Learning Analysis Based on Learners Learning Model

Wang Wangzhu[1], Liao Zhixin[2, *], Deng Yi[2], Xu Song[2], Guo Xiaoyu[2], Ye Junmin[2, *]

[1]School of Foreign Languages, South-Central University for Nationalities, Wuhan, China

[2]School of Computer, Central China Normal University, Wuhan, China

Email address:
liliwwz@163.com (Wang Wangzhu), zhixinl@mails.ccnu.edu.cn (Liao Zhixin), 798815708@qq.com (Xu Song),
425859679@qq.com (Guo Xiaoyu), 769896547@qq.com (Ye Junmin)
[*]Corresponding author

To cite this article:

Wang Wangzhu, Liao Zhixin, Deng Yi, Xu Song, Guo Xiaoyu, Ye Junmin. Learning Analysis Based on Learners Learning Model. *International Journal of Elementary Education*. Vol. 7, No. 1, 2018, pp. 1-6. doi: 10.11648/j.ijeedu.20180701.11

Abstract: In the era of big data and artificial intelligence, collecting and analyzing learners learning data can modle and predict learners learning trend and help learners avoid risks of academic failures. For those purposes, this paper presents a learning analysis method based on learners learning model. First, the teaching model of the curriculum is put forward to support the learning data analysis. Secondly, various methods including questionnaire are used in data collection and quantification of learners offline learning data so that all the meaningful data can be transformed into the numerical data that can be processed; thirdly, the linear fitting method is used to analyze the learning data and predict the learners learning trend. The results show that the linear fitting method can effectively describe learners learning trend.

Keywords: Learning Trends, Data Analysis, Linear Fitting

1. Introduction

Data analysis refers to the use of appropriate statistical method to analyze the large amounts of collected data, to extract useful information and to form conclusions, that is, to study the data in detail and summarize the process. In practice, data analysis can help people make judgments and take measures to make necessary adjustment. Technology for data analysis in the field of education is called learning analytics. In 2011, the first "Learning Analytics Technology and Knowledge Conference" proposed a definition of learning analytics that "measures, collects, analyzes, and reports datasets about learner learning behavior and learning environments to understand and optimize learning and its surroundings". Learning analytic techniques can be used not only for online data analysis, but also for offline data analysis.

1.1. Current Research Abroad

Learning Analytics started abroad much earlier, and researchers have achieved initial results. For example, the Horizon Report, first published in 2010 by the New Media Alliance of America and the United States Association for Higher Education for Information Technology, expands on forms of learning and predicts that new forms of learning will gradually emerge [1]. The Horizon Report 2015 confirms that quantitative learning will become one of the trends in learning analytics [2]. In 2012, the second session of "Seminar on Learning Analysis Technology and Knowledge" put forward the idea of quantifying self-reflective learning, knowing self by collecting information related to individuals, promoting continuous learning through regression and evaluation of past experiences [3]. In 2013, the 3rd session of the Conference on Learning Analysis Techniques and Knowledge proposed predictions of high-risk learners by virtualizing learning environments from historical activity data and other sources of data, by observing changes in users' activities in the environment, anticipating similar learning behavior to help learners to avoid failure [4]. In 2014, the 4th "Learning Analysis Technology and Knowledge Conference" proposed the research of early warning system, using the semester data of the learning management system and using the semester fixed-effects polynomial regression model to estimate the relationship between the curriculum resources utilization and the learners' possible evaluation [5]. The essence of these forecasting studies is to integrate and

analyze the real-time information of learners so as to obtain the learner's development status as the basis for predicting future achievements. The 5th session in 2015 and the 6th session in 2016 focused on the intersection between research and the practice and the multidisciplinary connections that effectively illustrate how learning analytics can provide critical insights into the individual and collective learning process [6]. The impact of students' exploration strategies on learning is investigated and the use of a probabilistic model jointly representing student knowledge and strategies is proposed [7]. Those analyses are based on data collected from an interactive computer-based game. Hlosta, Zdrahal and Zendulka [8] proposes a solution in the absencece of data from previous courses, focusing on the problem of identifying students who are at risk of failing their courses.

