An anomaly detection of learning behaviour data based on discrete Markov chain

Dahui Li*, Peng Qu, Tao Jin,
Changchun Chen and Yunfei Bai

School of Computer and Control Engineering,
Qiqihar University,
Qiqihar, Heilongjiang, 161006, China
Email: dahuilii@tom.com
Email: 2546693@qq.com
Email: 54223965@qq.com
Email: 45886732@qq.com
Email: 58766149@qq.com
*Corresponding author

Abstract: In order to overcome the problems of large anomaly detection error and long detection time in traditional learning behaviour data anomaly detection methods, this paper proposes a learning behaviour data anomaly detection method based on discrete Markov chain. This method analyses the types of learning behaviour data, and determines the influencing factors of learning behaviour data. With the help of support vector machine, the data extraction range is determined, and the data redundancy is determined to complete the data pre-processing. This paper analyses the basic principle of discrete Markov chain, constructs the discrete Markov chain model, and completes the detection of abnormal learning behaviour data. The experimental results show that the maximum detection error of the proposed method is about 2%, and the detection time is always less than 2.5 s.

Keywords: discrete Markov chain; learning behaviour data; support vector machine; redundancy.

Reference to this paper should be made as follows: Li, D., Qu, P., Jin, T., Chen, C. and Bai, Y. (2023) 'An anomaly detection of learning behaviour data based on discrete Markov chain', *Int. J. Continuing Engineering Education and Life-Long Learning*, Vol. 33, No. 1, pp.69–83.

Biographical notes: Dahui Li received his PhD in College of Information and Communication Engineering from Harbin Engineering University in 2012. He is currently a Professor in the School of Computer and Control Engineering of Qiqihar University. His research interests include network security, network detection and analysis.

Peng Qu received her Bachelor's and Master's from Northeast Agricultural University in 2005 and 2008. She has been engaged in management work at Qiqihar University since 2008. Her research interests include student management and education.

Tao Jin is a teacher in the School of Computer and Control Engineering, Qiqihar University, Qiqihar, Heilongjiang, 161006, China.

Changchun Chen has graduated from College of Computer Science and Technology of Harbin Normal University in 1996. He is currently an Associate Professor in the School of Computer and Control Engineering of Qiqihar University. His research interests include computer control, and embedded system applications.

Yunfei Bai received his Bachelor's in College of Urban construction from Yangtze University in 2018. He is currently a graduate student in the School of Computer and Control Engineering of Qiqihar University. His research interests include knowledge graph and entity relation extraction.

1 Introduction

With the continuous popularisation of higher education in China, the number of college students is increasing year by year. The increasing number of students has brought some difficulties to the management of the school (Sun and Lu, 2018). At the same time, it is difficult to guarantee the quality of students' learning. Students' learning behaviour in school determines their learning quality. In recent years, with the continuous development of Internet technology, the traditional learning environment and lifestyle of students have been changed. The mixed information in the Internet has a certain impact on the immature students. University administrators need to find out students' abnormal behaviours in time and guide them correctly. Effective detection of students' abnormal learning behaviour can help maintain the stability of colleges and universities and improve the quality of students' learning (Li et al., 2018). Therefore, it is very important to detect students' abnormal learning behaviour data. For this reason, related researchers in this field have carried out a lot of research on the detection of abnormal learning behaviour data, and have achieved certain results.

In Liu et al. (2020), an improved multi-objective regression method for students' classroom behaviour data detection was proposed. Aiming at the diversity of students' behaviour in class, a new behaviour data detection method is designed. First of all, according to the data set to be detected, it is transformed into a uniform sampling data set. Using the improved neural network algorithm and clustering algorithm, we design a method to detect the students' changeable behaviour in class. According to the neural network algorithm, we train the sample data and design a data detection model to complete the detection of students' classroom behaviour data. This method can effectively improve the training speed of students' learning behaviour data, but the selection of key indicators in the construction of student behaviour data set is less, which easily leads to low accuracy of learning behaviour data detection, and has certain limitations. In Li et al. (2018), this paper proposes an evaluation model based on LMS. The model takes into account the important characteristic variables that affect students' learning behaviour, and designs the remote automatic evaluation of students' learning behaviour data. From students' online participation, self-monitoring and learning performance, the variables of students' learning behaviour are determined. The reliability and validity of learning behaviour variables were tested, and the online detection model of students' learning behaviour data was constructed by multi-level regression analysis. The design of the model can effectively detect the evaluation indicators of distance online learning, and the evaluation accuracy is high. However, the data of the model evaluation mainly comes from the collection of students' online data, and the scope of evaluation is relatively limited and has certain limitations. Zhao et al. (2019) proposed a method for identifying personality traits based on online learning behaviour data. This method first obtains the students' personality characteristics, and takes the important feature data as the research basis. In the recognition of learning behaviour data, assuming students' preference, the learning behaviour data is classified by machine learning classification algorithm, and then the basic operation principle of decision tree is analysed by rapid miner data mining tool, and five kinds of personality data of students are effectively identified according to naive Bayes and support vector machine algorithm. This method can effectively analyse the personality characteristics of students' online learning behaviour, but the measured data feature data is less, and the operation process is complex, and there are some problems such as the time-consuming of students' learning behaviour data detection.

