



RESEARCH ON STUDENT BEHAVIOR ANALYSIS AND GRADE PREDICTION SYSTEM BASED ON STUDENT BEHAVIOR CHARACTERISTICS

QIANG FU*

Abstract. In the era of big data, the traditional governance model of student behavior management gradually shows the disadvantage of “post positioning”. Therefore, the research uses relevant data mining algorithms to extract and analyze students’ behavior characteristics, and constructs a SPC system model for students’ behavior analysis and performance prediction. At the same time, the experiment verifies its effectiveness. The experimental results show that in the factor analysis, the overall variance of the first seven indicators is 69.942%, which shows that there is a significant correlation between students’ learning behavior and academic performance. In the model performance analysis of SPC system, the accuracy of the five algorithms is kept at 40% - 60%, while SVM has higher stability than other methods. In addition, the prediction accuracy and response rate of SPC model reached 86.90% and 81.57% respectively. When the sequence length was increased from 1 to 20, the accuracy of SPC model exceeded 70%; However, when the feature dimension exceeds 50, the model representation ability will decline. Therefore, appropriate feature dimensions are needed to predict student performance. To sum up, the SPC model built by the research is effective in analyzing student behavior and predicting students’ performance, and practical in actual student management.

Key words: Internet of Things; path Behavior characteristics; Performance prediction; Data mining; SPC model

1. Introduction. The progress of science and technology has made people’s lives increasingly networked and information-based, and the information management system has also replaced the traditional text recording method with the development of technology, so as to promote the exponential growth of the recorded data on the trajectory of human daily behavior [1]. In education management, the characteristics of students’ behavior data are effective indicators to promote the improvement of school teaching. As an effective tool for extracting data features, data mining technology has been widely used in education management [2]. Trakunphutthirak et al. used a progressive temporal data mining method of educational information on the basis of data mining technology to effectively improve the prediction accuracy of students’ academic performance [3]. To effectively evaluate the learning achievements of students in online distance education, Bütüner et al. conducted in-depth research on artificial neural networks and deep learning algorithms in data mining methods, so as to effectively improve the prediction accuracy of students’ learning achievements [4]. Shreem et al. have improved on the basis of genetic algorithm to achieve effective mining of educational data and effective prediction of student performance, thus effectively realizing the prediction classification of student learning performance [5]. Under this background, based on the recurrent neural network (RNN), a sequence-based performance classifier (SPC) was built by using support vector machines (SVM) and a hybrid encoder decoder network (HRNN) built on the basis of attention. The purpose is to effectively analyze students’ behavior, so as to effectively predict their achievements, so as to provide help for teachers to improve teaching and help students to improve their achievements.

2. Related Work. The promotion of digital application system enables students’ learning dynamics and life trajectory to be recorded in an all-round way in the form of digital information. These data are of great help to education managers in analyzing students’ activity trajectory [6]. In the field of education management, the ultimate purpose of analyzing students’ behavior characteristics is to improve students’ academic performance and promote their all-round development, and to teach based on the characteristics of each student [7]. Therefore, the majority of scholars at home and abroad have conducted in-depth research on students’ performance prediction. To realize the early prediction of students’ performance, Xiong et al. organically combined convolutional neural network and recurrent neural network to propose a mixed depth learning model,

*Principal’s Office, Zhengzhou Finance School, Zhengzhou, 450007, Henan China (hq18115139713@126.com)

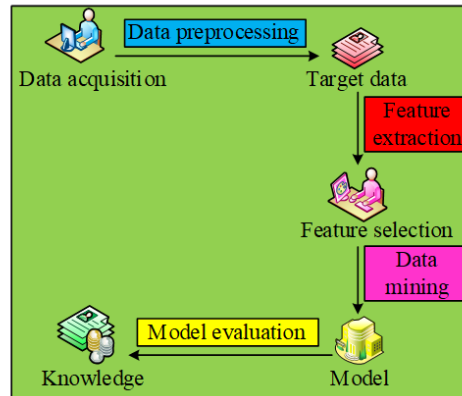


Fig. 3.1: The process of data mining

which accurately predicted students' behavior and improved their learning performance [8]. To promote the development of education, Keser et al. proposed a new hybrid integrated learning algorithm using data mining technology, so as to achieve effective prediction of student performance and help the development of education [9]. To explore the impact of social media on students' performance, Nti et al. built a prediction framework for learning students' performance by using the learning algorithms of decision tree (DT) and random forest (RF), thus realizing effective prediction of students' performance [10]. Teng M F et al. conducted a confirmatory analysis of students' English writing scores by using the data collection method based on the metacognitive academic writing strategy questionnaire, thus achieving an effective prediction of students' English writing performance [11]. In addition, to improve the quality of students' academic achievements in the School of Education, Britum and others conducted statistical tests to solve the problem of students' self-esteem, so as to effectively improve students' achievements on the basis of improving their sense of self-efficacy [12]. Serrano et al. analyzed in detail the role of the five-factor model on the personality level to improve the accuracy of the prediction of students' academic performance [13]. Su has effectively improved the data processing algorithm on the basis of machine learning and neural network algorithm to improve the accuracy of learning achievement prediction [14]. To improve students' academic performance, Deeba et al. analyzed the self-concept of English majors in detail, thus providing help for improving students' self-concept and achieving effective prediction of performance [15].

