Research on effectiveness model of online learning for college students in big data era

Xiaojun Zhang*

School of Marxism,
Wuhan University,
Wuhan, Hubei, 430072, China
and
Wuhan Technology and Business University,
Wuhan, Hubei, 430065, China
Email: xiaojun@mls.sinanet.com
*Corresponding author

Xiaoji Yang

Wuhan Academy of Educational Science,
Wuhan, Hubei, 430000, China
Email: 48895632@qq.com

Abstract: In order to enhance the effectiveness of online learning of college students in the era of big data, this paper puts forward the research on the effectiveness model of online learning of college students. This paper establishes a fusion clustering model for the evaluation of online learning effect, and uses fuzzy fusion grouping method to analyse the panel data of online learning effect evaluation of college students combined with big data mining method, and analyses the characteristics of online learning behaviour of college students and mining association rules. Linear programming model is used to optimise the scheduling of online learning resources for college students. The experimental results show that the design method has high accuracy and reliability in predicting the online learning effect of college students in the era of big data, and improves the convergence and optimisation ability of learning process.

Keywords: learning analysis; big data era; online learning; effectiveness assessment.

Reference to this paper should be made as follows: Zhang, X. and Yang, X. (2023) 'Research on effectiveness model of online learning for college students in big data era', *Int. J. Continuing Engineering Education and Life-Long Learning*, Vol. 33, No. 1, pp.1–9.

Biographical notes: Xiaojun Zhang is a PhD candidate at the School of Marxism, Wuhan University. She received her MSc Econ in International Management from the University of Wales, Aberystwyth in 2004. Currently, she is an Associate Professor in the School of Economics and Business Foreign Languages of Wuhan Technology and Business University. Her research direction is international management, ideological and political education.

Xiaoji Yang received his PhD in Computer Software and Theoretical Engineering from the Wuhan University in 2012. Currently, he is a Research Scientist in the Centre for Modern Education Technology of Wuhan Academy of Educational Science. His research direction is information technology teaching and information technology curriculum integration.

1 Introduction

Research on higher education reform has attracted great attention. Using online learning methods can improve the teaching effect and efficiency. As a necessary means for teaching, online learning methods for college students have played an irreplaceable role. Conduct analysis of higher education curriculum reform, and establish online education and learning models for college students. Combining resource optimisation scheduling and big data mining methods, transforming teaching achievements and optimising scheduling of teaching resources can improve the effectiveness of teaching. Studying the online learning effectiveness model of college students (Zhou et al, 2015) is of educational significance in improving college students' online teaching and optimising control capabilities. Related research methods of online learning effectiveness models for college students have attracted attention.

The analysis of online learning effectiveness of college students is based on the fusion of big data information and feature analysis of online learning. The combination of phase information and the method of detection to conduct online student research and reliability analysis. The traditional effective analysis methods for college students online learning mainly include fuzzy information detection methods, feature analysis methods and empirical analysis methods (Xu et al., 2017; Miao et al., 2012). Combining association rules mining method to carry out online learning information mining and big data fusion of college students. The fuzzy information clustering method is used to grid the collected online learning resource information. Ruan and Lin (2020) apply association rule mining to online learning integration. Combined with fuzzy c-means clustering analysis, online learning resources fusion and feature clustering are carried out, which improves the effectiveness of the education process. However, this method has a high degree of fuzziness in the analysis of online learning effect and its feature recognition ability is not strong. Ni et al. (2016) proposed an online learning effect analysis method for college students in big data era based on rough set feature matching, and used rough set feature matching method to analyse online learning resource fusion and effectiveness of college students. This method is not suitable for the evaluation of online learning effect. Li et al. (2007) proposed a statistical analysis method based on statistical analysis method for online learning effect analysis. This method has poor ability to optimise the process of learning information fusion.

Although the above methods to a certain degree to achieve the analysis of online learning effect, there are some problems such as weak feature recognition ability, poor adaptability and poor optimisation ability. For this reason, this paper constructs a prediction model based on linear programming and resource scheduling. In this paper, a fusion clustering model of college students' online learning effect evaluation panel data is established, and the learning process is optimised by using template feature matching method. Thesis method uses linear programming model to study the resource scheduling

and resource optimisation of college students' online learning, and analyses the effectiveness. Finally, the effectiveness of the design method is verified.

