Features of Group Online Learning Behaviours Based on Data Mining

https://doi.org/10.3991/ijet.v17i04.29583

Zhu Tanbo[1](✉), Wang Lei[2], Wang Die[3]
[1] State Grid Shandong Electric Power Company, Jinan, China
[2] State Grid Shandong Electric Power Company, Tai'an Power Supply Company, Taian, China
[3] State Grid Shandong Electric Power Company, Electric Power Research Institute, Jinan, China
zhutanbo@sd.sgcc.com.cn

Abstract—With the development of information technology, how to scientifically and properly organizing and guiding learners to learn actively and efficiently has become a research subject for domestic and foreign scholars. However, existing research on online learning behaviours studied little about learning attitudes, learning preferences, student-student interaction, teacher-student interaction and so on. To this end, this paper studies the features of group online learning behaviours based on data mining. In this paper, a K-means-based group online learning behaviour feature selection model and an AdaBoost-based group online learning behaviour classification model were constructed, and the processing methods, execution processes and algorithm functions of the two models were described in detail. Finally, the effectiveness of the constructed models was verified through an experiment.

Keywords—data mining, online learning, analysis of learning behaviour features, learning group

1 Introduction

The development of information technology has been promoting continuous changes in modern teaching models and learning forms [1-6]. In this context, how to scientifically and properly organize and guide learners to learn actively and efficiently has become a research subject for domestic and foreign scholars [7-12]. To achieve this goal, it is particularly important to analyze and study the features and patterns of learners' learning behaviours [13-18]. The analysis of learning behaviour features includes two parts - group feature analysis and individual feature analysis. The latter is the basis and refinement of the former, and the two types of features analyses support each other and develop together [19-25]. Luckily, with the increase of online open course learning platforms, online open courses and online learners, the learning behaviour data generated in the learning process of learners are gradually accumulating, further supporting the exploration of the learning behaviour features and patterns of learners in the online learning environment.

In order to identify and analyze the behavioural patterns that affect students' academic performance in the undergraduate computer programming course, Premchaiswadi et al. [26] integrated various subjects such as process mining, e-learning, and educational data mining to discuss the opportunity to apply event modelling and process management technologies in e-learning systems. Umam et al. [27] showed the design of an ubiquitous learning model and demonstrated the learners' experiences in improved engagement and behaviour when IIMG was used for learner-lecturer interaction, with the aim of identifying ubiquitous learning scenarios, understanding learners' and lecturers' impressions about engagement and behaviours and their contributions to learning. Som et al. [28] used a simple deep learning-based machine learning model to automatically determine the overall collaboration quality of a group based on the notes about the personal role, level and behaviours of each individual student in the group, and also explored the use of an ordinal cross-entropy loss function and studied its effect with and without mixup. Sacharidis et al. [29] proposed the idea of how to extract a more suitable model to explain and predict group learning decisions by observing the decision results. Online learning via online education platforms is a way that effectively integrates education with information technology, which has a great impact on the learning methods and modes used by students. Wang and Zhang [30] first obtained typical observation indicators as the original data set, then extracted six that can objectively reflect the learning behaviour features through correlation analysis, and finally put forward appropriate suggestions from the perspective of teaching management according to the learning behaviour features of different groups.

The existing domestic and foreign research on online learning behaviours mostly focused on analyzing the differences between online learning behaviour features and traditional teaching and learning behaviour features or predicting learners' learning styles or individualized learning needs from a psychological perspective based on the analysis results of learning behaviours. However, little research has been done on the learning attitudes, learning preferences, student-student interaction and teacher-student interaction. In addition, traditional research methods such as questionnaires cannot objectively reflect the group cognitive behaviours of learners during online learning, and further research is much needed. To this end, this paper studied the features of group online learning behaviours based on data mining. First, a K-means-based group online learning behaviour feature selection model and an AdaBoost-based group online learning behaviour classification model were constructed. Then, the processing methods, execution processes and algorithm functions of the two models were described in detail. Finally, the effectiveness of the constructed models was verified through an experiment.

2 Selection of group learning behaviour features

Figure 1 shows the analysis process of group online learning behaviours. It can be seen that group online learning behaviours can be analyzed based on the content of the group online learning environment and the accumulated data of online learning behaviours. In this paper, two group online learning behaviour analysis models based on data mining were constructed, namely the K-means-based group online learning behaviour feature selection model and the AdaBoost-based group online learning behaviour classification model. The processing methods, execution processes and algorithm functions of the two models are described in detail below.

Figure 2 shows the workflow of the group online learning behaviour feature selection model. Its implementation consists of three steps: initializing the cluster centre and determining the number of clusters, determining whether the optimum is achieved and selecting features.