From the research of scholars at home and abroad, the current research method is mainly to manually extract the statistical characteristics of data, which cannot make deep use of effective information. Therefore, on the basis of correlation analysis of campus behavior data, the proposed performance prediction modeling method based on student behavior sequence makes full use of the attributes of education data. On the basis of using data mining technology, it fully makes up for the shortcomings of traditional methods, and is innovative to a certain extent.

3. Research on Student Behavior Analysis and Grade Prediction System Based on Student Behavior Characteristics.

3.1. Analysis of relevant theoretical algorithms of student behavior characteristics. To help teachers improve students' academic performance in a personalized way, the research constructs a classification system model of performance prediction based on students' behavior by extracting students' behavior characteristics. Data mining technology is widely used in student behavior feature extraction. Data mining refers to the use of algorithms to find information hidden in the massive data. It is usually deeply related to computer science, and can effectively mine data through statistical analysis, machine learning, etc [16]. The current data mining process has formed a relatively mature process system, as shown in Figure 3.1.

From Figure 3.1, the process of data mining starts with data collection; Secondly, the target data is

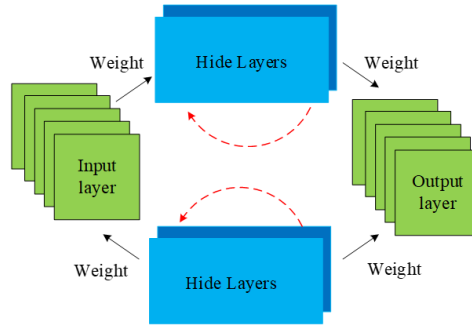


Fig. 3.2: Structure diagram of recurrent neural network

obtained after data preprocessing; Then the feature of target data is extracted to complete the feature selection; Then the corresponding model is obtained through data mining; Finally, the model is evaluated to obtain the corresponding knowledge. In the actual data mining, SVM is selected to analyze and predict the student’s behavior sequence and performance. SVM is a generalized binary classification method based on supervised learning. Its decision boundary is the maximum margin hyperplane. On the basis of structural risk minimization, it is formalized to solve a convex quadratic programming problem, and its local optimal solution must be the global optimal solution [17]. In addition, when SVM is linearly nonseparable, a kernel function is introduced to solve the problem of sample classification. On the sequence of students’ learning behavior, the research chooses the sequence model of RNN to model the characteristics of students’ short-term campus behavior sequence. RNN can capture the characteristics of the input sequence and adapt to various length changes by iteratively updating its internal hidden state. Its structure is shown in Figure 3.2.

From Figure 3.2, a basic RNN is composed of three layers, namely the input layer, hidden layer and output layer. In RNN, the connecting line between the input layer, hidden layer and output layer not only appears among the three, but also exists between multiple hidden layers in the upper time dimension. Among them, the expression of the update value formula of the hidden layer is shown in equation 3.1.

$$h_1 = f(h_{t-1}, x_t) \tag{3.1}$$

In equation 3.1, h represents the hidden state; $f(\cdot)$ represents the nonlinear function; t represents time; x represents the input sequence. Specifically, the expression of the state calculation process of the hidden layer is shown in equation 3.2.

$$h_t = f(Ph_{t-1} + Qx_t + b) \tag{3.2}$$

In equation 3.2, P represents the weight matrix of the state of the hidden layer; Q represents the weight matrix of the state from the hidden layer to the input layer; b represents the bias item. Among them, the parameters of RNN model have the same weight in the time dimension, that is, under different time scales, each parameter of the model has the same consistency, thus reducing the training parameters of the model [18]. The learning of network parameters is usually updated by back-propagation algorithm. Therefore, RNN has the ability of short-term memory when processing time series samples of any length, and its optimal variable is the optimal parameter of the recurrent neural network.