Based on the problems of the above methods, this paper proposes a new anomaly detection method for learning behaviour data. With the help of discrete Markov chain, the learning behaviour data anomaly is effectively detected. The technical route of this paper is as follows:

1 Firstly, the types of learning behaviour data are analysed, and the influencing factors of learning behaviour data are determined;

2 On the basis of the extracted learning behaviour data, the redundancy of learning behaviour data is determined, and the different attributes of learning behaviour data are obtained. The independent data in the learning behaviour data are processed for redundancy to complete the data pre-processing.

3 This paper analyses the basic principle of discrete Markov chain, constructs the discrete Markov chain model, and completes the detection of abnormal learning behaviour data.

2 Design of anomaly detection method for learning behaviour data

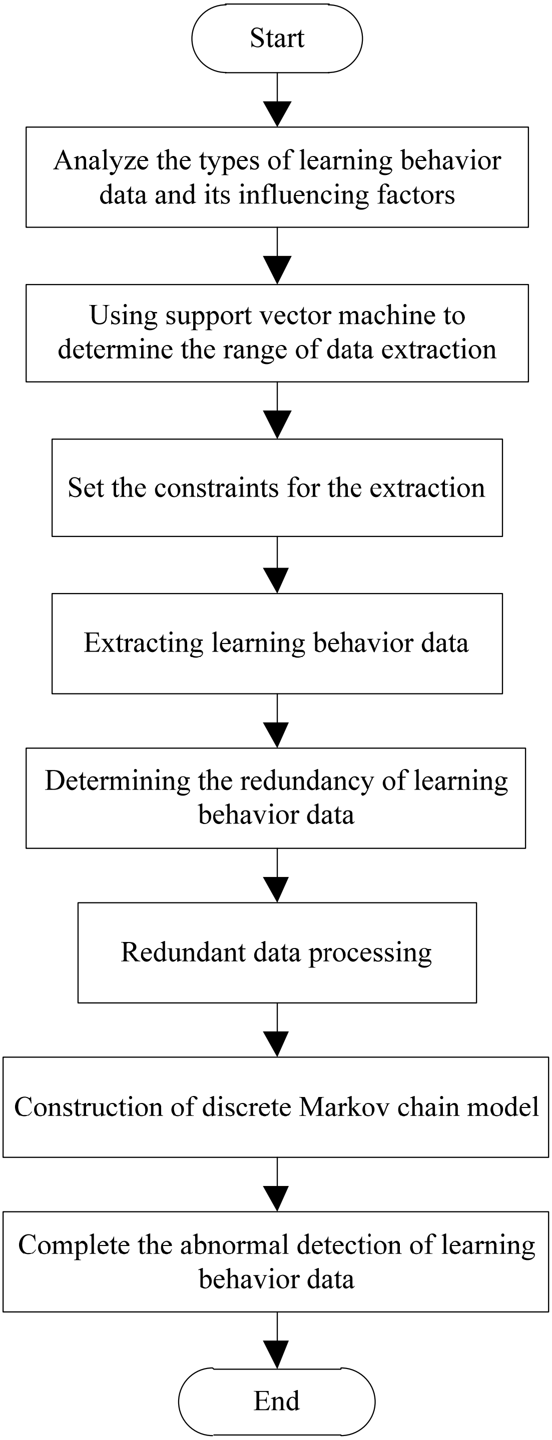
In order to solve the problems of traditional methods, this paper designs a learning behaviour data anomaly detection method based on discrete Markov chain. The overall design process of this method is shown in Figure 1.

2.1 Classification and influencing factors of learning behaviour

Students' learning behaviour is mainly manifested in daily learning and life, which is the general name of a series of learning actions (Wang et al., 2020). Students are a changeable individual, and their learning behaviour types are more diverse. According to different definition standards and learning States, they are divided into learning state/non learning state, and according to students' behaviour in the classroom, they are divided into positive behaviour and negative behaviour. According to the S-T (student teacher) behaviour analysis method, students' learning behaviour is divided, as shown in Figure 2.

In Figure 2, s represents students' learning behaviour and T represents teacher's classroom behaviour. Students and teachers should be considered in learning behaviour data.