2 Statistical analysis and feature extraction of big data for online learning of college students

2.1 Statistical analysis of big data for online learning of college students

The realisation of the analysis of the effectiveness model of online learning for college students, combined with the method of fuzzy information fusion, the practice effect analysis of the college student's online learning method is conducted. Combined with the method of systematic reconstruction, it conducts feature modelling and hierarchical analysis, and performs reliability and quantitative feature analysis of the practical effect of online learning methods for college students. Compressed sensing combined with multi-attribute method is used to study online learning methods of college students (Li et al., 2007). Using analytic hierarchy process to conduct multi-attribute research, the statistical analysis results are:

$$MSD_{a\to b} = \sqrt{\left(d_{a,i} - \overline{d}_a\right)^2} \tag{1}$$

In the formula, $MSD_{a\to b}$ is the multiple mapping model of online learning for college students, $d_{a,i}$ is the learning effect data, and $\overline{d}_a$ is the average learning effect data. The $\boldsymbol{x}^{(0)}$ of the practical effect distribution of college students' online learning methods is divided into N grades, $\boldsymbol{x}^{(1)}$, $\boldsymbol{x}^{(2)}$, …, $\boldsymbol{x}^{(N)}$, which can also be expressed as $\boldsymbol{x}^{(0)} = \bigcup_{i=1}^{N} \boldsymbol{x}^{(i)}$.

utilising the method of fuzzy information programming, the characteristic parameter matrix of data recovery is carried out. Using block fusion and linear programming methods and multi-attribute derivation method (Leng et al., 2016), the practical effect distribution fusion model of college students' online learning methods is:

$$p = \boldsymbol{x}^{(0)} MSD_{a\to b} \tag{2}$$

Using the method of finite element feature analysis to conduct reliability analysis and adaptive parameter simulation of online learning for college students. Through the method of multi-state fusion, the fuzzy set function is constructed as $R_2^T R_2 = V_2 \sum\nolimits_2 V_2^T$ and the effective statistical information component of the online learning method of college students meets $I(R^N; \varphi_g \mid Z^N) = 0$. The statistical data set obtained is:

$$H = R_2^T R_2 \cdot R^N \tag{3}$$

Using the sparse representation method, the state parameter set of the online learning method of college students is expressed as:

$$F = \{f_1, f_2, \cdots, f_n\} \tag{4}$$

In the quantitative evaluation of the practical effect of college students' online learning methods (Li et al., 2016), the fuzzy set scheduling set of the practical effect distribution of college students' online learning methods under the linear programming model is:

$$f(t) = \frac{1}{2\pi} H \cdot F \tag{5}$$

Under different coding methods, the decision feature quantity for constructing the online learning effect of college students is:

$$Q = f(t)E(i) + b \tag{6}$$

where $E(i)$ is the expected value of the template data distribution, b is the decision coefficient. The fuzzy scheduling set for the effectiveness analysis is:

$$H = \begin{cases} H_0 : Q + w(t) \\ H_1 : \sqrt{E(i) + w(t)} \end{cases} 0 \le t \le T \tag{7}$$

where $w(t)$ is the fuzzy scheduling period. The fusion clustering model of college students' online learning effectiveness evaluation panel data is established, and the fuzzy correlation fusion clustering method is used to analyse the college students' online learning effectiveness evaluation panel data. Using the method of constraint parameter analysis, online learning research is conducted under different construction modes (Li et al., 2017).

