Research on Students' Consumption Behavior Patterns

FangYao Xu[1], Shaojie Qu[2(✉)], and ZhiQiang Li[3]

[1] School of Mathematics and Statistics, Beijing Institute of Technology, Beijing, China
1120180084@bit.edu.cn
[2] School of Automation, Beijing Institute of Technology, Beijing, China
qushaojie@bit.edu.cn
[3] Network and Information Center, Beijing Institute of Technology, Beijing, China
lizq@bit.edu.cn

Abstract. Many studies have examined the impact of students' learning behavior on their performance, however, consumption behaviors have not been fully explored. This study uses the consumption behavior data of 3616 students over one month to analyze. First, we processed the data and proposed the characteristics that reflect students' living habits and campus behaviors, and conducted qualitative analysis. Second, we introduced multiple regression model to filter out features that had a significant impact on performance, and compared it with the SVR and random forests methods. The results showed that the linear regression model had a better quantitative characterization effect. We also found that some behavior patterns can distinguish students with excellent grades from those who lag behind. We further built a graph with significant features, and introduced graph convolutional neural networks for classification to verify above findings in educational data mining. Consequently, this study can effectively help tutors grasp students' learning status and intervene promptly when necessary.

Keywords: Consumption data · Behavior pattern · Graph neural network

1 Introduction

Among many problems in the field of educational data mining (EDM), predicting students' academic performance by using the significant amount of data generated during learning matters most [13]. In fact, one of the most important purposes of predicting students' performance is to determine whether a student will fail in a subject, so as to facilitate early interventions that can reduce the number of students who fail certain subjects [19]. Nevertheless, the question of which behavioral characteristics can distinguish outstanding students from those that lag behind has not been well resolved, and the learning tendencies of outstanding students have yet to be determined. Overall, little research has unveiled the behavior mode of outstanding students.

K. Arai (Ed.): FICC 2022, LNNS 438, pp. 44–58, 2022.
https://doi.org/10.1007/978-3-030-98012-2_5

As an important behavior mode, consumption behavior can well reflect the scope and trajectory of students' on-campus activity, their behavioral trends over a long time period. Some research has examined data on students' consumption behavior; however, few studies have related it to students' academic performance. Some scholars, such as Qian Yang and Ming Li [25], have explored the relationship between consumption behaviors and green lifestyles and the low-carbon economy. However, how to use consumption data to reflect students' activities and which behaviors are unique to or frequently displayed by excellent students needs to be studied. Thus, the issue of determining these behaviors on students' consumption behaviors remains.

In addition to determining behavior patterns, this study attempts to quantify the relationship between these behavior patterns and students' grades. Many researchers have used different methods to predict and explore. Generally speaking, machine learning algorithms are the most widely used tools in EDM. The use of machine algorithms such as Naive Bayes and decision trees to mine students' data has received significant attention [6,14]. To meet students' individual needs, the use of data mining to support distance education is also particularly important [24]. In this study, we employ multiple linear regression to quantify the relationship between students' consumption behavior and students' grades. Further, we consider constructing graph composed of these features and perform graph neural network to verify our result of linear regression.

The article is structured as follows: In Sect. 2, we further introduce current developments in EDM. In Sect. 3, we present the original data, describe the experimental method we adopted and illustrate the pseudo-code process. In Sect. 4, we describe the relationship between students' consumption behaviors and their grades both qualitatively and quantitatively, analyze the results and indicate the behavioral characteristics that can distinguish outstanding students from lagging students; further we carry out graph neural network to further illustrate the goodness of selected features. Section 5 discusses the results in more detail. Finally, we summarize the study, propose its limitations, and provide some directions for future research.

2 Related Work

Many forms of data are generated by students, for example, when they access the Internet, their test results, and their time of learning on MOOC platforms. EDM focuses on making good use of this information to predict students' performance, evaluate teachers' class quality and so on. To render the academic performance prediction more intelligent, some scholars have focused on the framework of data mining to facilitate teaching improvement [8,17]. Agaoglu [2] used students' feedback, such as course feedback questionnaires, to assess teachers' class quality. Jovanovic [12] used common machine learning algorithms, such as Naive Bayes and multilayer neural networks to predict students' academic performance. Furthermore, it was necessary to warn students who might fail examination [22]. Injada [10] focused on the graduation rate of students, while some scholars have

also studied the impact of new teaching models such as flipped classrooms on students' academic performance [1,21]. Part of the previous research on performance prediction has focused on improving the present model's accuracy [4,11]. However, although there has been much research on students' grade prediction, the behavioral characteristics that can distinguish students' academic performance have not been well elucidated.