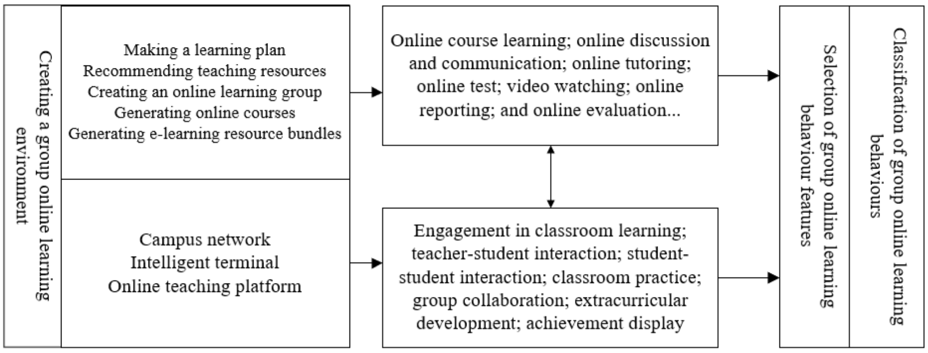

Fig. 1. Analytical process of group online learning behaviours

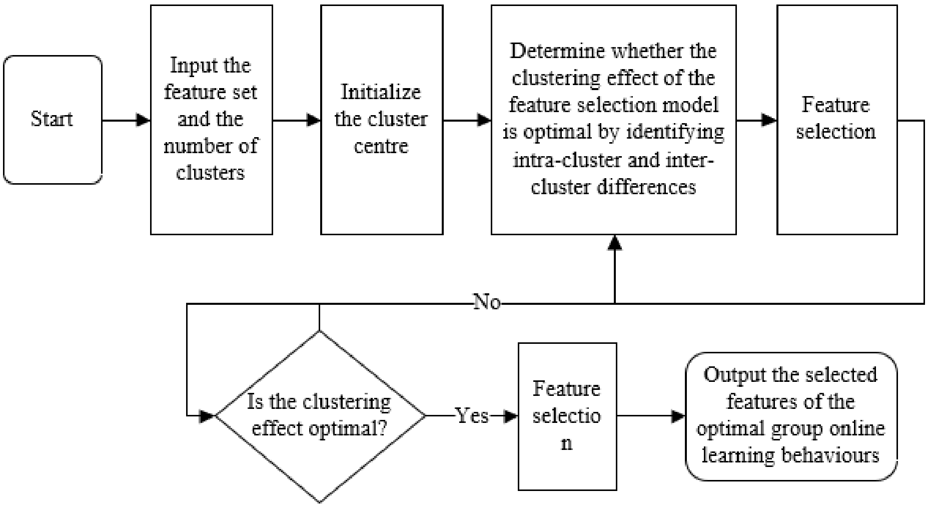

Fig. 2. Workflow of the group online learning behaviour feature selection model

Suppose there is a dataset of online learning behaviour features of m groups, represented by $A=\{a_1,a_2,...,a_m\}$. Each group has λ online learning behaviour sub-features, which are represented by $a_i=\{a_{i1},a_{i2},....a_{im}\}$. Divide A into l clusters

$RG_j(j=1,2,...,l,l<m)$, and then the $j(1\leq j\leq\lambda)$-th learning behaviour feature of the $i(1\leq i\leq m)$-th online learning group g_i can be defined as a_{ij}.

Calculate the Euclidean distance between online learning groups g_i and g_j $(1\leq i\neq j\leq m)$ according to Eq.(1):

$$\xi\left(g_i,g_j\right)=\sqrt{\sum_{x=1}^{\lambda}\left(g_{ix}-g_{jx}\right)^2} \tag{1}$$

Suppose that the distance density function corresponding to the learning behaviour feature data of the *i*-th online learning group in the group online learning behaviour feature set is represented by $DD(g_I)$, which can be defined as Eq.(2):

$$DD\left(g_i\right)=\sum_{v=1}^{m}\frac{\xi\left(a_u,a_v\right)}{\sum_{u=1}^{m}\xi\left(a_u,a_v\right)} \tag{2}$$

Suppose that the distance density function corresponding to the *i*-th online learning group is represented by S_i, and the adjustment coefficient by $\phi(1\leq zS\leq m)$, then S_i can be defined as shown in Eq.(3):

$$S_i=m^{\varphi}*\frac{1}{m}\sum_{i=1}^{m}e^{-DD\left(g_i\right)} \tag{3}$$

where, when the value of ϕ is 0.15, the clustering effect of online learning behaviour features is optimal. Suppose that the point density of the online learning group g_i in the group online learning behaviour feature set A is represented by $PO(g_i)$, which is an area with g_i as the centre and S_i as the radius. In the spherical area, the higher the point density of the online learning group, the greater the $PO(g_i)$, and there is:

$$PO\left(g_i\right)=\left|\left\{\lambda\mid\xi\left(g,\lambda\right)\leq S,\lambda\in A\right\}\right| \tag{4}$$

Suppose that the mean density $NP(a)$ of all feature data and the online learning groups in the group online learning behaviour feature set A can be calculated according to Eq.(5):

$$NP\left(g\right)=\frac{1}{m}\sum_{g\in A}PO\left(g\right) \tag{5}$$

In this paper, the discriminant function that can identify whether a new sample or piece of data falls within the same class was selected as the objective function of the group online learning behaviour feature selection model. It is used to identify intra-cluster and inter-cluster differences to make the clustering effect of the feature selection model optimal, and at the same time to obtain the clustering result and the number of clusters when its function value reaches the minimum.