Figure 1 Abnormal detection process of learning behaviour data


Start
Analyze the types of learning behavior data and its influencing factors
Using support vector machine to determine the range of data extraction
Set the constraints for the extraction
Extracting learning behavior data
Determining the redundancy of learning behavior data
Redundant data processing
Construction of discrete Markov chain model
Complete the abnormal detection of learning behavior data
End

Due to the influence of many factors on students' behaviour, we should pay attention to the key factors affecting students' learning behaviour (Tao et al., 2019). Among them, the key influencing factors of students' learning behaviour mainly come from the outside, and the influencing factors of learning behaviour are shown in Figure 3.

Figure 2 Learning behaviour

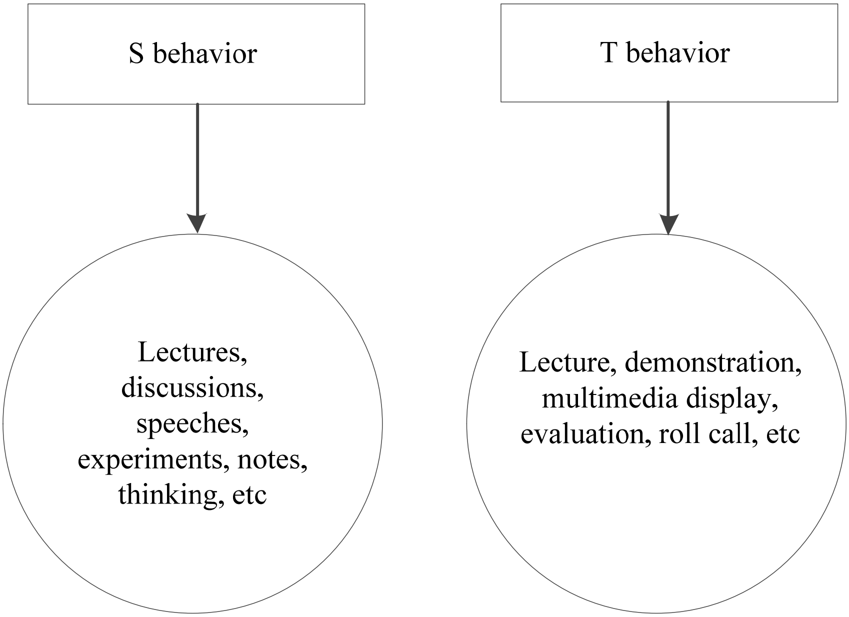
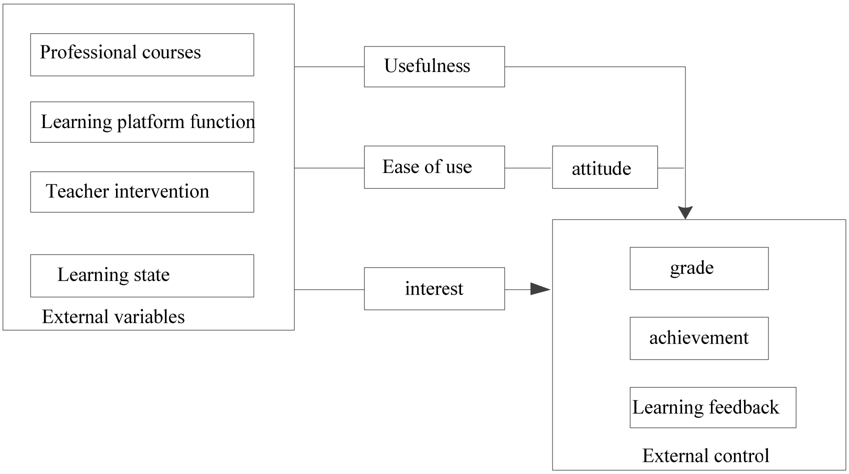

S behavior
T behavior
Lectures, discussions, speeches, experiments, notes, thinking, etc
Lecture, demonstration, multimedia display, evaluation, roll call, etc

Figure 3 Influencing factors of learning behaviour



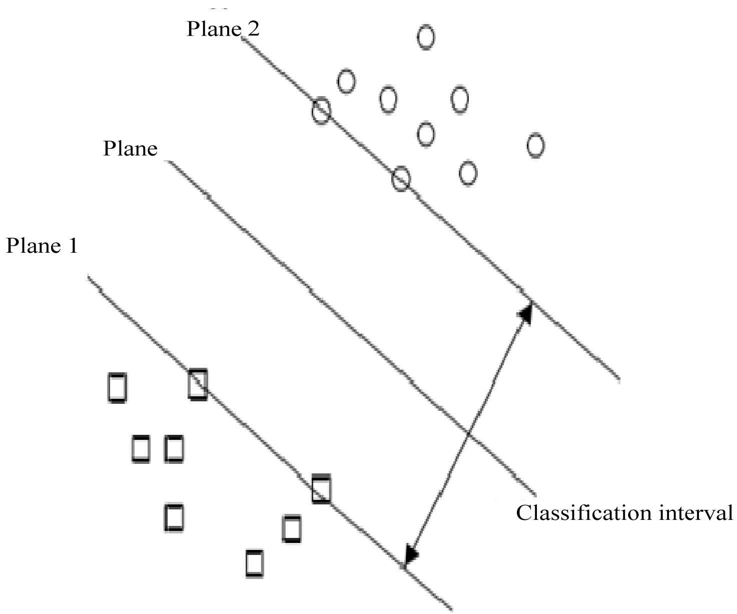
The external influencing factors of students' learning behaviour mainly include: the importance of professional courses, the function of learning platform, teachers' intervention and students' learning state (Hu et al., 2020). External factors affect the change of students' learning behaviour data.