Insufficient attention has been paid to the mining of students' consumption behaviors and their correlation with grades, not to mention the behavioral characteristics and activity trajectories of outstanding students. Prabhu [16] studied the situation of students who use mobile apps for food shopping. Zhedi Wan [23] also studied the characteristics of college students' financial behavior and the main factors that affect their investment propensity. Chang Che-Chang [5] investigated the factors that affect students' shopping consumption, such as age and monthly disposable income. Meanwhile, Auksorncherdchoo [3] attempted to develop software and systems to monitor students' attendance in class, but failed to classify their behavioral patterns. Though above research has examined students' consumption, it has not well distinguished students' behaviors and grades, thus, cannot contribute to the understanding of the behavioral characteristics of outstanding students.

Through the observation of student's behavior, we found that students with excellent performance are more likely to keep regular schedules and have more obvious learning tendencies. This inspired us to further explore the daily behaviors and activity trajectories of outstanding students, as well as those who are relatively weak in learning. Through such analysis, the relevant results can help teachers better focus on those students who are behind in learning and provide appropriate help promptly.

3 Method

3.1 Data Description

To analyze students' behavioral characteristics, we firstly took a qualitative approach. We collected data over one month from 3616 university students, as well as their final grades for some course. Detailed original consumption data are shown in Table 1:

We randomized the ID numbers of the students first; in the Table 1, the "Consumption amount" represents the amount of money a student spent on consumption items in RMB cents. The data represents the consumption of this students for the entire month of May, in which "Action" means what the student had done. The number of occurrences of an action and the order of the occurrence of certain behaviors were emphatically considered.

Table 1. Example raw data collected from students

ID number	Consumption amount (RMB cent)	Time	Action
6277	2	1/5/2018 14:35:44	Dormitory boiled water
6277	4	2/5/2018 08:26:03	Dormitory boiled water
6277	920	2/5/2018 09:30:03	Breakfast
6277	800	2/5/2018 11:35:50	Lunch
6277	150	2/5/2018 11:59:34	Lunch
6277	44	2/5/2018 16:52:19	Dinner

3.2 Data Processing and Method

We conducted experiments in three stages. The first involved the preprocessing, cleaning and feature extraction of the data. Second, scatter plots and box diagrams were used for intuitive analysis, and to draw preliminary conclusions. Finally, the regression model was applied to perform model fitting and to assess prediction effects, graph neural network was utilized to further verify and illustrate that these features are distinguishing from graph classification level as a trial.

We first de-duplicated and cleaned the data. If the two rows of data were exactly the same or the students' consumption took place in the gymnasium, management office or school bus, we did not consider this type of behavior, which has nothing to do with learning. Further, two identical behaviors that occurred within five minutes of each other were considered as the same event. We did not consider the data of students who only had a small number of consumption records and regarded that the kind of daily performance was abnormal. We then extracted features related to the students' life habits and behavior trajectories in the processed data, for example, the number of early rises, which represents that there is consumption record before 8:00 a.m.

A qualitative analysis of the relationship between students' grades and certain behaviors (e.g., the actual number of times they studied) was performed through scatter plots and box plots. When building the model, we employed the below equation where i represented the number of sample points and k represented the number of independent variables to fit the multiple linear model.

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} \cdots + \beta_k x_{ki} + \sigma_i, i \in 1, 2, \ldots, n \tag{1}$$

Where Y_i represented the grade of i_{th} student, and x_{ki} represented the i_{th} value of the k_{th} feature we extracted. Through the least squares method, for binary linear regression, only the system of equations must be solved.

$$\begin{cases} \sum Y_i = n\beta_0 + \beta_1 \sum x_{1i} + \beta_2 \sum x_{2i} \\ \sum x_{1i} Y_i = \beta_0 \sum x_{1i} + \beta_1 \sum x_{1i} + \beta_2 \sum x_{1i} x_{2i} \\ \sum x_{2i} Y_i = \beta_0 \sum x_{2i} + \beta_1 \sum x_{1i} x_{2i} + \beta_2 \sum x_{2i} \end{cases} \tag{2}$$

The independent variables were tested for significance. We also tested the properties of heteroscedasticity and autocorrelation to obtain the best model. Finally, we used machine learning algorithms to fit the model and compare prediction effects on the test set with the linear regression model. The corresponding pseudo-code is shown as Algorithm 1 below.