1.2. Domestic Research

Mo Jingqi [9] pointed out that the implementation of classroom teaching evaluation is to diagnose and improve teaching, so attention should be paid to the interpretation and application of evaluation results. Liu Gang [10] put forward the idea that the evaluation of the function, the subject evaluation, evaluation criteria and evaluation results help understanding the basic direction of classroom teaching evaluation reform. However, the evaluation of these classroom teaching did not state explicitly how to further analyze and apply the evaluation results. The specific methods for quantitative analysis of learner learning data, such as the PAD Class model, can collect and describe the data of the learning process, but lack a holistic assessment of learner learning status [11]. A model of learning alert based on outlier mining and analysis [12] and an intervention model based on tracing, accumulating and selecting the online learning behavior data in Moodle [13] are proposed in recent studies.

From the above research, the relevant researches mainly focus on the collection and processing of online data, because the existing technology fully supports this meaningful work. For the analysis of learner learning data from the perspective of classroom teaching mode, the research is relatively weak. The main work in this field is still primarily based on the statistical analysis of learners performance results, mainly because of the fact that learner learning data is subjective and difficult to quantify, and because of lack of an effective analytical method. On the basis of the above research background, how to further quantify the evaluation of offline teaching and analyze the data of these evaluation results to complete the overall evaluation and prediction of learner learning status is a topic worthy of attention, which is the focus of this research.

2. A Method Based on Learner Learning Model Analysis

2.1. Structure Level Course Learning Model

In view of the selected target course, studying and constructing the course learning model on the basis of the characteristics of teaching, learning, assessment and knowledge system structure is essetial to learning analytics. In this research, the course of "College English for non-English Majors, band 2" is studied. For the course's characteristics of combining knowledge with practice, the course learning model constructed is a hierarchical structure. The learning model of this course takes the knowledge unit as the basic unit and divides the activity module into two parts: in-class learning and extra-curricular learning. In-class learning includes classroom instruction, answering questions, attendance and other activities in the classroom. From the learner's perspective, the effectiveness of classroom instruction can be indicated by learner's attentiveness; answering questions refers to the learner effectively answering questions raised by the instructor in the classroom; attendance refers to the learner attending class in the classroom during the stipulated time. In this research these three types of classroom activities are given corresponding weight according to their importance. It is assumed that extracurricular learning includes extracurricular activities such as homework, the quiz, self-study, etc. The homework mainly includes doing homework and the effect of doing homework. The quiz mainly refers to finishing the quiz and the time it takes to finish it. Self-study mainly includes the preview, review and extra-curricular knowledge learning. According to the importance of three kinds of activities in extra-curricular learning, instructors can set corresponding weight according to their experience.
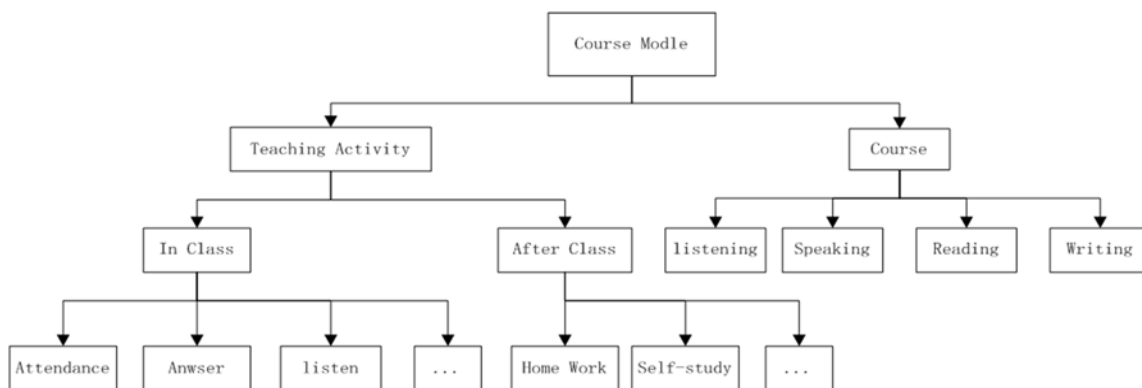

Figure 1. Structure level English course learning model.