Suppose that the group online learning behaviour feature set is represented by $A=\{a_1,a_2,...,a_m\}$, that the set of l classes by $U=\{u_1, u_2,...,u_l\}$, where $u_i(1\leq i\leq m)$ is the centre of the *i*-th category, then the degree of difference of intra-cluster feature data can be calculated according to Eq.(6):

$$q(u)=\sum_{i=1}^{l}q(u_i)=\sum_{i=I}^{l}\sum_{a\in u_i}\varepsilon(a,u_i)^2 \tag{6}$$

Assuming that the centres of the *i*-th cluster and the *j*-th cluster are represented by ε_i and ε_j, respectively, the difference $DC(u)$ between the two clusters can be calculated according to Eq.(7):

$$DC(u)=\sum_{1\leq j\leq i\leq l}\varepsilon(u_j,u_i)^2 \tag{7}$$

Assuming that the intra-cluster difference is denoted as $q(u)$ and that the inter-cluster difference as $DC(u)$, the discriminant function used is expressed as Eq.(8):

$$Q(u,l)=\frac{1}{1+e^{DC(u)-q(u)}} \tag{8}$$

Considering that there are many features of group online learning behaviours that need to undergo dimension reduction, in this paper, the weights of all group online learning behaviours were sorted through the feature selection operation, and then the forward algorithm was used to obtain the optimal feature subsets so as to avoid the impact of redundant data on the clustering results of group online learning behaviours. Further, the optimal group online learning behaviour feature set was obtained.

Let the group online learning behaviour feature set containing redundant data be represented by $E=\{e_1,e_2,...,e_m\}$, where each group online learning behaviour feature contains τ sub-features, represented by $e_i=\{e_{i1},e_{i2},...,e_{i\tau}|1\leq i\leq n\}$. The set of l classes is represented by $Z=\{z_1,z_2,...,z_l\}$, where $z_i\in Z$. First, select an online learning group h_i from the feature set E, and then select ε pieces of learning behaviour feature data from each cluster that are closer to h_i than other classes. The ε pieces of group online learning behaviour features that are of the same category with h_i constitute the set $F(z)$, and those of the different classes from h_i constitute the set $N(z)$. Based on $F(z)$ and $N(z)$, the weight vectors $\omega=\{\omega_1,\omega_2,...,\omega_\tau\}$ of group online learning behaviour features can be updated. Assuming that the data are sampled by m times, and that he function of difference between the online learning groups h_i and $h_j(1\leq i\neq j\leq n)$ in the r-th online learning behaviour feature is represented by $DV(r,e_i,a)$, then the weight of the $r(1\leq r\leq\tau)$-th group online learning behaviour feature can be calculated according to Eq.(9):

$$q_r^{i+1} = q_r^i - \sum_{a \in F(z)} \frac{DV(r, h_i, g)}{(m \cdot \varepsilon)} + \frac{\sum_{z \neq \Gamma(h_i)} \left[\frac{\eta(z)}{1 - \eta(\Gamma(h_i))} \right] \sum_{a \in N(z)} DV(r, h_i, g)}{(m \cdot \varepsilon)} \tag{9}$$

If the group online learning behaviour features are continuous, assuming that the minimum and maximum values of r in E are represented by min_i and max_i, respectively, there is:

$$DV(r, h_i, a) = \left| \frac{h_{ir} - h_{jr}}{max_r - min_r} \right| \tag{10}$$

If the group online learning behaviour features are discrete, there is:

$$DV(r, h_i, a) = \begin{cases} 0, h_{ir} = h_{jr} \\ 1, h_{ir} \neq h_{jr} \end{cases} \tag{11}$$

Given a hidden Markov model $\gamma=(C,D,\psi)$, assuming that the observation sequence of group online learning behaviours is represented by GC_1, GC_2, ..., GC_φ, and that the forward probability is defined as the probability of the state being w_φ, there is:

$$\langle\langle \beta_\phi(i) = \eta(GC_1, GC_2, ..., GC_\phi, E_\phi = w_i \mid \gamma) \rangle\rangle \tag{12}$$

The forward probability $\beta_\varphi(i)$ and the observation sequence probability $\eta(GC|\gamma)$ can be calculated through the recursion of the forward algorithm. Eq.(13) shows the initial value calculation formula of the forward probability $\beta_\varphi(i)$:

$$\beta_{\phi=1}(i) = \beta_1(i) = \psi_i y_i(GC_1), (i = 1, 2, ..., m) \tag{13}$$