2.2 Feature extraction

Combined with big data mining methods, it conducts feature analysis and association rule mining of college students' online learning behaviours, and analyses the effects (Sun and Lin, 2012). The regression analysis model for evaluating the effectiveness is expressed as:

$$S_C = \theta \frac{Q}{(Nr)^2} + \sum_{l=-\infty}^{\infty} \sin(f(t) - H) \tag{8}$$

In the above formula, r is the distribution threshold of prior knowledge, θ is the ambiguity coefficient of repeated iterations for college students' online learning, and N is the distribution set of college students' online learning behaviour features. On the basis of formula (8), the online learning behaviour characteristics of college students are extracted, and the binary programming model of online learning is constructed:

$$\begin{pmatrix} S_{C1} \\ \vdots \\ S_{Ct} \end{pmatrix} = \frac{x_m(t)}{a_{mi}} \tag{9}$$

where a_{mi} is the number of regressions, and $x_m(t)$ is the set of regression distributions. Combined with the method of sparse regression analysis, the statistical analysis model of college students' online learning party correlation is:

$$W = S_{Ct} \cdot w_j^k \tag{10}$$

where w_j^k represents the statistical collection of college students' online learning methods. Construct a fuzzy state feature distribution set for the effect distribution of college students' online learning methods. Using the method of dual planning, described as:

$$O_p = \sigma\sqrt{\ln(S_c)}\cdot W \tag{11}$$

Based on this method, the online learning behaviour characteristics of college students are extracted, and the actual effect of online learning of college students is analysed according to the dual programming method. On this basis, the actual effect of online learning method for college students is quantitatively analysed, and the scheduling of online learning resources for college students is realised.

3 Analysis and optimisation of the effectiveness of online learning methods for college students

3.1 Resource scheduling for online learning

Through the template feature matching method, the learning process is optimised, and the linear programming model is used for resource scheduling and resource optimisation of online learning for college students (Yu et al., 2014). Extract the relevant feature information of the effectiveness of college students' online learning methods, and use the process optimisation control method to evaluate the effectiveness of college students' online learning methods. Data link $V = \{1, 2, \ldots, n\}$ for the effectiveness evaluation of college students' online learning methods, generating metadata $\Upsilon = (r_{ij})_{N*N}$. Using the association rule mining method, the statistical features obtained by the effectiveness evaluation of online learning methods for college students are:

$$k(r) = V \cdot \Upsilon \cdot \Delta r_{ij} \tag{12}$$

Δr_{ij} is the length of the exponential distribution. If $r_{ij} > 0$, the autocorrelation feature quantity of college students' online learning method effectiveness evaluation is expressed as:

$$r_{ij} = \sum_{j \neq i} r_{ij} \tag{13}$$

Combined with the optimal mode selection method, the evaluation of the effectiveness of online learning methods for college students and the analysis of the optimised education mode are realised (McNelles et al., 2015). Then the educational effect of online learning is predicted.

3.2 Predictive evaluation of educational effectiveness

Through the template feature matching method, the learning process is optimised, and the linear programming model is used for resource scheduling and resource optimisation of online learning for college students. Process optimisation control method is used to analyse the effectiveness of online learning methods for college students (Abualigah

et al., 2017). The fuzzy control function for the effectiveness evaluation of online learning methods for college students is:

$$M_v = \frac{O_p}{k(r)} + M_h \tag{14}$$

The effective distribution feature quantity of college students' online learning method is M_h. The template matching method is used to achieve linear regression analysis of the effectiveness of online learning methods for college students. Through the piecewise regression analysis method, the statistical function of the effectiveness analysis of online learning methods for college students is:

$$Y = M_v \cdot \boldsymbol{p}_j(t+1) \tag{15}$$

In the optimal clustering centre, $\boldsymbol{p}_j(t+1)$ is used to represent the $t+1^{\text{th}}$ optimised feature distribution set, and the reliability analysis models of online learning methods for college students are $\boldsymbol{a}_1$, $\boldsymbol{a}_2$. In the M-dimensional random distribution space, the optimal convergence components for online reliability learning of college students are:

$$B(t) = \frac{Y}{f\left(\boldsymbol{p}_j(t)\right)} \tag{16}$$

In the formula, $f(\boldsymbol{p}_j(t))$ is the process function of college students' online learning. In the process of evaluating the statistical feature quantity of college students' online learning, the learning process is optimised through the template feature matching method. The linear programming model is used for resource scheduling and resource optimisation of online learning for college students and the fuzzy analytical model is used to predict the effectiveness of online learning for students and optimise the scheduling of learning resources (Hao et al., 2020).