2.2. Calculation Method

In this research the learners' overall level of listening, speaking, reading and writing is based on the learners' total score of each unit in the English course，The core calculation method adopted is the linear fitting method，results are further analyzed with the combination of clustering method. The specific method is described as follows:

1) Calculating the total score of ith-learner in unit j:

$$Sum_{unit}[i][j];$$

2) According to the linear fitting equation shown in the function (1)(2), the linear fitting linear graph of the learner's total score change is drawn. The fitting formula for each learner is given as follows;

$$\bar{y} = \hat{b} \cdot \bar{x} + \hat{a} \tag{1}$$

$$\begin{cases} \hat{b} = \frac{\sum_{j=1}^{m}\{(x_i-\bar{x})\times(y_i-\bar{y})\}}{\sum_{j=1}^{m}\{(x_i-\bar{x})^2\}} = \frac{\sum_{j=1}^{m}\{x_i \cdot y_i\}-n\cdot\bar{x}\cdot\bar{y}}{\sum_{j=1}^{m}\{(x_j)^2\}-n\cdot(\bar{x})^2} \\ \hat{a} = \bar{y} - \hat{b} \cdot \bar{x} \end{cases} \tag{2}$$

3) The interpretation of experimental results includes two aspects: first, analyzing the learning trend of each learner; second, K-means clustering for related samples, then analyzing the learning trend of similar learners. The main idea of K-means clustering method is:

Step 1: Set the number of cluster centers K;

Step 2: Random select K samples as cluster center;

Step 3: Calculate the distance from all the rest of the samples to the K cluster center, and divide the remaining samples into the categories represented by the nearest samples;

Step 4: The central point of the K classes is calculated as a new cluster center. If the new cluster center is the same as the last cluster center, it will enter step 5, otherwise, the new K cluster centers can be subdivided into steps 3;

Step 5: With the final cluster center of K clustering(s) as the final cluster center, the complete dataset can be divided into k categories.

2.3. Data Collection

Compared with the data collected online by learners, it is more difficult to collect offline data (this is because there are fewer tools to collect data and their subjectivity is evident). It is more difficult to ensure the reliability and credibility of the data. At present, there are mainly the following forms for offline data collection: questionnaire [14]; combining questionnaire with website monitoring behavior [15]; the questionnaire is combined with special data acquisition tools [16]. Because these data acquisition methods have some problems because of the incompleteness and subjectivity of the data, and there are also some problems in the process of recording the psychological cognitive process of learners. Therefore, on the basis of the questionnaire survey, this study is supplemented by the following data collection methods:

collecting data by means of live video recording and replay analysis. Through the analysis of video recording, effective learning time is manually annotated by analyzing learners' facial expressions, gestures and cognitive status. Data are collected by using live video and facial expression analysis tools. Combined with expression analysis algorithm, such tools can automatically annotate effective lectures time, collect learners data from their facial expressions, gestures, cognitive and psychological status: a combination of manual and automatic data collection. Experiments show that the comprehensive application of these methods can obtain the psychological and cognitive data of learners learning process more accurately at this stage.

3. Empirical Research

Based on the linear fitting method, taking "College English band 2" as an example, this paper makes a statistical analysis of the pre-study of 100 non-English majors of two undergraduate College English reading classes, band 2, The distribution of the scores is as follows: 30% of the learners have excellent grades (grades 86-100), 40% of the learners' have fair grades (grades 71-85) and 20% of the learners have poor grades (grades 60 -70), 10% of the learners have very poor grades (grades 0-59). A total of 10 learners learning data from different learning levels (excellent, fair, poor, and very poor) were selected and cleaned up and screened based on the principle of authenticity, completeness and standardability from the 100 learners, which is an example of the method to illustrate the experimental conclusions, the specific conclusions are as follows:

First, obtain a linear fit function. Extract knowledge unit data of the 10 sample learners of the course, using scatter diagram to analyze the sample learners learning data. The fitted line (equation) shows the student's course learning process trend. According to the calculation method of section 2.2, the formulas are fitted, such as formula (1) to formula (10), as shown in Table 2.

Second, analyze learning trend. Table 1 gives 10 of the sample students specific test data. The meaning of symbols in table 1: S1 - S10 means ten students sample number, ave means the average score of Si students in listening, speaking, reading and writing from the first to the sixth unit tests, Total means the total score of Si students.