In the above equation, the probability of the state used at the first time being w_i and the roll being GC_1 is represented by ψ_i. Suppose that the probability of the state w_i being used at the φ-th time, and the roll being the observation sequence GC_1,GC_2,...,GC is represented by $\beta_\varphi(\varphi)$, and that the product of the probability of the roll being the observation sequence GC_1,GC_2,...,GC_φ in the first φ times multiplied by the probability of the state being w_j at the (φ+1)-th time is represented by $\beta_\varphi(i)\beta_{\varphi j}$. Eq.(14) gives the recursive calculation formula of the forward probability $\beta_\varphi(i)$:

$$\beta_{\phi=1}(j) = \left[\sum_{i=1}^{M} \beta_1(i) \beta_{ij} \right] y_j(GC_{\phi+1}), (j = 1, 2, ..., M) \tag{14}$$

Since there are M kinds of group online learning behaviour features at the φ-th time, if the state w_j is used at the φ+1-th time, there are M possibilities for the φth time.

Assuming that the probability that there are M possibilities in the sequence $GC=(GC_1,GC_2,...,GC_\varphi)$ generated when w_i is used at the δ-th time is represented by $\beta_\delta(i)$, then the final calculation formula of the observation sequence probability is as follows:

$$\left\langle \eta\left(GC \setminus \gamma\right) = \sum_{i=1}^{m} \beta_\delta\left(i\right) \right\rangle \tag{15}$$

Through the above operations, the weights of the group online learning behaviour features can be sorted, and in this way, the optimal group online learning behaviour feature subsets can be constructed using the forward algorithm, and finally the optimal selection of group online learning behaviour features can be obtained.

3 Classification of group online learning behaviours

After completing the selection of group online learning behaviour features based on K-means, this paper used the AdaBoost strong classification method to test whether the feature selection result is appropriate. Figure 3 shows the execution process of the group online learning behaviour classification model. The idea of the adopted Adaboost algorithm is to combine the outputs of multiple "weak" classifiers to achieve effective classification.

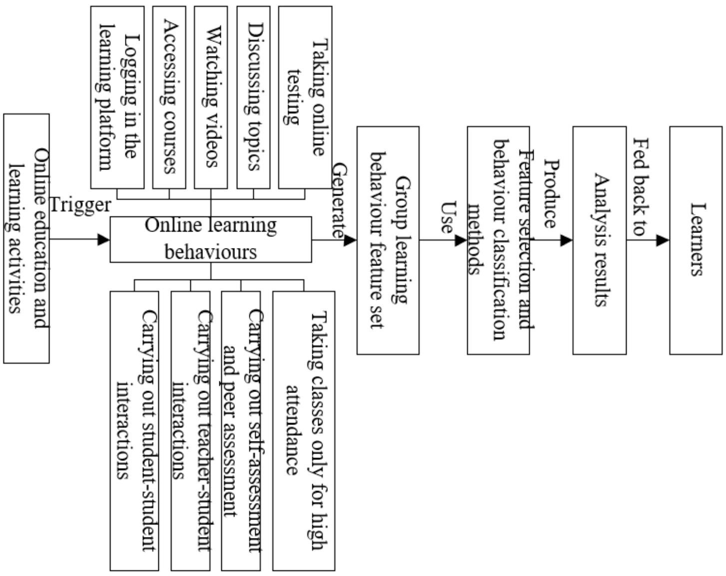

Fig. 3. Execution process of the group online learning behaviour classification model

For any given group online learning behaviour feature set Q={(a_1,b_1),(a_2,b_2), ...(a_M,b_M), the group online learning behaviour feature $a_j \in A \subseteq R^m$, and the expected classification result $b_i \in B$=(-1,1), and then perform N iterations,

Initialize the weight distribution of the group online learning behaviour feature set obtained in the previous section, and there is:

$$\Phi_1 = \left(\omega_{1,1}, \omega_{1,2} \cdots\cdots \omega_{1,i}\right) \omega_{1,i} = \frac{1}{M}, i = 1, 2, ..., M \tag{16}$$

Based on the above formula, other feature sets with weight distributions Φ_1, Φ_2, ..., Φ_n can be obtained, and through learning of the feature sets, the weak classifier $WC_n(a)$ can be obtained. Its classification error can be calculated according to Eq.(17):

$$\rho_n = \sum_{i=1}^{M} \omega_{n,j} I\left(WC_n\left(a_i\right) \neq b_i\right) \tag{17}$$

Based on the calculation result of the above classification error formula of $WC_n(a)$, the weight of the weak classifier in the strong classifier can be further calculated:

$$\beta_n = \frac{1}{2} log \frac{1-\rho_n}{\rho_n} \tag{18}$$

Assuming that the normalization factor is ζ_n, the weight distribution of the group online learning behaviour feature set is corrected according to Eq.(19):

$$\omega_{n+1,j} = \frac{\omega_{nj}}{\zeta_n} e^{-\beta_n b_i WC_n\left(a_i\right)}, i = 1, 2, ..., 10 \tag{19}$$

ζ_n can be calculated according to Eq.(20):

$$\zeta_n = \sum_{i=1}^{M} \omega_{nj} e^{-\beta_n b_i WC_m\left(a_i\right)} \tag{20}$$