2.2 Extraction of learning behaviour data

On the basis of the above classification of learning behaviour types and the determination of external influencing factors, the data of learning behaviour is extracted. In learning behaviour data extraction, due to the influence of external variables, a large amount of data needs to be extracted. Therefore, in order to facilitate the subsequent data detection, this paper uses support vector machine to extract learning behaviour data. Therefore, this learning behaviour data extraction method has many characteristics, such as fast speed and high precision.

Support vector machine (SVM) is an effective machine learning algorithm (Zhang et al., 2019), which can find the best data in the limited data and complete the effective extraction of data. It can find the most classified hyperplane in the set space, and extract two different kinds of data respectively. The principle of data extraction is shown in Figure 4.

Figure 4 Data extraction principle of SVM



In Figure 3, dots and rectangles represent two kinds of learning behaviour data. The plane division of data represents the edge distance of data acquisition, and the distance between planes is the distance of data extraction, which can ensure that the actual extracted data is within a certain confidence range (Qiu and Li, 2018).

When extracting learning behaviour data in the plane by support vector machine, certain constraints should be met to achieve the effectiveness of the extracted data:

$$\begin{cases} Ax_i + c \geq 1, \; y_i = 1 \\ Ax_i + c \leq -1, \; y_i = -1 \end{cases} \tag{1}$$

Where, x_i represents the extracted learning behaviour data sample, c represents the data extraction threshold, A represents the proportion of extraction, and y_i represents the boundary range of the extracted learning behaviour data.

The formula (1) is transformed into:

$$y_i(Ax_i + c) \geq 1 - \delta \tag{2}$$

In the formula, δ represent slack variables in learning behaviour data, and $\delta \geq 0$. Based on the above constraints, the minimum value of the learning behaviour data extraction function is obtained R (A) as follows:

$$R(A) = MIN\left(\frac{1}{2}\|A\|^2 + D\sum_{i=1}^{n} \delta\right) \tag{3}$$

In the formula, D is the data complexity of learning behaviour.

In order to solve the problem of convex optimisation, Lagrange multipliers are introduced to obtain the maximum range of learning behaviour data extraction. After satisfying the above constraints, the learning behaviour anomaly data are obtained according to the optimal classification function (Yu et al., 2019), that is:

$$F(x) = sign\left(\sum_{i=1}^{n} v_i y_i \left(x, x_i\right) + w\right) \tag{4}$$

In the formula, v_i represents the Lagrange multiplier, w represents nonzero coefficients. When the properties of the extracted learning behaviour data cannot be divided, it needs to be effectively divided by the optimal classification function, that is:

$$T(x) = sign\left(\sum_{i=1}^{n} v_i t_i \left(x, x_i\right) + A\right) \tag{5}$$

In the formula, K represent kernel functions, T (x) represents the decision function of learning behaviour data extraction.

In learning behaviour data extraction, support vector machine is used to determine the range of data extraction, set the constraints of extraction, and determine the maximum range of learning behaviour data extraction through Lagrange multipliers. The learning behaviour data is extracted by the optimal classification function.

2.3 Pre-processing of learning behaviour data

According to the data extracted from the above learning behaviour, it is pre-processed. Because of the conflict in the detection process of the extracted learning behaviour data, the redundancy of the data needs to be processed (Bi et al., 2018). First, determine whether there is redundancy in the extracted learning behaviour data, that is:

$$\varepsilon_i = \frac{\sum (Q - \bar{Q})(U - \bar{U})}{(n-1)\O_Q \O_U} \tag{6}$$

In the formula, n represents the number of learning behaviour data, $\bar{U}$ represents the average of learning behaviour data, $\emptyset_Q/\emptyset_U$ represent the standard deviation of different attributes of learning behaviour data. Among them,

$$\emptyset_Q = \sqrt{\frac{\sum\left(Q-\bar{Q}\right)^2}{n-1}} \tag{7}$$

$$\emptyset_U = \sqrt{\frac{\sum\left(U-\bar{U}\right)^2}{n-1}} \tag{8}$$

In the formula, there is a positive correlation between Q and U.