Algorithm 1. Data Processing and Model Fitting

Require: Original data X, X_i^1 represented i_{th} student ID,X_i^2 represented their consumption amount, X_i^3 represented time and X_i^4 represented what the student did
Ensure: Regression result
Group data by ID
while Student ID equals some ID **do**
if $X_i = X_{i+1}$ or X_i^4 =gym, management office, school-bus **then**
Delete a row
end if
if $X_i^4 = X_{i+1}^4$ and $X_{i+1}^3 - X_i^3 < 5$ **then**
Delete the second row
end if
end while
if Less than 15 consumption data records **then**
Do not consider the student's behavior data
end if
Construct features x_1,x_2,$x_3, \cdots, x_k$,
Perform intuitive qualitative analysis
while Not all features are significant, heteroscedasticity exists among $x_1, \cdots, x_k$ **do**
if Not all features are significant **then**
Delete insignificant feature x_i
Do regression
end if
if No heteroscedasticity exists **then**
Return results
end if
end while
Fit SVR model and random forest model
Do prediction on the test set and compare
Select significant variables based on multiple regression, construct graphs for graph convolutional neural network to perform classification for result verification.

4 Experiment and Result

4.1 Experiment

Based on the above method, we extracted features from two aspects to reflect the regularity of students' schedules and the sequences of their consumption behavior as Table 2 showed.

Table 2. Descriptive statistics of extracted features

Features and dependent variable (GPA)	Range	Max	Mean	Std	Var
GPA (y)	96.34	96.34	77.95	10.124	102.492
Study-actual (x_1)	144	144	6.00	13.987	195.644
Paid-days (x_2)	24	31	28.95	3.500	12.250
Getting-up (x_3)	31	31	9.75	8.189	67.060
Dinner (x_4)	31	31	15.68	8.154	66.481
Dinner-dormitory (x_5)	31	31	8.55	6.496	42.202
Classroom-lunch-or-dinner (x_6)	43	43	2.16	5.108	26.088
Breakfast-study-lunch (x_7)	27	27	0.88	2.376	5.644
Lunch-study-dinner (x_8)	19	19	0.55	1.511	2.285
Dinner-study-dorm-bathroom (x_9)	17	17	0.25	0.934	0.872
Classroom-lunch-study (x_{11})	22	22	0.31	1.292	1.669
Dinner-study-dorm-water (x_{12})	11	11	0.16	0.730	0.532
Dinner-study-dorm (x_{13})	20	20	0.41	1.341	1.799
Total-meals (x_{14})	272	272	54.82	24.708	610.487
Dor-and-super-actual (x_{15})	296	296	67.69	39.314	1545.628

We noticed that the students' average grade was only 77.95 points, and the corresponding standard deviation reached 10.124 points in Table 2. Thus, the average grade of students was relatively low, and the students' grades fluctuated greatly. Excellent students' scores were quite high, while others were likely to fail. We removed the data of students with grades close to 0, and finally filtered 3447 students' consumption data. The number of times students study should have a significant correlation with their academic performance. Inspired by this hypothesis, we drew the below scatter plots to compare the time students spent learning with their grades.

In the data we collected, different students had different records of consumption during the month. The number of days with consumption records can generally reflect the regularity of the students' lives, and we speculated that students with a regular study schedule would be more likely to achieve good results in their final exams. Figure 1 shows that when the number of times students studied in the library or classroom reached an average of at least twice a day, their academic performance fluctuated at around 80 points. When the students' average learning frequency was lower than twice a day, the scores were more densely distributed and fluctuated more sharply. Judging from the daily habits of ordinary students, it is very likely that some of these students suddenly spent more time studying in the period before the examination, and ultimately achieved a good score, which led to the sharp fluctuation.

