Student Performance Prediction Based on Blended Learning

Zhuojia Xu, Hua Yuan, and Qishan Liu

Abstract*—*Contribution: **This article explored blended learning by implementing a student-centered teaching method based on the flipped classroom and small private online course (SPOC). The impact of general online learning behavior on student performance was analyzed. This work is practical and provides enlightenment for learning analysis and individualized teaching in blended learning.**

Background: **Providing individualized teaching in a large class is an effective way to improve teaching quality, but the traditional teaching method makes it difficult for teachers to learn about each student's learning situation. Blended learning offers the possibility of individualized teaching for teachers. The combination of flipped classroom and SPOC is a good way to implement blended learning, but few studies have verified the predictability of learning performance in such a scenario to explore individualized teaching.**

Intended Outcomes: **Students' behavior in blended learning can be used to predict their learning outcomes, and the implementation method is reproducible. Teachers can implement individualized teaching in blended learning.**

Application Design: **The learning activities were designed and reconstructed to create a blended learning scenario, data that depict students' learning behavior were collected and used to predict their performance by a multiple regression model. Student performance was measured by the final offline exam, and its predictability in the 1/4, 1/2, and 3/4 semester was tested for early intervention.**

Findings: **The results show that students' online behavior can be predictors of their performance, and with the advance of the course, the predicted results are more stable and reliable.**

***Index Terms*—Blended learning, flipped classroom, individualized teaching, massive open online courses (MOOCs), small private online courses (SPOCs), student assessment, student performance.**

I. Introduction

IN HIGHER education, traditional face-to-face teaching is still the mainstream [1]. It is teacher centered that teachers are the providers of knowledge and students are the recipients of knowledge. As technology advances, the chalk in the teacher's hand has been gradually replaced by laser pointer with remote control, but the student's role of the audience did not change essentially. The traditional teaching method is increasingly difficult to adapt to the development of the new era, and its quality is hard to be guaranteed. College students of this generation are digital natives who are growing up in the ever-changing digital and Internet age. It is not easy to adapt them to traditional teaching. First, the Internet has become easier for students to access information and knowledge [2]. The sources of knowledge are diverse, and teachers are no longer the only imparter of knowledge. Once students are not satisfied with the lecture, the content, the way of teaching, or even the accent of the mother tongue in the lecture can become the cause of students not listening, sleeping in class, and skipping classes. Second, mobile phones have become the basic equipment for college students [2]. The excessive use of cellphone makes students become classroom phubbers, affecting their attention and learning in the class [3]. Traditional teaching needs to change with the interactions between technology and learning. Besides, large-class education is universal in some countries such as China [4]. Traditional large-class education may have difficulty in providing individualized teaching and could even have negative effects on student academic performance [5]–[7].

The emergence of the massive open online course (MOOC) seems to bring a wave of reform to traditional education. MOOC attracts students at diverse levels of knowledge and abilities, and it covers a wide range of knowledge and reduces the deepness of knowledge to set a lower threshold for learning [8]. Whereas, low completion, difficulty in mutual recognition of credits, low social recognition, and dishonesty [9] in MOOC have prompted college educators to introspect and explore constantly. The small private online course (SPOC) sets off an educational revolution of classroom teaching. SPOC is aimed at small-class teaching and is more suitable for further imparting professional knowledge [8]. There is some practice of SPOC that has achieved good results, such as Copyright [10] and The Architectural Imaginary [11] set up by Harvard University and Circuit Principle of Tsinghua University [12]. In recent years, SPOC has also been popular in blended learning [13]–[15].

Blended learning based on SPOC provides an opportunity for teachers to explore personalized teaching, but the practice has also exposed some problems. First, most teachers use questionnaires to learn about students' learning behavior and

Manuscript received March 13, 2019; revised November 12, 2019, January 20, 2020, and May 3, 2020; accepted July 4, 2020. Date of publication August 7, 2020; date of current version February 3, 2021. This work was supported in part by the Graduate Education Innovation Program of Guangdong Province under Project 2015JGXM-ZD04 and in part by the Guangdong Higher Education Teaching Reform Project "Research on Constructivism and Its Application in the Teaching of Computer Networks." *(Corresponding author: Zhuojia Xu.)*

Zhuojia Xu and Hua Yuan are with the Communication and Computer Network Lab of Guangdong, School of Computer Science and Engineering, South China University of Technology, Guangzhou 510641, China (e-mail: xu.zhuojia@mail.scut.edu.cn; hyuan@scut.edu.cn).

Qishan Liu is with the Theoretical Physics Department, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany (e-mail: qisliu@students.uni-mainz.de).