3.2. Research on short-term behavior feature extraction of students based on attention. The research chooses to mine and analyze the data from the perspective of classification and statistical characteristics, so as to explore the inner relationship between students’ behavioral characteristics and academic performance and realize the prediction of their performance. On this basis, the study uses relevant statistical analysis

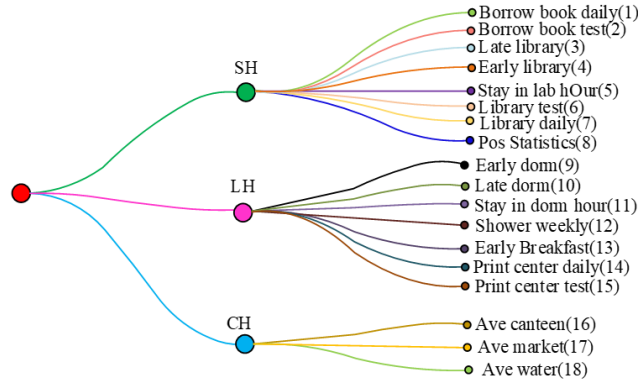


Fig. 3.3: Types of student characteristics proposed by traditional methods

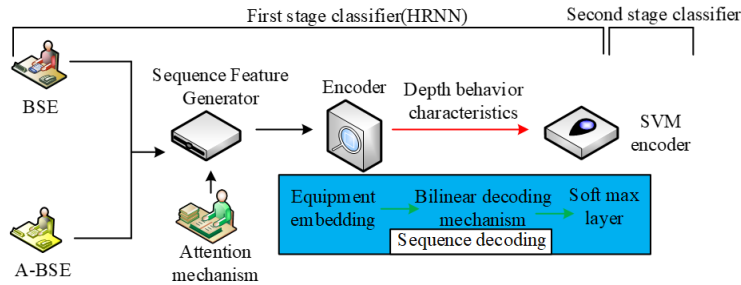


Fig. 3.4: Schematic Diagram of SPC Model Structure

methods and selects student behavior attributes related to grade rankings to extract 18 characteristics, the contents of which are shown in Figure 3.3.

From Figure 3.3 that students' behavior characteristics are mainly divided into three categories, namely, study habits (SH), living habits (LH) and consumption habits (CH). Among them, SH includes the times of borrowing books in the library during non exam and exam time, the times of entering the library before 8 a.m., the times of leaving the library at 22 p.m., the average daily time of staying in the library, the times of going to the library during non exam and exam time, and the times of using point of sales (POS) machines during class. The LH includes the times of leaving the dormitory between 8 a.m., returning to the dormitory after 22 p.m., daily average times of staying in the dormitory, weekly times of getting wet, times of eating breakfast between 8 a.m., and times of going to the printing center during non examination and examination time. CH includes the daily average times of canteen consumption, supermarket consumption and water fetching. In addition, on the basis of sequences, the study introduced a Hybrid Recurrent Neural Network (HRNN) model based on attention in the first stage of student performance prediction, which is used as the first learning classifier to learn the sequence characteristics of students' behavior. The Sequence based Performance Classifier (SPC) is shown in Figure 3.4.

From Figure 3.4 that the SPC model consists of two-stage classifiers. Among them, the base sequence encoding is combined with the output of the attention sequence encoder, which is input into the sequence feature generator. Both are basic networks built on the basis of Gate Recurrent Unit (GRU). The specific process expressions are shown in equation 3.3 and equation 3.4.

$$\begin{cases} z_t = \rho(W^{(z)}x_t + P^{(z)}h_{t-1}) \\ r_t = \rho(W^{(r)}x_t + P^{(r)}h_{t-1}) \end{cases} \quad (3.3)$$

In equation 3.3, Z_t represents the update gate; ρ represents the S -type growth curve function (Sigmoid); W represents the weight matrix of the state between hidden layers; r represents the reset gate.

$$\begin{cases} h'_t = \tan(Wx_t + r_t \odot Ph_{t-1}) \\ h_t z_t \odot h_{t-1} + (1 - z_t) \odot h'_t \end{cases} \quad (3.4)$$

According to the analysis of equation 3.3 and equation 3.4, this process linearly interpolates the existing hidden state with the existing hidden state, and the final hidden state contains all the information of all the original sequences. In fact, the last implicit state is used to express the behavior order of students, which is called Basic Sequence Encoder (BSE). The expression is shown in equation 3.5.

$$c_t^g = h_t = h_t^g \quad (3.5)$$

In equation 3.5, c_t^g denotes the base sequence encoder; h_t^g denotes the final hidden state. According to equation 3.5, it can be found that not all student behaviors are linked to their learning performance, so in the actual prediction, the SPC model can pay more attention to the behavioral interaction related to student performance. Based on this, the attention-based sequence encoder (Attention-based Sequence Encoder, A-BSE) is studied, and its expression is shown in equation 3.6.

$$c_t^l = \sum \alpha_{ti} h_i \quad (3.6)$$

In equation 3.6, c_t^l represents the context vector; α_{ti} represents the weighting factor; i represents a specific moment of time. Wherein, c_t^l enables the decoder to dynamically generate and linearly combine different parts of the input sequence. The calculation expression of the attention mechanism function is shown in equation 3.7.

$$\alpha_{ti} = \rho(W_\alpha[h_t; h_i]) \quad (3.7)$$

In equation 3.7, ρ the expressed Sigmoid function does not add up all hidden states learned from the RNN network to represent the student's behavior sequence. Instead, it is important for coders to use weight coefficients to represent hidden states / interactions. Finally, the student's behavior sequence is expressed by weighting and summing these hidden states. In order to better understand the context vector, its modified expression is shown in equation 3.8.