An obvious trend in Fig. 2 was that if the number of days with consumption records in a month was greater than 20 days, the student's grades had a

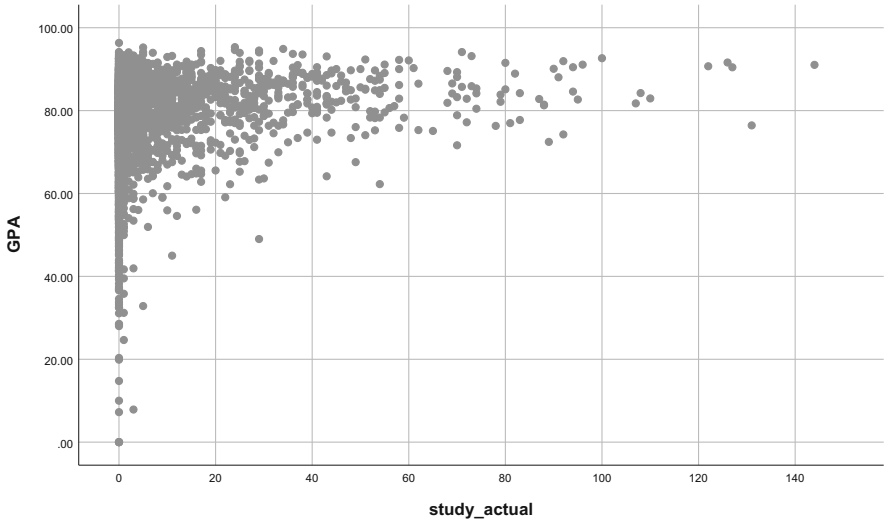

Fig. 1. Scatter plot between students' grades and the number of times students studied

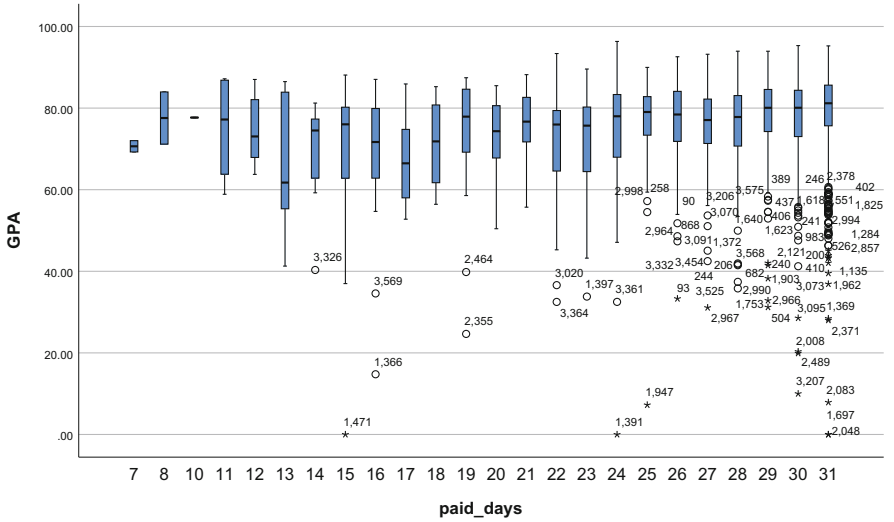

Fig. 2. Box plot between students' grades and the number of days with consumption record

slight upward trend as the number of consumption days increased. Moreover, the overall score trend was also higher than those with less than 20 days with consumption records. The data was collected over a limited time span. If this time span can include more months, this trend would be likely to be more obvious.

In general, a regular schedule was observed to have relationship with students' final exam results.

Further, we used regression models to further characterize the above findings quantitatively. Regression analysis is often used by scholars due to its simple ideas and good explanatory properties. For example, Hyun-il Lim [9] used multiple regression models to classify similar software, while Li Ting and Song Xinyuan [15] improved the classic cluster regression method and proposed a new model, that can be applied to mean regression, median regression, etc. The theoretical development of statistics is very mature, and some machine learning algorithms are established on the basis of related theories, such as logistic regression [7] and Naive Bayes [1]. Benefiting from statistical ideas, these algorithms usually have extraordinary effects on classification tasks. In our experiment, we used other machine learning algorithms to compare between models, including support vector regression (SVR) and random forest algorithms.

In multiple regression theory, the features left by the final model can indicate that at some level of significance, there indeed exists linear relationship between these features and students' performance. We hoped to further study whether these characteristics as a whole have additional effects on student performance. Inspired by the composition of graphs, we considered introducing graph convolutional neural networks for further data mining. We considered the significant variables as nodes, and constructed a complete graph by fully connecting, and also divided students' grades by 75 points. Finally, graph convolutional neural network for graph-level classification was utilized to test the overall influence of variables. In the literature survey, we found that the introduction of graph neural network models for data mining and feature analysis was relatively rare, which was an attempt in our study.

4.2 Results

We first tested the correlation between the respondents' habits and grade. The results showed that there was a significant linear relationship between them, apart from the "Dinner-study-dorm" pattern. Based on the existence of linear relationship, we performed the regression process and obtained the basic outcome show in Table 3.