The final classification result of group online learning behaviours is as follows:

$$G\left(A\right) = sign\left(\sum_{i=1}^{M} \beta_n WC_n\left(a\right)\right) \tag{21}$$

4 Experimental results and analysis

To test the proposed K-means-based group online learning behaviour feature selection model, this paper selected the relevant online learning behaviour data of 2455 undergraduates in 4 majors in a university from 2019 to 2020 as the test data and 7 indicators, namely platform access duration, number of platform accesses, achievement in video watching, achievement in online testing, number of engagements in topics discussion, number of student-student interactions and number of teacher-student interactions as the features of group online learning behaviours for analysis. Figure 4 shows the overall activity of group learning behaviours.

Through the proposed algorithm, the weights of all group online learning behaviour features are calculated and sorted. Since the weights of features will be affected by the samples randomly selected from the group online learning behaviour feature set, to eliminate the possible error caused by randomness, this paper chose to run the algorithm multiple times and obtain the mean values of the weights. Table 1 shows the weights of group online learning behaviour features. It can be seen that the weight values of the first few learning behaviour features are much larger than those of the last few ones, and thus, the last few learning behaviour features have less impact on the feature classification results.

Table 1. Weights of group online learning behaviour features

Running times	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Platform access duration	0.1628	0.1248	0.152	0.1362	0.1527	0.1659	0.1275
Number of platform accesses	0.2215	0.2149	0.2263	0.2518	0.2437	0.2519	0.2638
Achievement in video watching	0.0528	0.0492	0.0256	0.075	0.0415	0.052	0.0629
Achievement in online testing	0.0729	0.0529	0.0824	0.0925	0.0628	0.0573	0.0846
Number of engagements in topics discussion	0.2237	0.2149	0.1928	0.2235	0.1829	0.175	0.1837
Number of student-student interactions	0.075	0.0529	0.0746	0.1308	0.0826	0.0624	0.0915
Number of teacher-student interactions	0.2518	0.2149	0.2217	0.2649	0.2472	0.2631	0.225

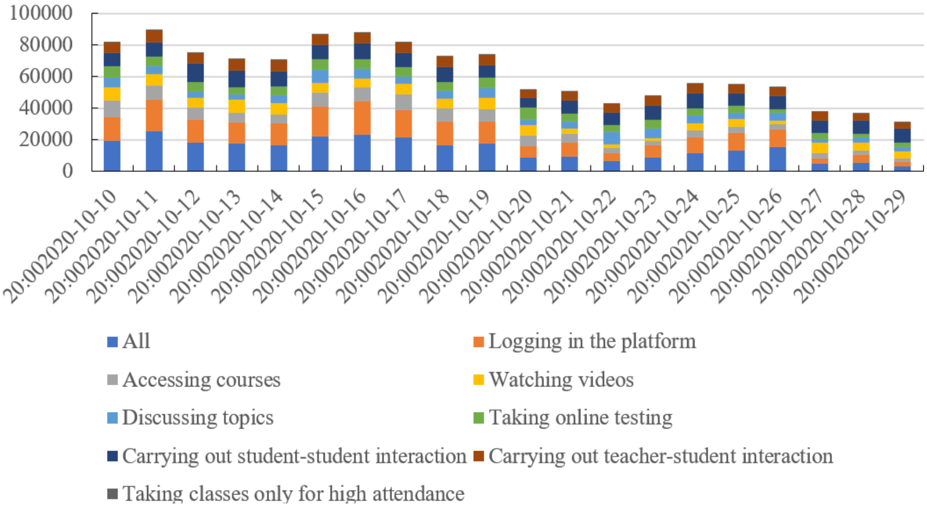

Fig. 4. Overall activity of group learning behaviours

Figure 5 presents the weight map of the major learning behaviour features. It can be seen that the weights of the number of platform accesses, the number of teacher-student interactions, and the number of engagements in topic discussions are greater than those of other features, which indicates that these features have major impact on the classification results of group online learning behaviours. On the other hand, the achievement in video watching and the achievement in online testing have less impact on the classification results.

Figure 6 shows the curve of behaviour classification accuracy with different numbers of features. After 6 features were added, the behaviour classification accuracy tended to be stable, which further verified the effectiveness of the data dimensionality reduction process after the feature weight sorting. In addition, the generated optimal feature subsets showed good classification performance, and the computational complexity was also effectively reduced.