4 Simulation experiments and results

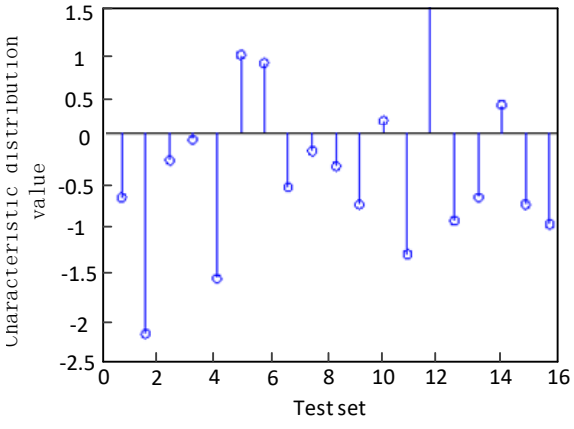
MATLAB is used for simulation test and analysis of the application effect of the design method. There are 1,000 samples of online teaching resources information of college students, which are tested four times respectively. The matching coefficient of fuzzy correlation features is 0.24, and the regression value of big data mining samples of online learning of college students is 1.46. The statistical analysis results of efficacy evaluation are shown in Table 1.

Table 1 Statistical analysis results of the effectiveness evaluation of online learning methods for college students

Data set	*Number of feature categories*	*Fuzzy distribution set/MBit*	*Correlation dimension*
Test set 1	46	57,000	3
Test set 2	45	34,000	4
Test set 3	24	43,000	5
Test set 4	56	21,000	15

According to the statistical results in Table 1, the linear programming model is used for resource scheduling and resource optimisation of online learning for college students. The characteristic distribution of online learning is shown in Figure 1.

Figure 1 Distribution of online learning effectiveness features (see online version for colours)



According to the distribution results of Figure 1, combined with the fuzzy analytical model, the online learning effectiveness prediction of college students and the optimal scheduling of learning resources are obtained. The scheduling results are shown in Figure 2.

Figure 2 Online learning resource scheduling results (see online version for colours)

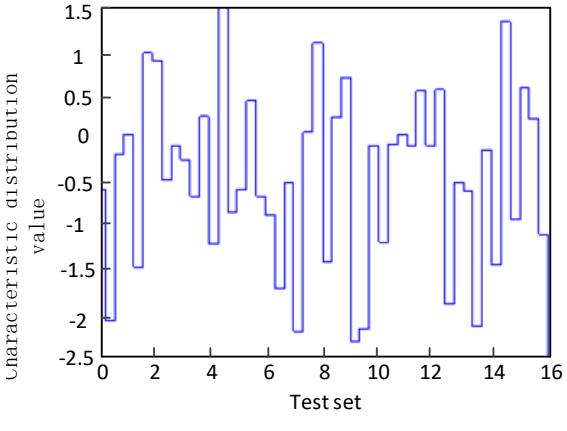
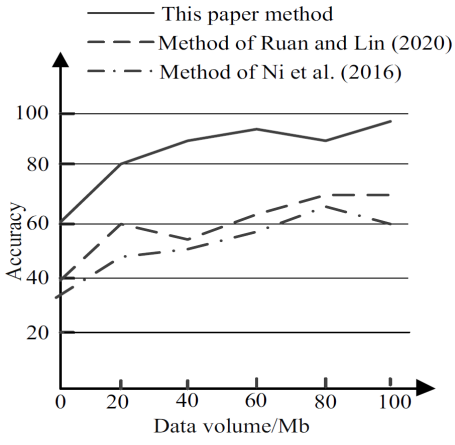


Figure 2 shows that the design method can be implemented effectively the effective management and optimised scheduling of college students' online learning methods. In order to test the accuracy of online learning validity prediction, the methods of Ruan and Lin (2020) and Ni et al. (2016) are used as experimental comparison methods, the result shown in Figure 3. It shows that the design method has a higher accuracy in predicting the effectiveness of online learning for college students, and has better convergence and ability to seek optimisation.