Digital Object Identifier 10.1109/TE.2020.3008751

their attitude toward the curriculums [16], which is time consuming and not time for teaching feedback. Is there a more direct and automated way to help teachers learn about students' learning and their performance for timely intervention? Second, in the learning process, students' learning behaviors, such as study time and assignment grades, are recorded and showed as diagrams on the SPOC platforms. What is the relationship between these data and student performance? Can this information be used to predict their performance to help individualized teaching?

To solve the above issue, this article created a blended course *computer networks* based on SPOC and flipped classroom to explore the possibility of using online learning behavior to predict student performance in blended learning. Considering the blend of online and offline learning, student performance was measured by the offline final exam and predicted by online learning data, which were generated on the SPOC platform. Multiple linear regression was used to analyze the impact of online behavior on student performance and the possibility of early prediction. It is noteworthy that student performance is not necessarily predictable in this case, as offline learning behavior is not taken into account and online learning data just involve partial learning activities in blended learning. If student performance can be predicted only through online learning behavior, teachers can save time and effort in collecting offline data and make full use of educational data mining to assist personalized teaching in blended learning.

There are previous works on student performance prediction in blended learning [17]–[19], but most of them discussed this issue without teaching context. As the types of blended learning vary with the dimensions of the blend [20] and the prediction is data driven [21], it is not that practical to predict student performance without a specific context. Besides, online learning data in such research were mostly collected through private and hidden learning logs [17]–[19], [22]–[25], which are unlikely to be obtained by teachers, bringing more difficulties to practice. Compared to related work, the contributions of this article are as follows.

1) This article explores the possibility of using online learning behavior to predict student performance and provide personalized teaching, providing an example for the design of blended learning and the application of educational data mining.
2) The online behavioral data in this article were general and accessible for teachers, making the research more practical.
3) Linear and nonlinear models were compared to further discuss the generalization of the predictions.

II. Related Work

The prediction of student performance has been a hotspot in online learning. Many studies are devoted to finding online learning behavior that can be used to predict student performance. In [22], six types of features were extracted from click-stream logs within an MOOC and used to predict students' grades of next assessments. Brinton *et al.* [23] developed frameworks to extract event-based and position-based sequences from student video-watching clickstreams in MOOC. Their experiments demonstrated that video-watching behavior can help improve student performance prediction. Due to declining participation over time in MOOC [26], Jiang *et al.* [27] utilized students' assessment performance and social data in week 1 to predict students' certificate obtaining.

In addition to MOOC, there is research on the prediction of student performance in SPOC. The work [28] developed a linear regression model and a deep learning model to predict student performance in SPOC, and it stated that the model can be generalized for MOOC or other online learning. Wan *et al.* [29] used logical regression to predict the weekly test pass of students in SPOC. Marcos *et al.* [25] used features related to platform visiting and interactions with videos and quizzes to predict whether a student can pass the admission test, providing enlightenment for student performance prediction in SPOC-based blended learning. Overall, there is a lack of work concerning studies on student performance prediction in SPOC-based blended learning.

Some studies predicted student performance in blended learning by using data directly from the learning management system (LMS). Raga and Raga [17] developed a deep neural network model for early prediction of student performance in blended learning. Kim *et al.* [18] developed linear and nonlinear prediction models based on the pedagogical types of blended learning. Figueira [19] used the number of online course access, coverage of digitally provided learning material, and the differences in student sequences with their golden standard to predict the online score.

In the literature, student performance is often quantified by final grade [17]–[19], [24], [30], [31], course engagement [32], certificate obtaining [33], or student dropouts [34], [35]. Linear models, such as multiple linear regression [18], [24], [30], [31] and nonlinear models, such as gradient boosting decision tree (GBDT) [36], *K*-nearest neighbors (KNNs) [37], [38], and decision tree (DT) [39]–[41] have been widely used in the prediction of student performance. Multiple linear regression is popular due to its simplicity and convenience in analyzing the influence of multiple factors. Students' interactions with the course, such as video-watching clickstreams, visiting boards, login frequency, the number of posts, and behavior in online quizzes, are generally considered as predictors. Whereas, most of the studies tend to use hidden and more detailed learning logs [17]–[19], [22]–[24], which may be hard to obtain due to privacy protection. Their proposals may be difficult to apply widely in practice. Second, some studies tend to pay attention to the prediction accuracy of student performance while ignoring the teaching context [17]–[19]. As argued by [21], the predictors of student performance are data driven and the results are closely related to the teaching context.