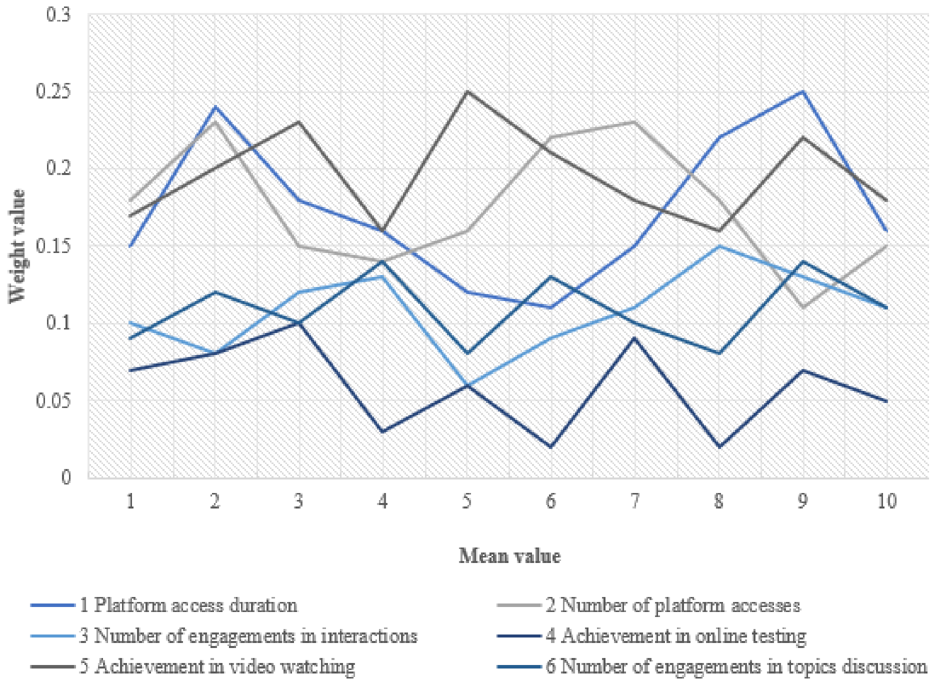

Fig. 5. Weight map of major learning behaviour features

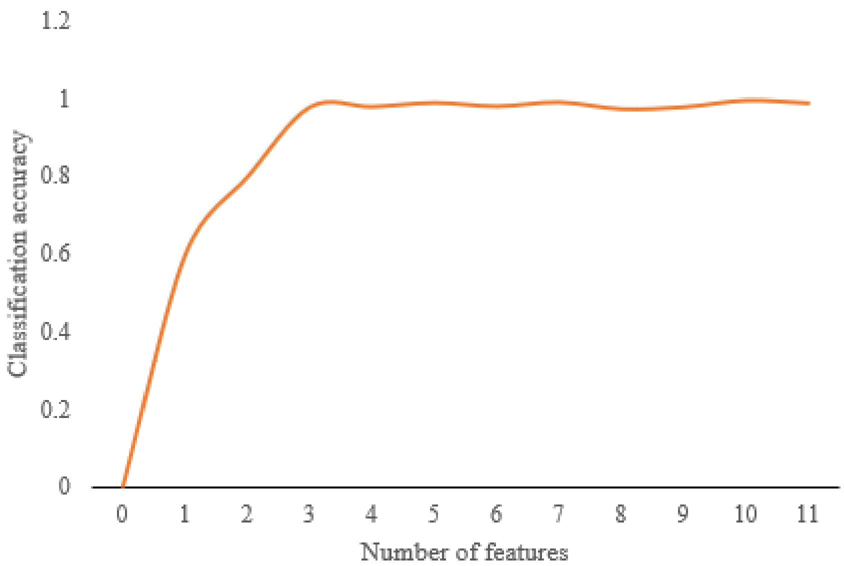

Fig. 6. Curve of behaviour classification accuracy with different numbers of features

In order to make the group learning behaviour feature selection results and feature classification results reasonable and portable, regression analysis was carried out in this paper, with the results shown in Table 2. According to the significance analysis at the level of multivariate regression coefficients, based on the previously generated optimal feature subsets, each feature indicator is very important. In collinear statistics, the closer the tolerance is to 0, the more collinear the feature indicators will be, and if the variance inflation factor is greater than 10, it means that there is limited collinearity between feature indicators. The tolerances of the above feature indicators are all less than 0.5, and the variance inflation factors are all less than 5, verifying that there is no significant multivariate collinearity among the selected feature indicators.

Table 2. Multivariate regression analysis of group learning behaviour feature selection results

Features		**Platform access duration**	**Number of platform accesses**	**Achievement in video watching**	**Achievement in online testing**	**Number of engagements in topics discussion**	**Number of student-student interactions**	**Number of teacher-student interactions**
Unnormalized estimate of coefficient	*Estimated value of B*	0.225	0.315	0.426	0.612	0.385	0.241	0.124
	Standard error	0.165	0.215	0.25	0.162	0.115	0.136	0.289
Normalized coefficient	β value	0.28	2.75	1.28	3.62	0.37	2.16	1.85
t value		5.074	8.042	7.538	4.182	5.294	3.628	4.275
Significance level		0.001	0.0005	0.0002	0.0001	0.0002	0.0004	0.0001
Collinearity Statistics	Tolerance	0.285	0.316	0.286	0.115	0.436	0.218	0.128
	Variance inflation factor	3.628	1.752	3.429	4.825	1.342	2.859	3.627