Figure 3 The prediction results of college students' online learning effectiveness



5 Conclusions

Establish the online education and learning mode of college students, the teaching results transformation and teaching resources optimisation scheduling are carried out by combining the resource optimisation scheduling and big data mining methods. The state parameter set of online learning method is obtained by sparse representation method, and the fusion clustering model of the panel data of online learning effect evaluation is established. The linear regression analysis of the effectiveness of online learning method is realised by using template matching method, and the effect of online learning is analysed. The learning process is optimised by the template feature matching method, and the online learning resource allocation of college students is scheduled and optimised by using linear programming model, realise the prediction of online learning effect and the optimal scheduling of learning resources. The experiment shows, design method can effectively manage and optimise the scheduling, also has high accuracy in predicting the effect of online learning, and has better convergence and optimisation ability.

References

Abualigah, L.M., Khader, A.T., Al-Betar, M.A. and Alomari, O.A. (2017) 'Text feature selection with a robust weight scheme and dynamic dimension reduction to text document clustering', *Expert Systems with Applications*, Vol. 84, No. C, pp.24–36.

Hao, F., Xu, G., Xue, Y., Gao, Z. and Zhang, H. (2020) 'An energy consumption optimization and evaluation for hybrid cache based on reinforcement learning', *Journal of Computer Research and Development*, Vol. 57, No. 6, pp.1125–1139.

Leng, X., Zhang, S.Q. and Lei, Z.Y. (2016) 'Application of improved artificial bee colony algorithm in neural network', *Computer Engineering and Applications*, Vol. 52, No. 11, pp.7–10, p.25.

Li K., Yang, S.L. and Li, G.Q. (2007) 'Improved algorithm for uniform parallel-machine scheduling with minimizing total completion times', *Journal of System Simulation*, Vol. 19, No. 11, pp.2618–2621.

Li, J., Pan, Q. and Duan, P. (2016) 'An improved artificial bee colony algorithm for solving hybrid flexible flowshop with dynamic operation skipping', *IEEE Transactions on Cybernetics*, Vol. 46, No. 6, pp.1311–1324.

Li, J., Wang, J., Pan, Q., Duan, P., Sang, H., Gao, K. and Xue, Y. (2017) 'A hybrid artificial bee colony for optimizing a reverse logistics network system', *Soft Computing*, Vol. 21, No. 20, pp.6001–6018.

McNelles, P., Lu, L. and Abi-Jaoude, M-J. (2015) 'Dynamic Flowgraph methodology assessment of an FPGA-based postaccident monitoring system for Westinghouse AP1000 nuclear power plants', *ASME J. of Nuclear Rad.*, Vol. 1, No. 3, pp.34501–34501.

Miao, D.Q., Xu, F., Yao, Y. and Wei, L. (2012) 'Set-theoretic formulation of granular computing', *Chinese Journal of Computers*, University of Regina S.S.A. Canada, Vol. 35, No. 2, pp.351–363.

Ni, Z.W., Pang, S.S., Wu, Z.J., Li, R.R. and Li, J.M. (2016) 'Parallel machine scheduling problem based on improved glowworm swarm optimization algorithm', *Computer Engineering and Design*, Vol. 37, No. 6, pp.1531–1536, p.1548.

Ruan, C. and Lin, J. (2020) 'Multi-task logistic survival prediction method for time-dependent time-to-event data', *Journal of Computer Applications*, Vol. 40, No. 5, pp.1284–1290.

Sun, X.Y. and Lin, Y. (2012) 'Improved artificial bee colony algorithm for assignment problem', *Microelectronics and Computer*, Vol. 29, No. 1, pp.23–26.

Xu, J., Miao, D., Zhang, Y. and Zhang, Z. (2017) 'A three-way decisions model with probabilistic rough sets for stream computing', *International Journal of Approximate Reasoning*, Vol. 88, No. 49, pp.1–22.

Yu, Y., Wang, Z.Y. and Xu, D.G. (2014) 'Speed and current sensors fault detection and isolation based on adaptive observers for induction motor drivers', *Journal of Power Electronics*, Vol. 5, No. 14, pp.967–979.

Zhou, P., Du, L., Wang, H., Shi, L. and Shen, Y. (2015) 'Learning a robust consensus matrix for clustering ensemble via Kullback-Leibler divergence minimization', *Proceedings of the 24th International Conference on Artificial Intelligence*, AAAI Press, Palo Alto, pp.4112–4118.