In learning behaviour data, the value of the *Q* increases, the value of the *U* increases, the larger the value, the more kinds of attributes the learning behaviour data contains. When it increases enough, the attribute redundancy of the learning behaviour data is deleted. If the learning behaviour data attributes are independent of each other, there is no redundancy problem (Du and Chen, 2018). The redundancy of learning behaviour data is processed by formula (9), that is:

$$\mu' = \frac{\mu-\bar{Q}}{\emptyset_Q} \tag{9}$$

In the formula, μ represents normalised learning behaviour data.

In the pre-processing of learning behaviour data, the redundancy of learning behaviour data is determined, the different attributes of learning behaviour data are obtained, and the non-independent data in learning behaviour data are processed with redundancy to complete the data pre-processing. Therefore, the above process has the characteristics of high speed and high precision.

3 Exception 0 Markov Chain

3.1 Discrete Markov chains

Markov process is a representative stochastic process, which is suitable for interval series and time series. This method mainly studies the state and transition law of things. Through the study of the state and operation law of the research object, it is predicted. This method is widely used in various fields of society. Markov chain is a relatively simple Markov process. In its prediction process, the numerical value is continuous. The two adjacent objects can be infinitely divided and the states of multiple objects can be decomposed. The state parameters of the research object in Markov chain are discrete values, and the obtained state is limited. In this paper, in the anomaly detection of learning behaviour data, the time series of learning behaviour data is regarded as discrete state. This method has no aftereffect of Markov process in the study of research objects (Chen et al., 2018)

The discrete time set in discrete Markov process is $\{\varphi\ (x),\ t\in T\}\ T$, among T = {0, 1, 2…}, the state space of the object is P, at this point, the finite dimension distribution function of the object is:

$$F(\varphi(x)) = \varphi \sum i_{n+1} \tag{10}$$

In the formula, the finite dimension distribution function *F*, φ represents the decentralised dimension value of the study object.

After obtaining the finite dimensional distribution function, the no aftereffect of Markov chain can be effectively analysed. In this case, the prediction results are not affected by the past state changes, only the existing state is analysed.

After analysing its inefficiency, this method has the advantage of predicting the transfer probability of the research object. Assuming that the random process in this method is a Markov chain, whose state space is $Z = \{z_1, z_2, \ldots, z_n\}$ and at any time point, the state of the research object is different, and there is only one running state. At this time, the state of the research object can be expressed as:

$$z_i \rightarrow z_1, z_i \rightarrow z_2, \ldots, z_j \rightarrow z_n \tag{11}$$

On this basis, each of its states is transferred, namely:

$$\Lambda i, j \in P\{z_{n+1} = j | z_n = i\} = z_{ij} \tag{12}$$

After completing the state transfer of the research object, it is necessary to obtain the probability of the existence of the current state of the research object. At this time, all its states are transferred to the probability set to realise the state analysis of the research object, that is:

$$\rho = \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{1n} \\ \rho_{21} & \rho_{22} & \rho_{2n} \end{pmatrix} \tag{13}$$

In formula (13), ρ represents the research state matrix of the object of study.

Through the analysis of the relationship between Markov process and Markov chain, and determine the advantages of this method, then according to its working principle, the learning behaviour data is taken as the object of this analysis, and the discrete Markov chain is used to detect the data anomaly.

3.2 Abnormal detection of learning behaviour data

Because the value of the research object in Markov chain is discrete state, this paper constructs a discrete Markov chain model to detect the abnormal learning behaviour data.

The time series of learning behaviour data is assumed to be *T*, where the detected data is the N_t, contains normal and abnormal behaviour data in detected learning behaviour data. At this point, the discrete value of learning behaviour data is s, when abnormal, that is $S = \{S_1, S_2, \ldots, S_3\}$. Assuming that the probability of learning behaviours data anomalies in this state is G:

$$G = P(S_t = t) \tag{14}$$

Through the passivity of Markov chain, we can get:

