} + b_{uv} + \frac{\sum_{u \in N} \left( R_{uk} - \bar{R}_u \right) \text{sim}(u,i)}{\sum_{u \in N} \text{sim}(u,i)}
$$

(3-5)

Among them, $\bar{u}$ is the overall average score.

4. Experiment and Result analysis

The Date data set comes from a certain physics exam in a middle school, and no other corresponding data processing is done, which maintains the completeness and authenticity of the data. The test questions are from school exams and practice papers. The main knowledge points are as follows: 1. The law of universal gravitation, 2. The action of force is mutual, 3. Uniform linear motion, 4. Centripetal force, 5. Friction force, 6. Support force, 7. Force decomposition, 8. Composition of forces,
9. Law of conservation of mechanical energy, 10. Description of motion, 11. Conservation of kinetic energy.

| Data set | Student | Knowledge points | Test questions |
|----------|---------|------------------|----------------|
| Sparseness | 120     | 11               | 109            |

**Tab.1** Specific information about the data set

According to the chi-square test, PDate<0.001, rejecting the null hypothesis H0, we can get the prediction that $\eta_{uv}$ and $R$ are not independent of each other, that is, the predicted score of the student is related to the actual student score, so it is proved that the collaborative filtering method can meet the required requirements. It is observed that the curve of correct answer rate with $\alpha$ decreases as the difficulty of recommendation decreases, and the correct answer rate of test questions corresponding to the recommended knowledge points of related courses continues to increase, indicating that the individual The personalized learning mechanism can recommend knowledge points that suit the students' own needs and difficulty by setting the value of the parameter $\alpha$. Assuming that $\alpha$ is 0.6, the corresponding difficulty course knowledge points corresponding to the corresponding difficulty test questions recommended on the personal personalized learning mechanism, the actual correct answer rate of students is close to 0.6, and the recommendation effect of personal personalized learning mechanism curriculum knowledge points is steadily increasing.

| Simple questions | Difficult questions |
|------------------|---------------------|
| 20% | 40% | 60% | 20% | 40% | 60% |
| Accuracy 0.9000 | 0.9028 | 0.9531 | 0.9743 | 0.9791 | 0.9773 |
| Recall rate 0.3523 | 0.3421 | 0.3701 | 0.4329 | 0.4401 | 0.4367 |
| F1 0.5412 | 0.5312 | 0.5926 | 0.5976 | 0.6069 | 0.6532 |

**Tab.2** The accuracy, recall, and F1 of the recommendation effect under the same data set

The precision rate, recall rate, and F1 value of the difficult and simple questions are different under the same test question set ratio, which is higher than that of the simple questions, indicating that the personalized characteristics of the difficult questions are smaller than that of the simple questions.

5. Conclusion

This mechanism is based on collaborative filtering methods and cognitive diagnosis. Method to recommend, grasp the individual's knowledge mastery through cognitive diagnosis, and then use collaborative filtering method for score prediction and course knowledge point recommendation, and flexibly set the two parameters $\alpha$ and $\beta$, one is the course knowledge point The difficulty range is used to adjust the size of the forecast affected by personalization, which basically takes into account
the personal characteristics of the individual. In the end, it achieves the purpose of personal and individualized learning. It is different from traditional methods and uses a small point of course knowledge of learning resources as recommended content, which simplifies the time required for learning and improves the enthusiasm of students in learning. Based on the data of a middle school physics test, relevant experiments are carried out to prove whether the personal adaptive learning mechanism described in this article is reasonable, specifically through the chi-square test, the change of parameter $\alpha$, and the recommendation of difficult and easy questions. The overall result is that Said basically meets the requirements.

Acknowledgement
This work was supported in part by the National Natural Science Foundation of China, grant number 72073041. Open Foundation for the University Innovation Platform in the Hunan Province, grant number 18K103; 2011 Collaborative Innovation Center for Development and Utilization of Finance and Economics Big Data Property, Universities of Hunan Province. Open project, grant number 20181901CRP03, 20181901CRP04, 20181901CRP05. 2020 Hunan Provincial Higher Education Teaching Reform Research Project under Grant HNJG-2020-1130, HNJG-2020-1124. 2020 General Project of Hunan Social Science Fund under Grant 20B16.

References
[1] Ishii M, Imaoka Hand Sato A 2017 Fast k-nearest neighbor search for face identification using bounds of residual score (FG2017: American Washington) pp 194-199
[2] Koren Y, Bell R and Volinsky C 2009 Matrix factorization techniques for recommender systems IEEE Computer (vol 42 Iss 8) pp 30-37
[3] Do la Torre J 2009 DINA model and parameter estimation: Adidactic. Journal of Educational and Behavioral Statistics (vol 34 Iss 1 ) pp 115-130
[4] Di Bello L V, Roussos L A and Stout W 2007 A Review of (7)gnitivdy diagnostic a.*vscsmcnl and a summary of psychomclricmodels (vol 31) Rao C R. Sin ha ray S eds Handbook of Statistics (vol 26 Psychometrics Oxford) (UK: Elsevier B. V.) pp 979-1030