From the below Table 3, the variable x_{12} was excluded from the model. In addition, under the condition of a significance level of 0.05, the probability p value of coefficients of x_6, x_7, x_9, x_{11} and x_{13} was bigger than 0.05, meanings that these variables were insignificant in combination with other features. However, performing an F test on the regression equation showed that the overall influence of the behavioral patterns on students' grades was significant. In addition, R^2 was only 0.129, which drove us to improve the model. Noting that the p value of x_{11} was much larger than that of the other variables, x_{11} was removed. Through multiple removal operations, a regression equation in which all variables as a whole had a significant effect on student performance was finally obtained. The coefficients are shown in Table 4.

Table 3. Model parameters by fitting the data and T-Test result of the corresponding coefficients

	Unstandard coefficients		Standard coefficients	t	Sig.
	B	Ste	Beta		
Constant	73.213	1.1		66.532	0
x_1	0.05	0.019	0.099	2.602	0.009
x_2	0.133	0.045	0.061	2.94	0.003
x_3	0.21	0.022	0.237	9.461	0
x_4	−0.132	0.034	−0.148	−3.921	0
x_5	0.125	0.029	0.112	4.251	0
x_6	0.094	0.065	0.068	1.453	0.146
x_7	−0.129	0.088	−0.043	−1.457	0.145
x_8	0.244	0.122	0.052	1.997	0.046
x_9	0.201	0.154	0.026	1.306	0.192
x_{10}	−0.397	0.156	−0.072	−2.544	0.011
x_{11}	0.146	0.184	0.015	0.793	0.428
x_{13}	0.023	0.012	0.076	1.956	0.051
x_{14}	−0.008	0.004	-0.041	-2.047	0.041

Table 4. Final model result after removing all variables that had no significant impact on the prediction

	Unstandard coefficients		Stc	t	Sig.
	B	Standardized error	Beta		
Constant	73.641	1.078		68.284	0
x_1	0.04	0.011	0.079	3.512	0
x_2	0.097	0.042	0.045	2.329	0.02
x_3	0.203	0.021	0.229	9.506	0
x_4	−0.107	0.032	−0.12	−3.308	0.001
x_5	0.102	0.027	0.092	3.775	0
x_8	0.276	0.105	0.059	2.633	0.008
x_{13}	0.024	0.012	0.081	2.087	0.037

From the F test of the equation, F value equals 70.366. When the significance level was 0.05, the equation had passed the test significantly, indicating that as a whole, the remaining variables had a good linear predictive effect on students' grades. To ensure that the model could be used, we also performed a heteroscedasticity test using the grade correlation coefficient. In reality, the residuals that were ultimately obtained showed that the model had eliminated the heteroscedasticity. At the same time, all variables were significant to the

regression equation, and $R^2 = 0.505$; therefore, we determined the final regression equation as follows:

$$Y = 73.641 + 0.04x_1 + 0.097x_2 + 0.203x_3 - 0.107x_4 + 0.102x_5 + 0.276x_8 + 0.024x_{13} \quad (3)$$

Here, we chose random forest regression and SVR to compare with the linear regression model. For the data set, we used 90% of the data for training and 10% of the data for testing. The same processing was also performed in the linear regression; the comparison result is shown in the Table 5.

Table 5. Comparison between multiple linear regression model and part of machine learning algorithms

Model	Mean squared error	Mean absolute error	Parameters	R^2 score
Linear regression	5.0777	42.4911		0.505
SVR [18]	4.9893	42.7715	C = 0.6, kernel = rbf	0.3209
Random forest [20]	4.9027	37.6370	estimators = 300, criterion = mse, max features = 3, max depth = 3	0.4264

The square error of the SVR was 4.9893, while the absolute error was 42.7715; on the test set, the effect of SVR was good, but the R^2 was relatively small, which shows that the fitting effect of the SVR was poor on our dataset. The same analysis can be done on the performance of random forest algorithm, and we guessed that the two algorithms was sensitive to data. The multiple linear regression model had relatively large errors on the test set, while the fitting effect was better. In summary, the effect of linear regression was better than that of SVR and the random forests. Further, we constructed complete graph based on the multiple regression, performed graph classification, and selected 700 students as the test set. The training results were as follows:

From the final performance on the test set, the best classification accuracy of the graph convolutional neural network can reach 77.143%; due to the uneven distribution of labels, the weighted precision was 64.977%, the weighted recall score reached 69.42% and the weighted f1 score achieved 66.824%. From Fig. 3, the reduction process was relatively smooth and stable. Considering from the perspective of graph classification, these selected significant variables could separate outstanding students from lagging students. We noticed that these constructed graph contained relatively few nodes and edges, and the information contained was also limited. This would also be the direction that graph neural networks could be improved in educational data mining in the future and more research was needed.