$$c_t^l = \sum \alpha_{ti} h_i = \sum \alpha_{ti} h_t^l \quad (3.8)$$

In equation 3.8, the last hidden state of BSE is used to encode the entire sequence behavior, and A-BSE is used to calculate the attention weight value of the previous hidden state. This hybrid approach can be built into the same generator, the sequential feature generator in the SPC model. Its expression is shown in equation 3.9.

$$c_t = [c_t^g; c_t^l] = [h_t^g; \sum_{i=1}^t \alpha_{ti} h_t^l] \quad (3.9)$$

On the basis of equation 3.9, the study applies an optional bilinear decoding mechanism between the current implicit representation and each campus card device to calculate the similarity score. The expression of this score is shown in equation 3.10.

$$S_i = \text{emb}_i^m T c_t \quad (3.10)$$

In equation 3.10, S_i represents the similarity score; T represents the combination matrix composed of the campus card embedding dimension and the sequence representation dimension. In sequence prediction, BSE is used to summarize students' weekly activities, while A-BSE can automatically select corresponding behaviors according to students' learning habits, so as to understand students' learning motivation.

3.3. Analysis of grade prediction classification model based on student behavior. In the SPC model constructed by the study, HRNN was selected for the first stage classifier study, while the SVM classifier was selected for the second stage study. It can find an optimal compromise between model complexity and machine learning ability by using limited sample data, so that it is more suitable for linear inseparable problems in multi-category situations. Among them, in the actual situation of linear separability, SVM will try to find the optimal classification hyperplane to maximize the interval separation. The expression of the hyperplane is shown in equation 3.11.

$$w^T \cdot x' + b = 0 \quad (3.11)$$

In equation 3.11, w denotes normal vector; T denotes transposition; x' denotes feature. Before finding the hyperplane, a quadratic programming problem needs to be solved first, and the expression of the problem is shown in equation 3.12.

$$\begin{cases} \min \Phi(w) & \frac{1}{2} \|w\|^2 \\ \text{subject to} & y_j [(w^T \cdot x'_j + b) - 1] \geq 0 \end{cases} \quad (3.12)$$

In equation 3.12, y_j represents the result label; j represents the dimension. At this time, the quadratic programming problem can be solved according to the Lagrangian dual solution method, and its expression is shown in equation 3.13.

$$\min L(w, b, \alpha') = \frac{1}{2} \|w\|^2 - \sum_{j=1}^l \alpha'_j [(w^T \cdot x'_j + b) - 1] \quad (3.13)$$

In equation 3.13, α' represents the Lagrangian multiplier. According to equation 3.13, the relationship between normal vector, offset term and Lagrangian multiplier can be obtained by using differential formula and reduction method, and its expression is shown in equation 3.14.

$$\begin{cases} \max & W(\alpha') = \sum_{j=1}^l \alpha'_j - \frac{1}{2} \sum_{k,j=1}^l \alpha'_j \alpha'_k y_j y_k x_j^T x_k \\ \text{subject to} & \sum_{j=1}^l \alpha'_j, \alpha'_j \geq 0 \end{cases} \quad (3.14)$$

Equation 3.14 is a quadratic function optimization problem with inequality constraints. Its objective function and linear constraint are convex functions, and there is a unique solution. The final expression of the optimal classification hyperplane is shown in equation 3.15.

$$f(x) = \sum_{j=1}^l (\alpha_j^* y_j x_j^T x + b^*) \quad (3.15)$$

In equation 3.16, $f(x)$ represents the optimal classification hyperplane function; α_j^* represents the support vector point. In the case of linear inseparability, you can choose to use kernel functions to map features into higher dimensions to achieve the goal of linear separability. Therefore, using a nonlinear mapping method to map the training samples to a high-dimensional feature space can make the nonlinear classification become a linear classification in the input space. The expression of this category is shown in equation 3.16.

$$\begin{cases} \min \Phi(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{j=1}^l \xi_j C > 0 \\ \text{subject to} \xi_j > 0, \quad y_j (w^T \cdot x'_j + b) \geq 1 \end{cases} \quad (3.16)$$

In equation 3.16, Φ represents the nonlinear mapping; $C \sum_{j=1}^l \xi_j$ represents the penalty term. The solution process of equation 3.16 is similar to the case of linear separability. At this time, in the high-dimensional space, according to the corresponding kernel function, a linear classifier is used to implicitly construct a classification surface in the high-dimensional space. Based on the theory of structural risk minimization, SVM constructs the optimal segmentation hyperplane in the feature space, so that learners can obtain the global optimal solution,