5 Conclusions

Little research on online learning behaviours has focused on learning attitudes, learning preferences, student-student interaction and teacher-student interactions. In view of this, this paper studied the features of group online learning behaviours based on data mining. In this paper, a group online learning behaviour feature selection model based on K-means and a group online learning behaviour classification model based on AdaBoost were constructed, and the processing methods, execution processes and algorithm functions of the two models were described in detail. The overall activity of group learning behaviours was also summarized based on an experiment. The weights of group online learning behaviour features were given, and the weight map of major learning behaviour features and the curve of behaviour classification accuracy with different numbers of features were drawn, which verified that the generated optimal feature subsets had good classification results. Finally, multivariate regression analysis was carried out on the selection results of group learning behaviour features, and it was verified that there is no significant multivariate collinearity among the selected feature indicators.

6 Acknowledgements

This paper was supported by the Science and Technology Project of State Grid Shandong Electric Power Company: Research on Human Resource Intelligent Decision Analysis and Early Warning Technology Application Based on "New Heights of Talent Development" (5206002000UR).

7 References

[1] Pang, L. (2020). Teaching Mode of Library Information Literacy Course Based on Hierarchical Embedded Service, International Journal of Emerging Technologies in Learning, 15(10): 27-41. https://doi.org/10.3991/ijet.v15i10.13995

[2] Zhou, N., Zhang, Z.F., Li, J. (2020). Analysis on course scores of learners of online teaching platforms based on data mining. Ingénierie des Systèmes d'Information, 25(5): 609-617. https://doi.org/10.18280/isi.250508

[3] Zhu, T. (2017). Creative Merging and Practical Application of" History of Modern Design" Course Online Teaching Platform Construction and Flipped Classroom Teaching Mode. International Journal of Emerging Technologies in Learning, 12(2): 118-136. https://doi.org/10.3991/ijet.v12i02.6318

[4] Zhao, C. (2012). Research on the mode design of writing teaching in college based on modern information technology. In 2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet), 2042-2045. https://doi.org/10.1109/CECNet.2012.6202256

[5] Cao, W., Wang, Q., Sbeih, A., Shibly, F.H.A. (2020). Artificial intelligence based efficient smart learning framework for education platform. Inteligencia Artificial, 23(66): 112-123. https://doi.org/10.4114/intartif.vol23iss66pp112-123

[6] Huang, W., Li, N., Qiu, Z.J., Jiang, N., Wu, B., Liu, B. (2020). An automatic recognition method for students' classroom behaviors based on image processing. Traitement du Signal, 37(3): 503-509. https://doi.org/10.18280/ts.370318

[7] Lacka, E., Wong, T.C., Haddoud, M.Y. (2021). Can digital technologies improve students' efficiency? Exploring the role of Virtual Learning Environment and Social Media use in Higher Education. Computers & Education, 163: 104099. https://doi.org/10.1016/j.compedu.2020.104099

[8] Makarova, I., Pashkevich, A., Buyvol, P., Mukhametdinov, E., Mavrin, V. (2020). Online Learning as a Necessary Measure During a Pandemic and as an Opportunity to Increase the Engineering Education Efficiency. In International Conference on Interactive Collaborative Learning, 608-620. https://doi.org/10.1007/978-3-030-68198-2_56

[9] Al Chibani, W. (2019). Investigating the efficiency of implementing active learning strategies in higher education courses in Lebanon: A multiple case study. IMSCI 2019 - 13th International Multi-Conference on Society, Cybernetics and Informatics, Proceedings, 2: 56-61.

[10] Mahdi, H.R., El-Hinawi, A.A. (2015). The Effectiveness of the Education College Programs at Al-Aqsa University in Develop Efficiencies Students in E-Learning. In 2015