According to the method of this paper, a typical trend chart of some of the ten students in the sample, is shown in Figure 2 to Figure 5. Among them, Figure 2 shows a sample S9, a student showing steady progress; Figure 3 shows a sample of a student S7 whose grades show a stable downward state; Figure 4 shows a sample S4 whose academic performance has improved in a volatile manner; Figure 5 shows a sample S3 in a state of waving backwards; Figure 6 shows a sample S6 with stable academic performance.

Then, we use the clustering method to get the similar learning status of the students. 5 categories of students are

shown below: student showing a sustainable progress in Figure 7; shown in Figure 8 is the category of continual backward students; Figure 9 the category of fluctuating progress; Figure 10 the category of backward fluctuating students; stable backward student category are shown in Figure 11.

Table 1. *Sample data used in the calculation.*

	Unit	Listening	Speaking	Reading	Writing	Total
	u1	16	20	20	7	63
	u2	20	17	22	5	64
	u3	17	20	23	4	64
S1	u4	18	20	23	5	66
	u5	17	17	21	12	67
	u6	18	18	23	10	69
	ave	17.67	18.67	22	7.17	65.51
	u1	24	9	24	23	80
	u2	23	10	19	25	77
	u3	21	15	15	25	76
S2	u4	20	17	14	24	75
	u5	18	17	15	23	73
	u6	12	21	13	23	69
	ave	19.67	14.83	16.67	23.83	75
	u1	18	20	21	16	75
	u2	8	16	20	18	62
	u3	8	19	13	12	52
S3	u4	16	16	18	13	63
	u5	15	11	16	16	58
	u6	12	11	12	14	49
	ave	12.83	15.5	16.67	14.83	59.83
	u1	17	14	12	16	59
	u2	5	20	17	19	61
	u3	12	18	19	14	63
S4	u4	10	20	17	13	60
	u5	6	19	14	23	62
	u6	15	19	12	21	65
	ave	10.83	18.33	14.83	17.67	61.66
	u1	22	16	0	24	62
	u2	24	16	6	20	66
	u3	22	14	12	21	69
S5	u4	25	13	5	20	63
	u5	21	18	8	21	68
	u6	23	18	10	23	74
	ave	22.83	15.83	6.83	21.5	66.99
	u1	15	9	15	19	58
	u2	11	8	17	20	56
	u3	10	10	13	18	51
S6	u4	11	5	16	18	50
	u5	7	5	14	21	47
	u6	10	12	17	17	56
	ave	10.67	8.17	15.33	18.83	53
	u1	21	21	10	20	72
	u2	21	23	10	16	70
	u3	21	11	23	12	67
S7	u4	12	23	16	14	65
	u5	23	10	15	15	63
	u6	12	19	14	17	62
	ave	18.33	17.83	14.67	15.67	66.5
	u1	19	23	24	7	73
	u2	22	19	21	7	69
	u3	25	8	24	6	73
S8	u4	25	15	19	8	67
	u5	23	16	14	8	61
	u6	20	11	11	9	51
	ave	22.33	17	18.83	7.5	65.66
S9	u1	8	18	14	17	57

	Unit	Listening	Speaking	Reading	Writing	Total
	u2	7	20	13	18	58
	u3	10	23	13	16	62
	u4	14	20	12	16	62
	u5	19	22	10	14	65
	u6	9	17	16	24	66
	ave	11.17	20	13	17.5	61.67
	u1	9	21	16	18	64
	u2	7	24	17	21	69
	u3	11	24	14	18	67
S10	u4	10	23	16	15	64
	u5	8	21	18	13	60
	u6	4	22	17	18	61
	ave	8.17	22.5	16.33	17.17	64.17