Table 4.1: Data field information content of experiment part

-	Student Library Borrowing Data						
-	1	2	3	4	5	6	7
Field Description	Student ID	Hours of loan service	Title	Call number	-	-	-
-	Data of student all-in-one card system						
Field Description	Student ID	Consumption category	Consumption place	Consumption mode	Dissipate	Consumption amount	Remaining amount
-	Student dormitory access control data						
Field Description	Student ID	Time of entry and exit	In and out direction (0 in and 1 out)	-	-	-	-
-	Access control data of student library						
Field Description	Student ID	Time of entry and exit	Access control No	-	-	-	-
-	Student performance ranking data						
Field Description	Student ID	College No	Score ranking	-	-	-	-

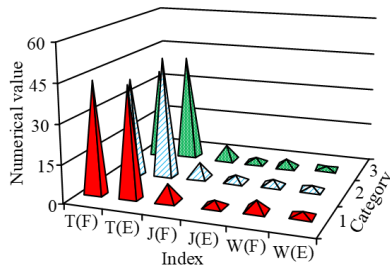
and the required overall risk meets a certain upper bound. In solving multi-classification problems, the research chooses to classify every two categories. According to Figure 4, the actual input of SCM is the complete sequence behavior feature constructed in the HENN classifier network model, and the output is the real grade ranking of students. Therefore, students' learning performance can be regarded as a short-term sequential model. By analyzing the weekly behavior of a specific student and using the SPC two-level classification system to predict it, students with learning crisis can be detected in time.

4. Student behavior data and grade prediction analysis based on student behavior characteristics. To better explain the behavior data extracted from the study, the study used Chint distribution method to divide students' scores into three grades, namely, good, medium and poor, which are represented by A, B and C. Before the experiment, the relevant data were collected and cleaned. It is worth noting that the content of the data set used in the study is the behavior data of students using the card in and out of the campus within two school years and the student score ranking data in the school's teaching management system. Some data field information is shown in Table 4.1.

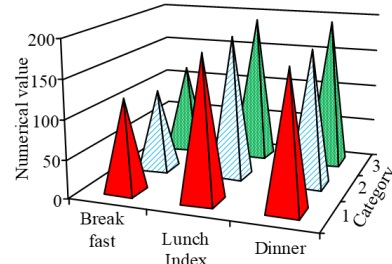
In Table 4.1, the numbers 1-7 are the field serial numbers. From Table 4.1, part of the data field information is mainly divided into four categories, namely, student library borrowing data, student card system data, student dormitory access control data, student library access control data, and student grades ranking data. Among them, the "consumption pattern" field in the all-in-one card data is the focus of research. When there is a "default" in the original data, the row where the "user" is located should be "deleted". If the borrowing time, consumption time and entry and exit time fields are the same twice, the row where it is located will be deleted. Because repeated records will analyze the data, resulting in errors. On this basis, the study combines the feature classification results in Figure 3 to give three levels of students who go to the library (T), borrow books (J) and go to the printing center (W) during the examination and non-examination periods. The number of times was compared, and the three meals of the students of the three levels were compared at the same time, and the results are shown in Figure 4.1.

In Figure 4.1, E and F represent examination and non-examination respectively. Figure 4.1b In order to reduce the actual error, the students of No. 1 college of this school were selected for analysis. From Figure 4.1a that the students of grade A borrow 52 books on average, the students of grade B borrow 48 books, and the students of grade C borrow 42 books on average. Likewise, as the grade decreases, the number of activities such as studying in the library decreases gradually. It shows that the better the students' academic performance, the more practical and feasible they are in their daily life on campus. From Figure 4.1b that the lunch of the three types of students is much larger than the breakfast, indicating that there is an irregular diet in the student group. In addition, in the comparison of the three types of students, students with grade A have the most breakfasts, which shows that students with academic performance have relatively good eating habits and are more self-disciplined. To verify the correlation between the behavior characteristics of students in Figure 5 and their academic performance, the study further used the statistical methods of factor analysis and principal component analysis to interpret the correlation between the two. Among them, the study used the measure (Kaiser-Meyer-Olkin, KMO) and Bartlett's sphericity (Bartlett's) test to select features when conducting factor analysis. The result is shown in Figure 4.2.

From Figure 4.2a that the statistic of KMO is 0.720, which is greater than 0.6, indicating that the research experiment can be used for factor analysis. At the same time, the approximate chi square value obtained by

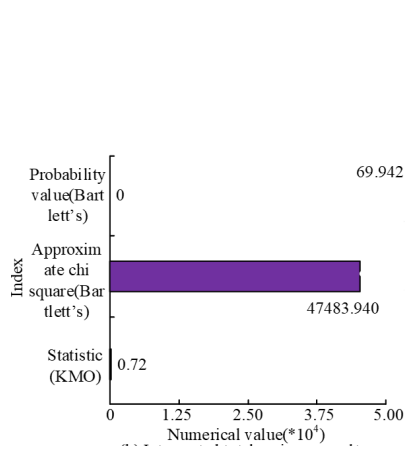


(a) Results of behavior comparison among three types of students

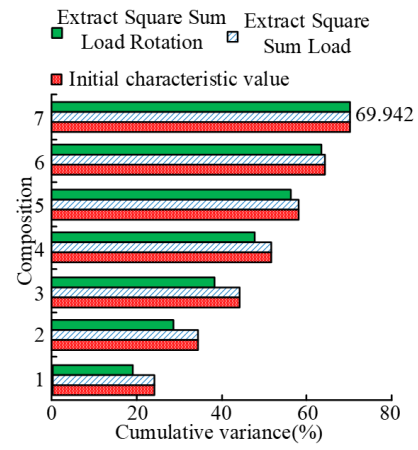


(b) Three Meals for Three Kinds of Students

Fig. 4.1: The number of times that three types of students go to three places and the diet comparison results of three meals