Fifth International Conference on e-Learning (econf), 539-550. https://doi.org/10.1109/ECONF.2015.93
[11] Yang, J., Huang, G., Ma, J., Howard, S.K., Ciao, M., Gao, J. (2021). Fuzzy contrastive learning for online behavior analysis. In 2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 1-6. https://doi.org/10.1109/FUZZ45933.2021.9494588
[12] Li, Y., Hu, W., Liu, F., Liu, T. (2021). Analysis of Learning Behavior Indicators Based on Python. In 2021 4th International Conference on Intelligent Autonomous Systems (ICoIAS), 211-215. https://doi.org/10.1109/ICoIAS53694.2021.00045
[13] Wang, S., Xu, Y., Li, Q., Chen, Y. (2021). Learning Path Planning Algorithm Based on Learner Behavior Analysis. In 2021 4th International Conference on Big Data and Education, 26-33. https://doi.org/10.1145/3451400.3451405
[14] Wang, S. (2021). Online Learning Behavior Analysis Based on Image Emotion Recognition. Traitement du Signal, 38(3): 865-873. https://doi.org/10.18280/ts.380333
[15] Altuwairqi, K., Jarraya, S.K., Allinjawi, A., Hammami, M. (2021). Student behavior analysis to measure engagement levels in online learning environments. Signal, Image and Video Processing, 1-9. https://doi.org/10.1007/s11760-021-01869-7
[16] Zamansky, A., Sinitca, A., van der Linden, D., Kaplun, D. (2021). Automatic Animal Behavior Analysis: Opportunities for Combining Knowledge Representation with Machine Learning. Procedia Computer Science, 186: 661-668. https://doi.org/10.1016/j.procs.2021.04.187
[17] Li, M. (2021). Recognition of Psychological Characteristics of Students' Behavior Based on Improved Machine Learning. Journal of Sensors, 2021. https://doi.org/10.1155/2021/8135942
[18] Lorås, M., Aalberg, T. (2021). Characteristics of the Student-Driven Learning Environment in Computing Education: A Case Study on the Interaction between Educational Design and Study Behavior. In Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education V. 1, 11-17. https://doi.org/10.1145/3430665.3456310
[19] Liu, B., Guo, W., Lu, X., Xu, M., Cui, L. (2019). A Method of Analysis on Consumer Behavior Characteristics Based on Self-supervised Learning. In CCF Conference on Computer Supported Cooperative Work and Social Computing, 605-617. https://doi.org/10.1007/978-981-15-1377-0_47
[20] Yang, B., Yao, Z., Lu, H., Zhou, Y., Xu, J. (2020). In-classroom learning analytics based on student behavior, topic and teaching characteristic mining. Pattern Recognition Letters, 129: 224-231. https://doi.org/10.1016/j.patrec.2019.11.023
[21] Hajiabadi, Z., Shalchian, M. (2020). Behavioral Modeling and STDP Learning Characteristics of a Memristive Synapse. In 2020 28th Iranian Conference on Electrical Engineering (ICEE), 1-5. https://doi.org/10.1109/ICEE50131.2020.9260770
[22] Xie, T., Zheng, Q., Zhang, W. (2018). Mining temporal characteristics of behaviors from interval events in e-learning. Information Sciences, 447: 169-185. https://doi.org/10.1016/j.ins.2018.03.018
[23] Pinhati, F., Siqueira, S.W. (2014). Music students' behavior on using learning objects closer to the domain characteristics and the social reality. Computers in human behavior, 30: 760-770. https://doi.org/10.1016/j.chb.2013.10.039
[24] Dubey, H., Sangwan, A., Hansen, J.H. (2017). Using speech technology for quantifying behavioral characteristics in peer-led team learning sessions. Computer Speech & Language, 46: 343-366. https://doi.org/10.1016/j.csl.2017.04.002
[25] Premchaiswadi, W., Porouhan, P., Premchaiswadi, N. (2018). Process modeling, behavior analytics and group performance assessment of e-learning logs via fuzzy miner algorithm.

In 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), 2: 304-309. https://doi.org/10.1109/COMPSAC.2018.10247

[26] Umam, K., Mardi, S.N.S., Hariadi, M. (2017). Ubiquitous learning model using interactive internet messenger group (IIMG) to improve engagement and behavior for smart campus. In Journal of Physics: Conference Series, 801(1): 012083. https://doi.org/10.1088/1742-6596/801/1/012083

[27] Som, A., Kim, S., Lopez-Prado, B., Dhamija, S., Alozie, N., Tamrakar, A. (2020). A Machine Learning Approach to Assess Student Group Collaboration Using Individual Level Behavioral Cues. In European Conference on Computer Vision, 79-94. https://doi.org/10.1007/978-3-030-65414-6_8

[28] Sacharidis, D., Delic, A., Neidhardt, J. (2015). Learning the Role and Behavior of Users in Group Decision Making. In DMRS, 25-28.

[29] Wang, J., Zhang, Y. (2019). Clustering Study of Student Groups Based on Analysis of Online Learning Behavior. In Proceedings of the 2019 International Conference on Modern Educational Technology, 115-119. https://doi.org/10.1145/3341042.3341065

8 Authors

Tanbo Zhu, born in 1988, received a master's degree at Xi'an Jiaotong University, and now works in Human Resources Department, State Grid Shandong Electric Power Company. He received 11 awards for provincial/ministerial achievement, e.g., the first prize of China electric power technology innovation, and the first prize of Shandong electric power technology progress.

Lei Wang, born in 1988, received a master's degree at Xi'an Jiaotong University, and now works in Human Resources Department, Tai'an Power Supply Branch, State Grid Shandong Electric Power Company. He received one award for provincial/ministerial achievement, i.e., the third prize of Shandong electric power technology innovation (email: wanglei0209@163.com).

Die Wang, born in 1985, received a doctor's degree at Qingdao University, and now works in Scientific Research Department, Electric Power Research Institute, State Grid Shandong Electric Power Company. She received 4 awards for provincial/ministerial achievement, e.g., the second prize of Shandong electric power technology progress (email: wangdie1985@163.com).

Article submitted 2021-12-16. Resubmitted 2022-01-18. Final acceptance 2022-01-19. Final version published as submitted by the authors.