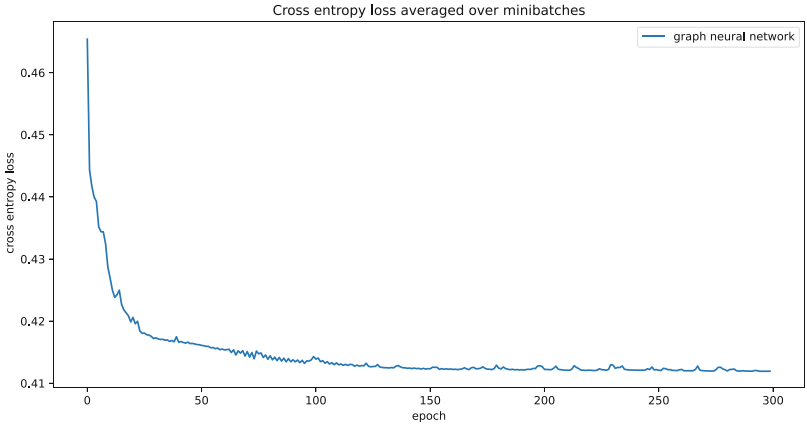

Fig. 3. Cross entropy loss reduction process of graph neural network in 300 epochs

4.3 Analysis

For linear regression models, fixing other variables, when the actual number of learning (x_1) increases by one unit, the corresponding learning score will increase by 0.04 points, which was consistent with the actual situation. The longer students spends studying, the better they master the knowledge, and the easier it is for them to achieve good grades in their exams. For the number of dinners (x_4), the coefficient was negative in the multiple linear regression. In fact, when we processed the collected data, we considered two identical actions that occurred within five minutes of one another as one occurrence. Even so, there still exists five or six dinner record during data collection, and to a certain extent, this may also have affected the coefficients in the model, making the regression model demonstrate that this behavior had a negative effect on performance prediction. As for the rest variables, a similar interpretation of the coefficients can be made. Taking R^2 into consideration, although it only reached 0.505, for the real data and the behavioral data extracted from real data to fit the model, it was difficult for R^2 to achieve above 0.8 in most cases, especially in economic filed. Therefore, the result was acceptable to some degree.

Though the model's performances of SVR and random forest algorithm were not as good as the multiple linear regression, the mean square error and mean absolute error were smaller than in the linear regression. In our experiment, the students' grade deviation predicted by machine learning was relatively less than that of linear regression model and performed better. The above results also encourage the consideration of how to use theories related to regression analysis for data processing and preliminary exploration.

The results in above Fig. 4 revealed that, when the number of times students got up early increased, their grades also showed an upward trend. Thus, it can be inferred that when the number of students' early morning rises exceeded 20 times in the month, their grade scores fluctuated at around 85 points. Further,

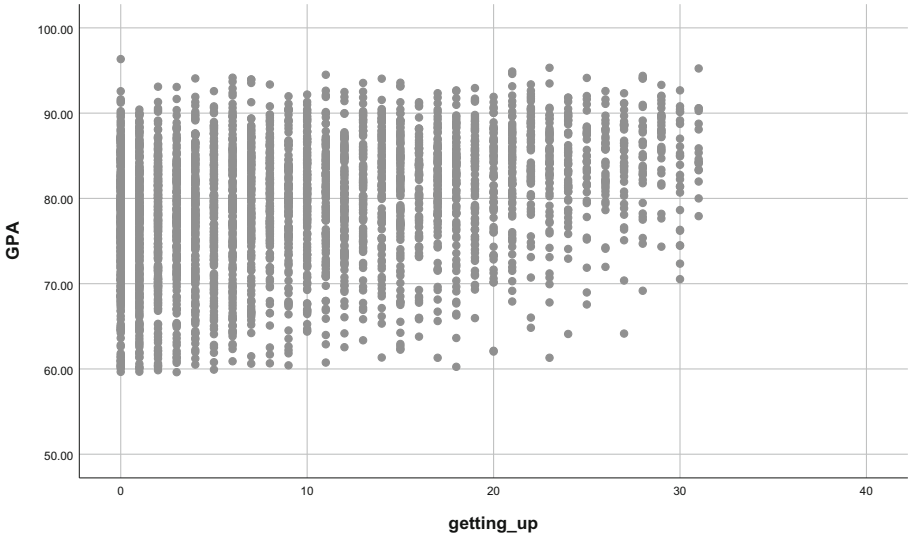

Fig. 4. The number of rising early is related to students' grades and could separate excellent student from those who lag behind