(a) Interpreted total variance results



(b) Interpreted total variance results

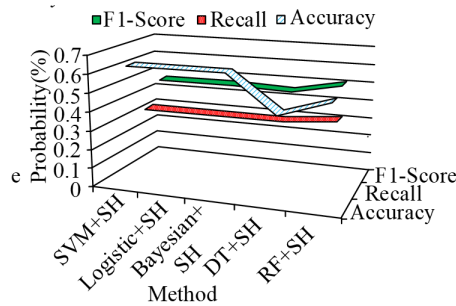
Fig. 4.2: KMO and Bartlett's test results and explained total variance results

Bartlett's test is 47483.940, and the probability value at this time is 0.000, less than 0.05, reaching the level of significance. The result shows that there is correlation between variables. Therefore, according to Figure 4.2b obtained by factor analysis in Figure 4.2a, the overall variance of the first seven factors is 69.942%, which meets the actual needs, effectively reflects the overall information and shows that there is a significant relationship between students' learning behavior and academic performance. On this basis, based on the results obtained in Figure 4.2 and the traditional behavior characteristics, the research built a multi category prediction model of Bayesian, logistic, DT, RF and SVM, and used three evaluation methods, namely, accuracy, recall and F1 score, to evaluate the prediction models of these students' scores. The result is shown in Figure 4.3.

From Figure 4.3 that the accuracy of the five methods is maintained between 40% and 60%, the recall rate is maintained at about 35%, and F1 Core is maintained at about 45%. In general, Bayesian model has the worst classification effect, the overall accuracy of logical regression and SVM is relatively more stable, and has a stronger mathematical interpretation of the research data. At the same time, it also proves that it is feasible to predict students' performance by using their behavior data, and verifies the effectiveness of SVM. After verifying the feasibility of the research direction, the research continues to analyze the actual performance of



(a) Classification of learning habits by different methods (b) Classification of living habits by different methods



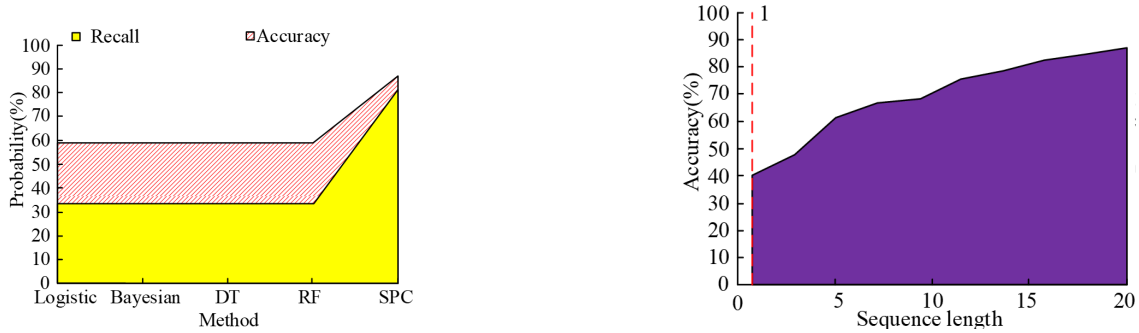
(c) Classification of consumption habits by different methods

Fig. 4.3: Results of the five traditional classification methods

the SPC model system and its feature modeling effect on the student behavior sequence. To reduce the impact of uncertain data on the actual results, the study extracted the behavior records of attacking more than 9000 students in 29 weeks. The performance verification test results are shown in Figure 4.4.

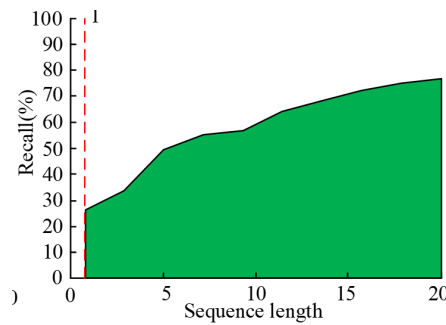
From Figure 4.4 that the SVM based on RNN can handle the sequences in multiple sequences well and complete the classification task efficiently in the process of processing multiple sequences. After considering the students' sequential behavior and the main goals of learning, the SPC proposed in the study can exceed all the benchmarks. The relative performance accuracy is 86.90%, and the response rate is 81.57%. In addition, when the length of behavior sequence increases from 1 to 20, the performance of SPC model exceeds 70% accuracy, indicating that it can achieve better prediction. On this basis, the research experiment analyzes the accuracy rate and recall rate of SPC model system in different cycles and dimensions, and the results are shown in Figure 4.5.