Table 2. *Learning trend results.*

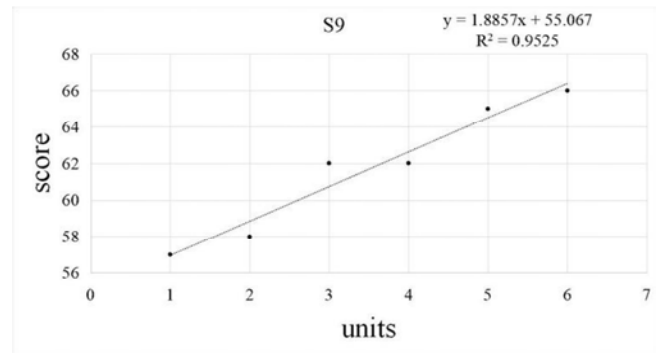| College English reading, band 2		
ID	**Learning Trend (Fitting Linear Equation)**	**Correlation R ^ 2**
S1	y = 1.1714x + 61.4	R2 = 0.9417
S2	y = -1.9429x + 81.8	R2 = 0.9437
S3	y = -3.7429x + 72.933	R2 = 0.5744
S4	y = 0.8571x + 58.667	R2 = 0.551
S5	y = 1.7143x + 61	R2 = 0.5357
S6	y = -1.06857x + 56.8	R2 = 0.2242
S7	y = -2.0857x + 73.8	R2 = 0.9823
S8	y = -4x + 79.667	R2 = 0.7836
S9	y = 1.8857x + 55.067	R2 = 0.9525
S10	y = -1.2857x + 68.667	R2 = 0.4917

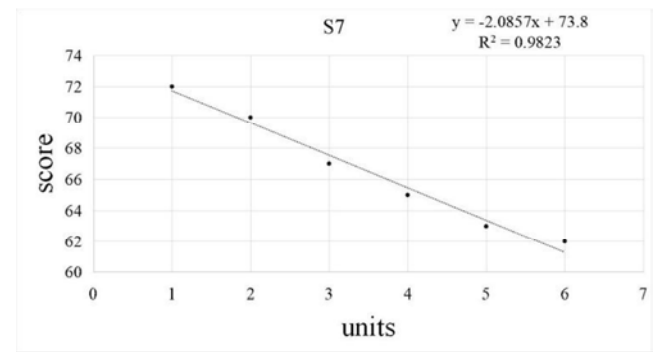**Figure 2.** *The learning trend of a stable progressive student type.*

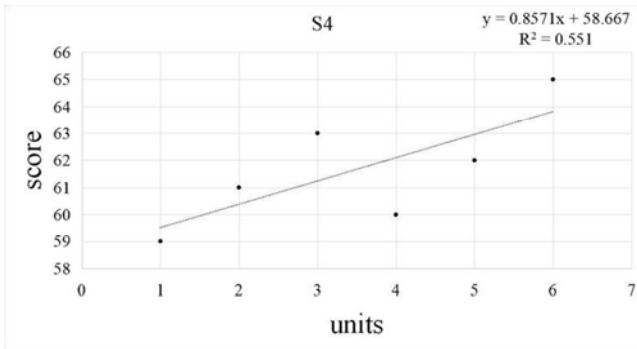**Figure 3.** *The learning trend of a stable downward student type.*

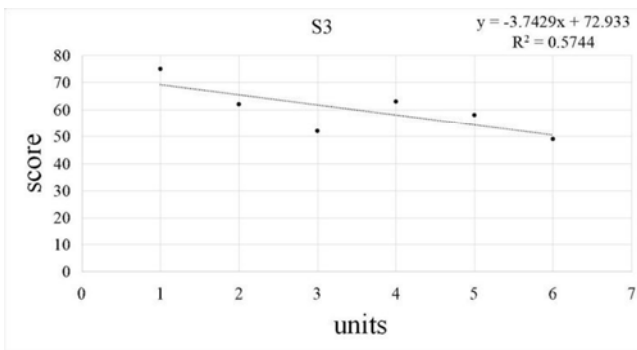**Figure 4.** *The learning trend of a student type with fluctuating progress.*

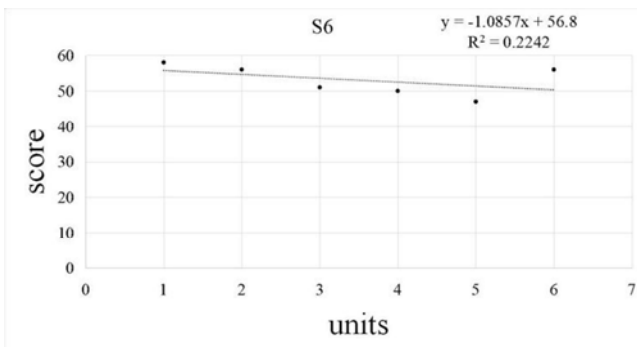**Figure 5.** *The learning trend of a student type with fluctuating downward trend.*