when the number of early rises was between 10 and 20 times, the scores fluctuated around 80 points. As the number of early rises increases, the fluctuation variance also tended to become smaller.

Besides, due to the limitations of regression model, we introduced the topology structure, and used the graph convolutional neural network to verify regression result. In fact, due to the relatively small number of features, the construction of the graph structure is relatively simple, which inspires more research on the introduction of graph neural networks into the field of data mining. The results of graph classification could also reflect that these significant variables are distinguishable, which provided tutors with useful information about the learning situation of students.

5 Discussion

Excellent students usually kept relatively regular schedules, which was mainly reflected in the number of days with consumption records. In addition, the number of early rises may reflect students' degree of diligence; the more diligent the students, the higher their final grades will be. At the same time, the regression equation and many significance tests showed that the patterns of "Lunch-study-dinner" and "Dinner-study-dorm" had significant effects on students' grades. This also showed that the regular pattern of outstanding students is to study after lunch and then eat dinner, then to study after dinner and return to the dormitory thereafter. The analysis of data from students who were lagging behind showed that they were more inclined to stay in the dormitory. We also used the

regression method to quantitatively describe the specific impact of significant features on students' grades. The above findings are in line with reality. Both the qualitative and quantitative results will provide university staff with more information about outstanding students.

In addition, our experiment conformed a method to deal with data preprocessing, that is, we can firstly use a variety of statistical methods to test the properties of correlation and heteroscedasticity. If necessary, data transformation can be applied in the original data, such as box-cox transformation and logarithmic transformation. After fully clarifying the information contained in the extracted features and the data used, an appropriate model can be selected to perform regression or classification tasks. In addition, graph neural networks are also developing rapidly, but there is a lack of research in data mining, which is also a direction that can be studied in depth in the future. Here, we provide one possible research direction, which refers to building larger graph and improving the internal structure of the graph, such as the existence of edges.

6 Conclusion

The existing research has represented findings on students' consumption behaviors that reflect the trajectory of their activities. In the qualitative analysis, we found a relationship between the distribution of students' grades, the number of times they study, and the number of times they get up early. Through quantitative analysis, we established an equation describing the relationship between students' consumption behaviors and their grades. The results show that if students often get up early and spend enough time studying in the afternoon and evening, they generally achieve better grade, which is interesting. However, our experiment also has shortcomings. Due to the significance test, some features could not be included in the model, which may have led to a waste of information. We then used the graph convolutional neural network to classify and confirm that these variables as a whole are indeed distinguishable for separating outstanding students and those who lag behind. Therefore, using more powerful tools to comprehensively explore the more complex relationships between various behaviors is particularly important. More research is needed to study the complex relationships between students' behaviors, and to conform the behavioral trajectories that affect student grades. We need to put some attention in applying graph neural network for local information mining, and more research will be needed.

References

1. Aditya, M.A., Helen, A., Suryana, I.: Naive Bayes and maximum entropy comparison for translated novel's genre classification, UK, vol. 1722, p. 012007, 9 pp. (2021)
2. Agaoglu, M.: Predicting instructor performance using data mining techniques in higher education. IEEE Access **4**, 2379–2387 (2016)