From Figure 4.5 that when the sequence length of the SPC model is divided by month, the excessive sequence length makes it impossible for managers to intervene in students' learning in time, resulting in a decline in accuracy, which is only 71.40%. It also shows that short-term continuity modeling can better predict students' learning situation and performance. In addition, as the feature dimension increases from 5 to 100, the performance of the model has been greatly improved, with the maximum accuracy rate exceeding 75% and the recall rate exceeding 60%. At the same time, the increase in the number of hidden units also prompted the curve of the two indicators to gradually show a gentle trend. Therefore, although the increase of hidden layer feature dimension can more comprehensively grasp the situation of students, the representation ability of the model will decline when the dimension exceeds 50. Therefore, it is necessary to set appropriate feature dimensions to better predict students' academic performance. And let educators better help students.



(a) Comparison of accuracy and recall of different methods

(b) Accuracy results of SPC model under different sequence lengths

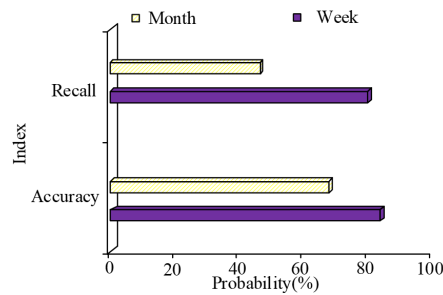


(c) Recall rate results of SPC model under different sequence lengths

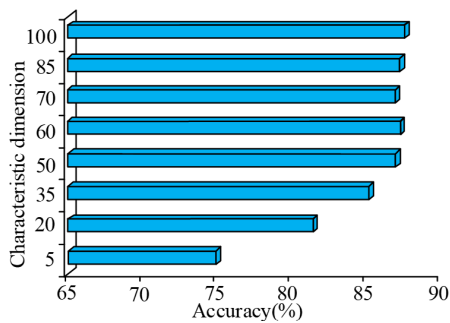
Fig. 4.4: SPC model performance test results

5. Research limitations and prospects. The selection of data mining in research still needs to be optimized. With the rise and development of artificial intelligence technology, applying artificial intelligence algorithms to SPC models may further improve their performance. At the same time, due to many restrictions, the research only applies it to education data. In fact, in the era of Big data, it can be popularized in other fields, such as short video user data, e-commerce, but it is undeniable that the research data can only be applied in education, and is affected by different educational environments. Therefore, in the future, the application of AI algorithms, Big data technology, cloud computing and other cutting-edge technologies in the SPC model will further enhance its applicability in the educational environment and expand.

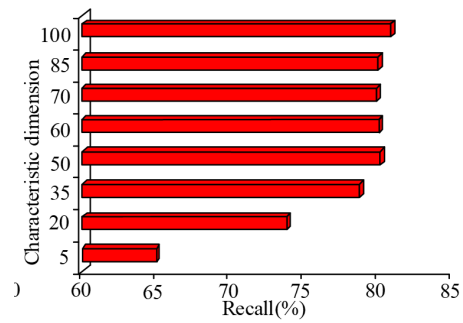
6. Conclusion. To realize the effective prediction of students' grades and help teachers to improve students, the study extracts the behavior characteristics of students, and constructs the SPC system model on the basis of HRNN and SVM, and analyzes the behavior of students. The features and performance of the SPC model were analyzed experimentally. The results of the experiment show that the average number of books borrowed by grade A students is 52, which is higher than that of the other two grades. At the same time, the number of breakfasts is also the most, which shows that students with good academic performance have good study habits. In the factor analysis, the overall variance of the first seven factors reached 69.942%, indicating that there is a significant correlation between students' learning behavior and academic performance. Therefore, in the performance analysis of the SPC system model, the accuracy rate of the five methods is maintained between 40% and 60%, the recall rate is maintained at about 35%, and the F1-Score is maintained at about 45%. At the same time, the stability of SVM is the highest, indicating that it has high effectiveness. In addition, the accuracy rate of the SPC model reached 86.90%, and the response rate reached 81.57%, both



(a) SPC Model Accuracy and Recall Results in Different Periods



(b) SPC model accuracy results under different feature dimensions



(c) SPC model recall results under different feature dimensions

Fig. 4.5: Accuracy and recall results of SPC model in different periods and dimensions

higher than other methods. When the sequence length and feature dimension are increased to 20 and 100, respectively, the accuracy of the model exceeds 70%, and the highest exceeds 75%. On the whole, the SPC system model constructed by the research is effective in extracting student behavior characteristics for performance prediction. However, the selection and application of data mining technology in the research has yet to be optimized and needs to be improved in the future. of SVM is the highest, indicating that it has high effectiveness. In addition, the accuracy rate of the SPC model reached 86.90%, and the response rate reached 81.57%, both higher than other methods. When the sequence length and feature dimension are increased to 20 and 100, respectively, the accuracy of the model exceeds 70%, and the highest exceeds 75%. On the whole, the SPC system model constructed by the research is effective in extracting student behavior characteristics for performance prediction. However, the selection and application of data mining technology in the research has yet to be optimized and needs to be improved in the future.