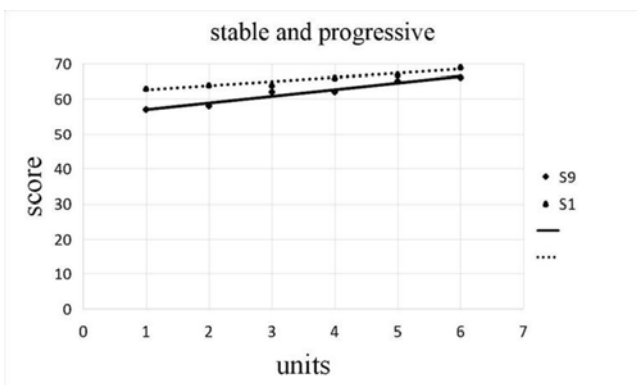**Figure 6.** *The learning trend of a stable backward student type.*

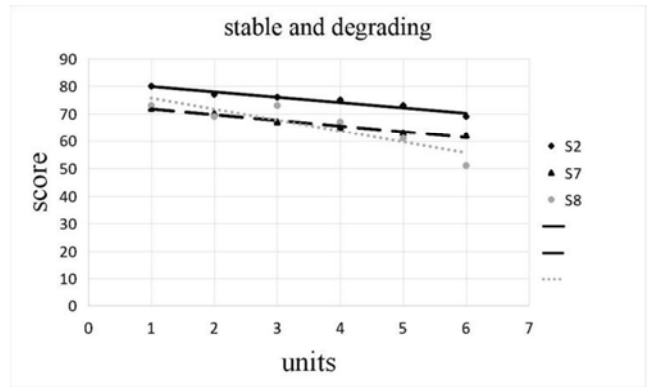**Figure 7.** *The category of students with continuous progress.*
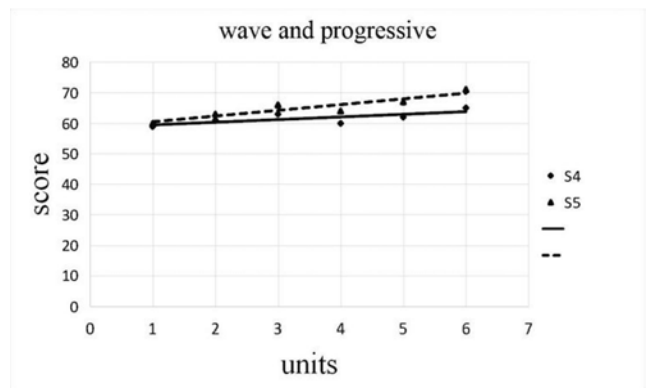
Figure 8. *A continual retrogression student type.*
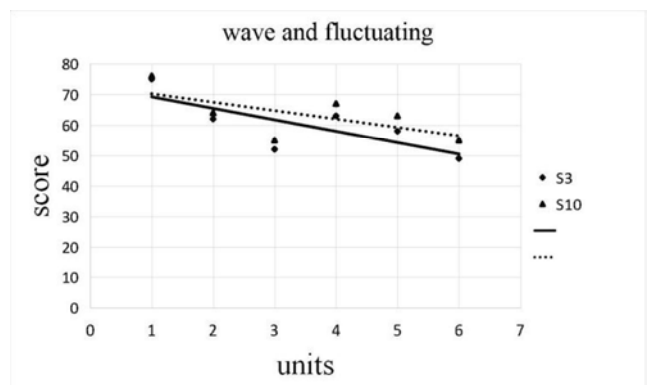
Figure 9. *The class of students with fluctuating progress.*