3. Auksorncherdchoo, S.: Software development for student behavior tracking system, San Diego, CA, United States, pp. 55–59 (2018)
4. Bonde, S.N., Kirange, D.K.: Educational data mining survey for predicting student's academic performance. In: Pandian, A., Senjyu, T., Islam, S., Wang, H. (eds.) ICCBI 2018. LNDECT, vol. 31, pp. 293–302. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-24643-3_35
5. Chang, C.-C., Hung, Y.-Y., Wang, Y.-C.: Partial least squares structural equation modeling in online shopping: the moderator effect between impulsive buying tendency and behavior. WSEAS Trans. Bus. Econ. **17**, 542–550 (2020)
6. Dewantoro, G., Ardisa, N.: A decision support system for undergraduate students admissions using educational data mining, Piscataway, NJ, USA, pp. 105–109 (2020)
7. Febrianti, R., Widyaningsih, Y., Soemartojo, S.: The parameter estimation of logistic regression with maximum likelihood method and score function modification, UK, vol. 1725, pp. 012014, 7 pp. (2021)
8. Goncalves, A.F.D., Maciel, A.M.A., Rodrigues, R.L.: Development of a data mining education framework for data visualization in distance learning environments, Pittsburgh, PA, United States, vol. 0, pp. 547–550 (2017)
9. Lim, H.: Interlacing data to classify software in linear regression approach. In: Park, J.J., Loia, V., Pan, Y., Sung, Y. (eds.) Advanced Multimedia and Ubiquitous Engineering. LNEE, vol. 716, pp. 25–31. Springer, Singapore (2021). https://doi.org/10.1007/978-981-15-9309-3_4
10. Injadat, M.N., Moubayed, A., Nassif, A.B., Shami, A.: Multi-split optimized bagging ensemble model selection for multi-class educational data mining. Appl. Intell. **50**(12), 4506–4528 (2020). https://doi.org/10.1007/s10489-020-01776-3
11. Johnson, W.G.: Data mining and machine learning in education with focus in undergraduate CS student success, Espoo, Finland, pp. 270–271 (2018)
12. Jovanovic, M., Vukicevic, M., Milovanovic, M., Minovic, M.: Using data mining on student behavior and cognitive style data for improving e-learning systems: a case study. Int. J. Comput. Intell. Syst. **5**(3), 597–610 (2012)
13. Karthikeyan, V.G., Thangaraj, P., Karthik, S.: Towards developing hybrid educational data mining model (HEDM) for efficient and accurate student performance evaluation. Soft Comput. **24**(24), 18477–18487 (2020). https://doi.org/10.1007/s00500-020-05075-4
14. Kumar, R., Kumar, M., Joshi, U.: Data mining-based student's performance evaluator. In: Choudhury, S., Mishra, R., Mishra, R., Kumar, A. (eds.) Intelligent Communication, Control and Devices. AISC, vol. 989, pp. 719–726. Springer, Singapore (2020). https://doi.org/10.1007/978-981-13-8618-3_73
15. Li, T., Song, X., Zhang, Y., Zhu, H., Zhu, Z.: Clusterwise functional linear regression models. Comput. Stat. Data Anal. **158**, 107192 (2021)
16. Prabhu, N., Soodan, V.: The effect of mobile app design features on student buying behavior for online food ordering and delivery. In: Stephanidis, C., et al. (eds.) HCII 2020. LNCS, vol. 12427, pp. 614–623. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-60152-2_45
17. Prada, M.A., et al.: Educational data mining for tutoring support in higher education: a web-based tool case study in engineering degrees. IEEE Access **8**, 212818–212836 (2020)
18. Smola, A.J., Schölkopf, B.: A tutorial on support vector regression. Stat. Comput. **14**(3), 199–222 (2004)
19. Tasnim, N., Paul, M.K., Sattar, A.H.M.S.: Identification of drop out students using educational data mining, Piscataway, NJ, USA, p. 5 (2019)

20. Teng, Z., Chu, L., Chen, K., He, G., Fu, Y., Li, L.: Hardware implementation of random forest algorithm based on classification and regression tree, Chongqing, China, pp. 1422–1427 (2020)
21. Urquiza-Fuentes, J.: Increasing students' responsibility and learning outcomes using partial flipped classroom in a language processors course. IEEE Access **8**, 211211–23 (2020)
22. Viloria, A., et al.: Data mining applied in school dropout prediction. J. Phys. Conf. Ser. **1432**(1) (2020)
23. Wan, Z.: Investigation on college students' financial management behavior and research on guiding strategies, France, vol. 233, p. 01167, 5 pp. (2021)
24. Xu, Y., Zhang, M., Gao, Z.: The construction of distance education personalized learning platform based on educational data mining. In: Abawajy, J., Choo, K.K., Islam, R., Xu, Z., Atiquzzaman, M. (eds.) ATCI 2019. AISC, vol. 1017, pp. 1076–1085. Springer, Cham. (2020). https://doi.org/10.1007/978-3-030-25128-4_134
25. Yang, Q., Li, M.: Research on college students' garment consumption behavior and low-carbon lifestyle, UK, vol. 1790, pp. 012093, 5 pp. (2021)