REFERENCES

- [1] Umer, R., Khan, S., Ren, J., Umer, S. & Shaykat, A. Prediction of students' failure using VLE and demographic data: case study on Open University data. *International Journal Of Business Intelligence And Data Mining*. **20**, 235-249 (2022)
- [2] AdrianChin, Y., JosephNg, P., Eaw, H., Loh, Y. & And, S. and learning behaviour dubious relationship. *International Journal Of Business Information Systems*. **41**, 548-568 (2022)
- [3] Trakunphutthirak, R. Lee V C S. Application of educational data mining approach for student academic performance prediction using progressive temporal data. *Journal Of Educational Computing Research*. **60**, 742-776 (2022)
- [4] B"ut"uner, R. & Calp, M. Estimation of the Academic Performance of Students in Distance Education Using Data Mining Methods. *International Journal Of Assessment Tools In Education*. **9**, 410-429 (2022)
- [5] Shreem, S., Turabieh, H., Al Azwari, S. & Baothmn, F. Enhanced binary genetic algorithm as a feature selection to predict

- student performance. *Soft Computing*. **26**, 1811-1823 (2022)
- [6] Mingyu, Z., Sutong, W., Yanzhang, W. & Dujuan, W. An interpretable prediction method for university student academic crisis warning. *Complex & Intelligent Systems*. **8**, 323-336 (2022)
- [7] Kittur, J., Bekki, J. & Brunhaver, S. Development of a student engagement score for online undergraduate engineering courses using learning management system interaction data. *Computer Applications In Engineering Education*. **30**, 661-677 (2022)
- [8] Xiong, S. Gasim E F M, Ying C X, Wah K K, Ha L M. *A Proposed Hybrid CNN-RNN Architecture For Student Performance Prediction*. **10**, 347-355 (2022)
- [9] Keser, S. & Aghalarova S. Helal, A. novel hybrid ensemble learning algorithm for predicting academic performance of students. *Education And Information Technologies*. **27**, 4521-4552 (2022)
- [10] Nti, I., Akyeramfo-Sam, S., Bediako-Kyeremeh, B. & Agyemang, S. Prediction of social media effects on students' academic performance using Machine Learning Algorithms (MLAs). *Journal Of Computers In Education*. **9**, 195-223 (2022)
- [11] Teng, M., Qin, C. & Wang, C. Validation of metacognitive academic writing strategies and the predictive effects on academic writing performance in a foreign language context. *Metacognition And Learning*. **17**, 167-190 (2022)
- [12] Britwum, F., Amoah, S., Acheampong, H. & Adiei, E. Self-esteem as a predictor of students' academic achievement in the colleges of education. *International Journal Of Learning And Teaching*. **14**, 29-40 (2022)
- [13] Serrano, C., Murgui, S. & Andreu, Y. Improving the prediction and understanding of academic success: The role of personality facets and academic engagement. *Revista De Psicodidáctica (English Ed)*. pp. 21-28 (2022)
- [14] Su, Y., Wang, S. & Li, Y. Research on the improvement effect of machine learning and neural network algorithms on the prediction of learning achievement. *Neural Computing And Applications*. **34**, 9369-9383 (2022)
- [15] Deeba, F., Ullah, S. & Saleem, A. An Analysis of Students' Academic Self-Concept in English as Predictor of their Academic Performance. *Pakistan Journal Of Humanities And Social Sciences*. **10**, 403-415 (2022)
- [16] Hidalgo, Á., Ger, P. & Valentín, L. Using Meta-Learning to Predict Student Performance in Virtual Learning Environments. *Applied Intelligence*. **52**, 3352-3365 (2022)
- [17] Ajibade, S., Dayupay, J., Ngo-Hoang, D., Oyebode, O. & Sasan, J. Utilization of Ensemble Techniques for Prediction of the Academic Performance of Students. *Journal Of Optoelectronics Laser*. **41**, 48-54 (2022)
- [18] Shin, J., Chen, F., Lu, C. & Bulut, O. Analyzing students' performance in computerized formative assessments to optimize teachers' test administration decisions using deep learning frameworks. *Journal Of Computers In Education*. **9**, 71-91 (2022)

Edited by: Mudasir Mohd

Special issue on: Scalable Computing in Online and Blended Learning Environments: Challenges and Solutions

Received: May 15, 2023

Accepted: Nov 16, 2023