Figure 10. *The category of students with a fluctuating downward trend.*
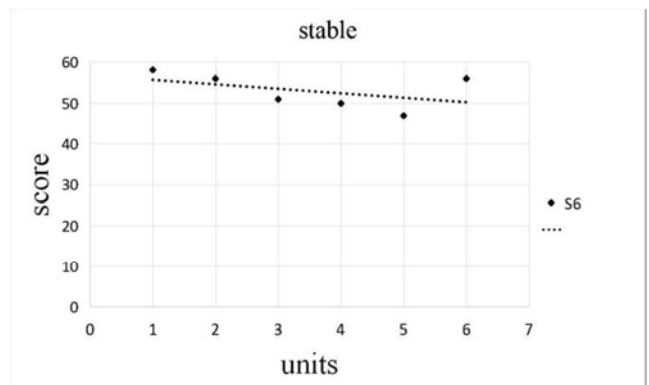
Figure 11. *A category of students with stably poor performance.*

4. Conclusion

Based on the collection and analysis of data generated in the traditional teaching process, this research analyzes and studies the learners' learning process. By analyzing the learning model of learners, instructors can detect the abnormal phenomenon of a learner's performance and issue warning to corresponding learners. Instructors can also give additional help to such learners by finding out their difficult points of the course. Experiment of this study indicates the validity and practicability of this method.

Further research could be the combination of online data analysis and offline data analysis which can describe the learner's learning trend more accurately, facilitating more accurate analysis of learners' current learning trend and prediction of learners' future trend, making possible timely intervention and failure avoidance.

Acknowledgements

References

[1] The New Media Consortium (NMC). 2011 -Horizon-Report-K12 [DB / OL].

[2] The New Media Consortium (NMC). 2015 -Horizon-Report-K12 [DB / OL].

[3] Verónica Rivera-Pelayo, Valentin Zacharias, Lars Müller, and Simone Braun. Applying Quantified Self Approaches to Support Reflective Learning [OL], LAK, 2012.

[4] Annika Wolff, Zdenek Zdrahal, Andriy Nikolov, Michal Pantucek. Improving retention: predicting at-risk students by analyzing and behavior in a virtual learning environment [OL], LAK, 2013.

[5] Richard Joseph Waddington, Sung Jin Nam. Practice Exams Make Perfect: Incorporating Course Resource Use into an Early Warning System [OL]. LAK, 2014.

[6] Adams B S, Cummins, Davis, et al. NMC Horizon Report: 2017 Higher Education Edition [J]. Journal of Open Learning, 2017.

[7] Hallinen N R, Schwartz D L. Modeling exploration strategies to predict student performance within a learning environment and beyond [C] International Learning Analytics & Knowledge Conference. ACM, 2017: 31-40.

[8] Hlosta M, Zdrahal Z, Zendulka J. Ouroboros: early identification of at-risk students without models based on legacy data [C]// International Learning Analytics & Knowledge Conference. 2017: 6-15.

[9] MO King Kee. How teachers carry out classroom instruction evaluation [J]. Curriculum, Textbook, Teaching Method, 2008 (11): 14-18.

[10] LIU Gang, Tian Jing. Several problems in the reform of classroom teaching evaluation [J]. Shanxi Normal University Press: Social Science Edition, 2012 (1): 144-147.

[11] FANG Haiguang, Wei Feng, Wang Xiaochun, Chu Yunhai. Data mining and analysis of digital classroom learning process based on PADClass model [J]. Environmental Construction and Resource Exploitation, 2014, (10): 110-120.

[12] JIN Yifu, Wu Tao, Zhang Zishi and Wang Weidong, Design and Analysis of Learning Alert System in Big Data Condition China Educational Technology 2016.2: 69-73.

[13] ZHAO Huiqiong, JIANG Qiang, ZHAO Wei, LI Yongfan and ZHAO Yan, Empirical Research of Predictive Factors and Intervention Countermeasures of Online Learning Performance on Big Data-based Learning Analytics. e-education research, 2017.1: 62-69.

[14] HE Guang-dong. Application of ID3 Algorithm in Psychological Education for College Students. Success, 2012, (20): 268.

[15] WANG Zheng, TAN Longjiang. Recommendation System Based on Customer Psychological Mining and Prediction [J]. Computer Engineering and Design, 2012, 31 (11): 4347-4350.

[16] ZHAO Xiaoyan. Psychological data mining system for higher vocational students and application [D]. Chengdu: University of Electronic Science and Technology, 2009.