$$G_i = G(z_i + 1 = j | z_i = i) \tag{15}$$

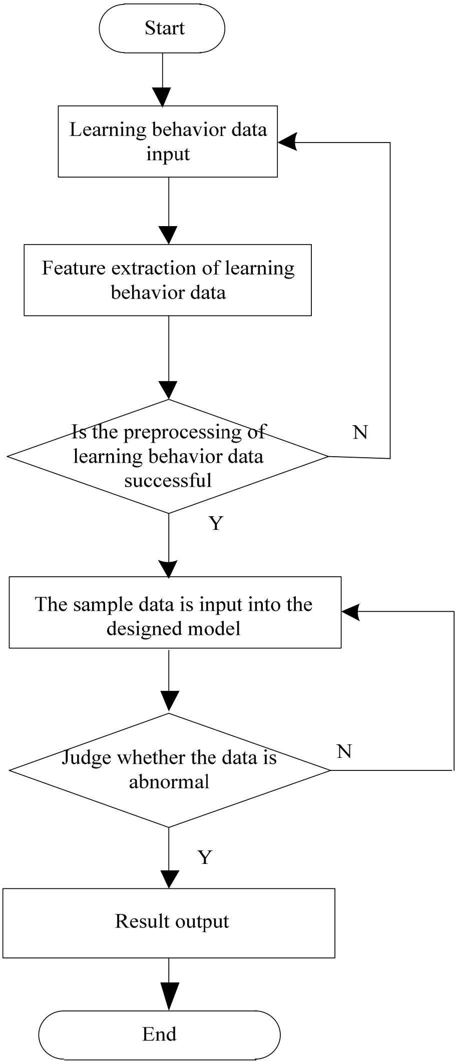
At this point, the state probability of learning behaviour data anomalies is expressed as:

$$c_t = [c_1, c_2, ..., c_n] \sum_{i=1}^{n} c_t = 1 \tag{16}$$

$$\rho = \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{1n} \\ \rho_{21} & \rho_{22} & \rho_{2n} \end{pmatrix}, \rho_{ij} \geq 0, i, j \in N \tag{17}$$

The data to be detected in learning behaviour data is input into formula (16) and formula (17) to obtain the probability of abnormal learning behaviour data. In order to ensure the effectiveness of anomaly detection of learning behaviour data, it is necessary to verify the probability of its occurrence. In this paper, Mahalanobis test method is used to verify it, so as to facilitate the subsequent construction of discrete Markov chain model to detect the accuracy of data anomaly (Guo et al., 2018).

Figure 5 Learning behaviour data anomaly detection process based on discrete Markov chain model


Start
Learning behavior data input
Feature extraction of learning behavior data
Is the preprocessing of learning behavior data successful
N
Y
The sample data is input into the designed model
Judge whether the data is abnormal
N
Y
Result output
End

In the anomaly detection of learning behaviour data, assuming that there is no abnormal state in the learning behaviour data, the marginal probability obtained from the transfer frequency matrix is the discrete Markov model.

$$H_K = \frac{\sum_{i=1}^{n} f_n}{\sum_{i=1}^{n} \sum_{j=1}^{n} f_n} \tag{18}$$

In formula (18), f_n exists when the value is large:

$$R(A) = MIN\left(\frac{1}{2}\|A\|^2 + D\sum_{i=1}^{n} \delta\right) \tag{19}$$

where τ represents the marginal probability and p_j represents the probability of occurrence of the normal state of the data.

After the marginal probability is obtained, the distribution state of the degree of freedom of learning behaviour data can determine that there is abnormal state in the data. The process of abnormal detection of learning behaviour data is shown in Figure 5.

To sum up, this method uses support vector machine to determine the range of data extraction, extract the learning behaviour data, and then determine the redundancy of learning behaviour data to complete data pre-processing. This paper analyses the basic principle of discrete Markov chain, constructs the model of discrete Markov chain, and completes the detection of abnormal learning behaviour data. Therefore, this method has the advantages of low detection error and short time cost.

4 Experimental analysis

4.1 Experimental scheme

In order to verify the effectiveness of the proposed learning behaviour data anomaly detection based on discrete Markov chain model, the simulation experiment is divided into three parts

4.1.1 Experimental environment

The experimental platform is MATLAB 7.2 platform, the operating system is WINDOWS XP system, the running memory of the experimental system is 32 GB, and the core processor is 3.6 GHz.

4.1.2 Experimental parameters

In order to verify the reliability of this method, the data used in this experiment are all from the network. The data acquisition parameters are shown in Table 1.

In order to ensure the accuracy of the experiment, the selection of experimental parameters is based on the experimental correlation, and the experimental results are the average results after several iterations. On this basis, the collected data are cleaned and filled, and the length of the experimental data is set to be suitable for the length of the input simulation software, so as to improve the accuracy of the simulation experiment.

Table 1 Data acquisition parameters

Parameter	*Value*
Sample data/GB	2
Normal behaviour data	2,000
Abnormal behaviour data/article	1,500
Data detection interval/s	0.2
Number of iterations/times	200

4.1.3 Experimental methods

The method proposed in this paper, the method in Li et al. (2018) and the method in Zhao et al. (2019) were used in the experiment.

4.1.4 Evaluation index

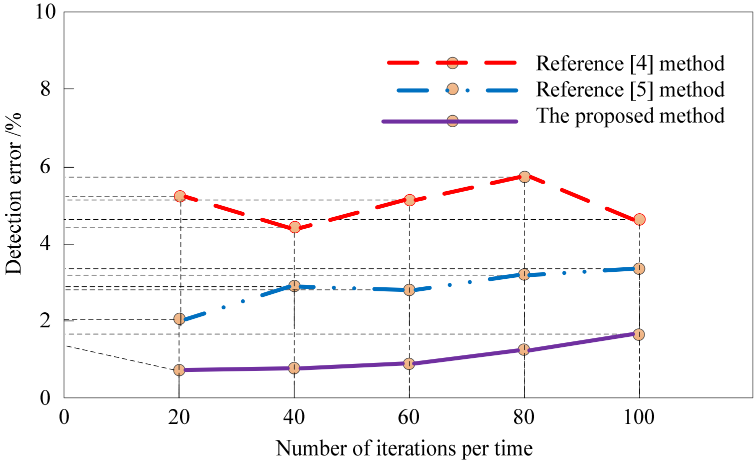
Compare the error and time cost of anomaly detection of learning behaviours data. The lower the error, the higher the accuracy of anomaly detection of learning behaviour data; the smaller the time cost, the higher the efficiency of anomaly detection. The accuracy of the experiment is guaranteed.

4.2 Analysis of experimental results

4.2.1 Comparison of detection errors

The error analysis of anomaly detection of learning behaviour data is an important measure to ensure the detection accuracy of the proposed method. The experimental results are shown in Figure 6.

Figure 6 Comparison of detection errors of learning behaviour data in three methods (see online version for colours)

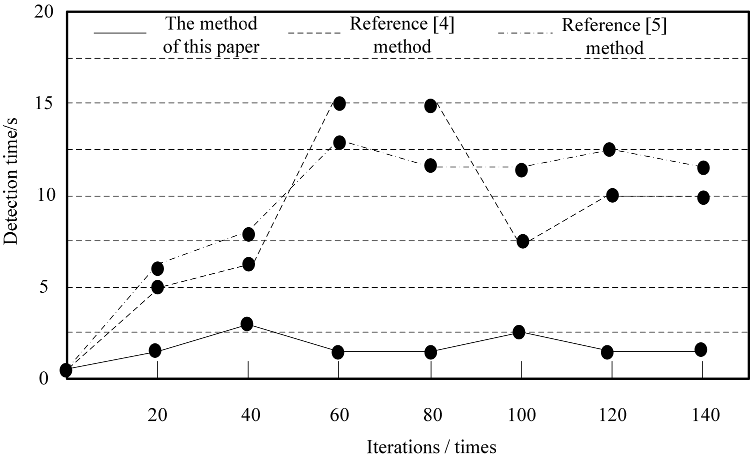


As can be seen from Figure 6, with the change of iteration times, the error of anomaly detection of sample learning behaviour data of the three methods is constantly changing. When the number of iterations is 60 When the number of iterations is 100, the error of the proposed method for anomaly detection of sample behaviour data is about 1.8%, the error of Li et al. (2018) method for anomaly detection of sample learning behaviour data is about 4.3%, the error of Zhao et al. (2019) method for anomaly detection of sample learning behaviour data is about 3.7%, when the iteration number is 100, the error of the proposed method for sample behaviour data anomaly detection is about 2%, and the error of Li et al. (2018) method for sample learning behaviour data anomaly detection is about 2%, the error of abnormal data detection is about 3.5%, and the error of the method in Zhao et al. (2019) is about 3.8%. Compared with the method in this paper, the detection error of this method is lower. This is because the proposed method extracts the characteristics of learning behaviour data, determines the abnormal amount of learning behaviour data, and uses discrete Markov chain model to detect it, which improves the performance of the proposed method Data anomaly detection accuracy.

4.2.2 Time cost comparison

In order to further verify the feasibility of the proposed method, the time cost of the proposed method, the method in Li et al. (2018) and the method in Zhao et al. (2019) for anomaly detection of sample learning behaviour data are analysed:

Figure 7 Comparison of time cost for exception detection of learning behaviour data



From the analysis of Figure 7, it can be seen that under the same experimental environment, the time cost of the three methods for anomaly detection of learning behaviour data is different. Among them, the detection cost of the proposed method is always less than 2.5s, while the detection cost of the other two methods is always higher than the proposed method. This is because the proposed method checks the sample data abnormally and repeatedly determines the abnormal behaviour of the data, so the detection time is shortened.

5 Conclusions

In view of the problems existing in the anomaly detection of learning behaviour data, this paper proposes a new detection method. This method analyses the types of learning behaviour data, and determines the influencing factors of learning behaviour data; uses support vector machine to determine the range of data extraction, sets the extraction constraints, and completes the extraction of learning behaviour data; on the basis of the extracted learning behaviour data, determines the redundancy of learning behaviour data, obtains different attributes of learning behaviour data, and makes sure that the learning behaviour data has different attributes the independent data are processed redundancy to complete the data pre-processing. This paper analyses the basic principle of discrete Markov chain, constructs discrete Markov chain model, and completes the detection of abnormal learning behaviour data. Compared with traditional methods, this method has some advantages

1 The maximum error of the proposed method is about 2%, and the error is low; the time cost of the proposed method is always less than 2.5s, and the detection speed is fast, so the method has many characteristics such as fast detection speed and low detection error.

2 This method can help university administrators find students' abnormal behaviour in time, and guide students timely and correctly. It can improve the quality of students' learning, ensure the safe and stable operation of universities, and has significant reference significance for the field of modern education. In the future, we need to design students' correct guidance strategies combined with this research, in order to comprehensively improve the quality of students' learning.

Acknowledgements

This work was supported by Key Entrustment Project of Heilongjiang Higher Education Teaching Reform Research grant no. SJGZ20190067, Archival Project of the 13th Five-Year Plan for Education Science of Heilongjiang grant no. GJC1318118, Research project of Science and Technology Bureau of Qiqihar under grant no. GYGG-201626, Scientific Research Project of Basic Scientific Research Operating Expenses of the Education Department grant no. 135309464 and 135409421.

References

Bi, M., Wang, A., Xu, J. and Zhou, F-C. (2018) 'Anomaly behavior detection of database user based on discrete-time Markov chain', *Journal of Shenyang University of Technology*, Vol. 40, No. 1, pp.70–76.

Chen, H., Wang, G. and Song, J. (2018) Research on anomaly behavior classification algorithm of internal network user based on cloud computing intrusion setection data set', *Netinfo Security*, Vol. 15, No. 3, pp.1–7.

Du, G.Y. and Chen, M.J. (2018) 'Research on anomaly detection algorithm of moving objects based on intelligent video analysis', *Video Engineering*, Vol. 42, No. 12, pp.23–26.

Guo, Z., Peng, H., Niu, S., Shao, K., Lyu, Z. and Wang, W. (2018) 'Analyzing user and network behaviors for host-based anomaly detection', *Journal of Beijing Jiaotong University*, Vol. 42, No. 5, pp.40–46.

Hu, Z., Zhao, M. and Xin, B. (2020) 'Video anomaly detection algorithm combining global and local video representation', *Pattern Recognition and Artificial Intelligence*, Vol. 33, No, 2, pp.133–140.

Li, H-B., Li, Q., Tang, R-M., Wu, J., Lv Z-Y., Pei. D., Shi, J-J., Dong X., Fand, S-D., Yang Y-F. and Wu, Y. (2018) 'User behavior anomaly detection for database based on unsupervised learning', *Journal of Chinese Computer Systems*, Vol. 39, No. 11 pp.246–-2472.

Li, S., Li, R-Q. and Yu, C. (2018) 'Evaluation model of distance student engagement: based on LMS data', *Open Education Research*, Vol. 24, No. 1, pp.91–102.

Liu, X-Y., Ye, S-P. and Zhang, D-H. (2020) 'Improved multi-objective regression student classroom action detection method', *Computer Engineering and Design*, Vol. 41, No. 9, pp.2684–2689.

Qiu, H-W. and Li, X-Y. (2018) 'Simulation of accurate detection of abnormal data in interactive network', *Computer Simulation*, Vol. 35, No, 5, pp.375–378.

Sun, L-H., and Lu, Y. (2018) 'Multivariate autoregression based algorithm for anomaly detection in rating data', *Computer Engineering and Design*, Vol. 39, No. 6, pp.1629–1652.

Tao, T., Zhou, X., Ma, B. and Zhao, F. (2019) 'Abnormal time series data detection of gas station by Seq2Seq model based on bidirectional long short-term memory', *Journal of Computer Applications*, Vol. 39, No, 3, pp.924–929.

Wang, X., Feng, A., He, F., Ma, H. and Yang, J. (2020) 'Research of database user behavior anomaly detection based on k-means and naive bayes', *Application Research of Computers*, Vol. 37, No. 4, pp.1128–1131.

Yu, L-H., Zhang, K., Cai, Y. and Jing, H-F. (2019) 'Abnormal behavior detection algorithm of moving target', *Computer Engineering and Design*, Vol. 40, No. 12, pp.3443–3450.

Zhang, M., Yu, Z-W., Guo, B., Ren, S-Y. and Yue, C-G. (2019) 'An anomaly detection system based on express big data', *Computer Engineering and Science*, Vol. 41, No, 2, pp.224–232.

Zhao, H., Liu, Y., Li, S., Xu, P. and Zheng, Q. (2019) 'Online learning behavior based personality recognition', *Education Research*, Vol. 25, No. 5, pp.110–120.