Based on related work, this article mainly used multiple linear regression to analyze the predictability of student performance and verified the generality of the prediction through typical nonlinear models, which are GBDT, DT, and KNN. Students' online learning data were more general and

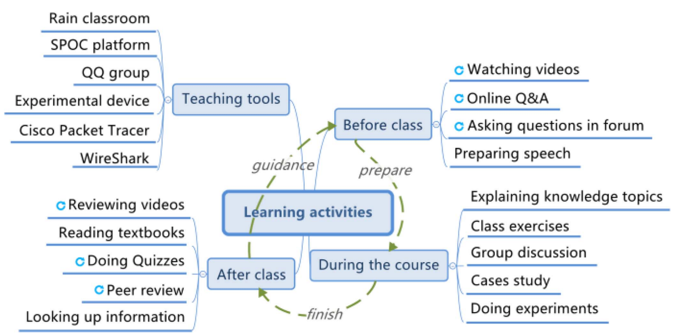

Fig. 1. Blended learning activities of computer networks.

accessible and were analyzed within the teaching context, making the research more practical.

III. Teaching Context and Data Collection

Understanding the teaching context is essential to obtain *a priori* knowledge for the analysis. Therefore, this section will introduce the design and implementation of blended learning and data collection to help understand the teaching context.

A. Design of Blended Learning

This article adapted MOOC resources, SPOC, flipped classroom, and student-centered teaching to blend self-paced learning, live and collaborative learning, which is the second dimension of the blend as introduced by Singh [20]. To realize the deep alignment of online and offline learning, a variety of online and offline activities were designed, as shown in Fig. 1. These activities include self-paced learning, such as watching videos, doing quizzes, and collaborative learning, such as peer review and group discussion. The activities with the "○" symbol were online learning activities and conducted on the SPOC platform. The other activities were mainly offline activities and carried out through the flipped classroom.

Before the class, students mainly learned on the SPOC platform. They can preview the course in advance by watching videos, answering the Q and A in the videos, or discussing with classmates in the forum. They have the freedom to choose when and where to explore the online course.

In-class activities were mainly driven by the flipped classroom. The teacher taught key and difficult knowledge points according to students' learning situations and used a built-in mobile application, called "rain classroom" tools to carry out random roll calls, case studies. Students were put into the center. They can discuss peer to peer with teachers and peers, do exercises, or send on-screen comments through rain classroom. Online and offline learning were seamlessly connected and blended through classroom teaching.

After-class learning activities were designed to help students digest what they had learned. Multiple-choice questions (MCQs) and peer-review assignments corresponding to each unit of teaching content were deployed on the SPOC platform. Questions and options of MCQ were randomly generated by the SPOC platform. Students were given two chances to complete MCQ assignments in each unit. They can decide to answer twice or only once. Peer review helps students understand the solutions of their peers, and comprehend and consolidate the knowledge once again. For these after-class assignments, students can selectively review the videos, read textbooks, or consult other relevant materials according to their situation.

B. Implementation and Data Collection

The author launched the undergraduate-level on-demand MOOC *computer networks* on CNMOOC[1] in the autumn of 2017 and the spring of 2018, and implemented blended learning according to the above methods. The course contains eight units of teaching content. Videos and the courseware were published by the teacher on the SPOC platform in advance. The spring class and the fall class had 55 and 72 students, respectively. Students in these two classes were in the same grade and exposed to blended learning for the first time. These two classes were taught by the same teacher.

To test the reliability of online learning data, the requirements for the spring class and the fall class were a little different. The time for doing MCQ assignments was limited in the spring class, but not in the fall class, and the online learning requirements of the fall class were fewer than that of the spring class. In the empirical study and experiments in Section IV, the reliability of the data will be verified.

During the whole semester, the SPOC platform recorded online learning data for the teacher. In general, it is difficult for teachers to get hidden and detailed behavioral data, such as clickstreams, time spent in each activity directly from third-party platforms. For the practicability of the study, the author only considered some general behaviors shown in Table I, and learning activities out of SPOC were not considered as predictors. Note that there were multiple units of assignments in the course, each student would have multiple assignment-related records. At the end of the semester, the students were required to take the final exam offline and their final grades were recorded and used to measure their performance.

IV. Data Analysis of Blended Learning

Section IV examined whether the common online behavior in Table I can be used to predict student performance. A linear analysis was used to explore the relationship between online behavior and student performance. The predictability of student performance in the 1/4, 1/2, 3/4, and the whole semester was tested by multiple linear regression. Finally, the goodness of fit of the linear model and nonlinear models was compared.

The goodness of fit of the linear model was often measured by R^2 (1) and adjusted R^2 (2). In the following formulas, y represents the true value, $\hat{y}$ is the predicted value, n is the number of samples, and p is the number of predictors:

$$R^2 = 1 - \frac{\sum \left(y - \hat{y}\right)^2}{\sum \left(y - \bar{y}\right)^2} \tag{1}$$

$$\text{Adjusted } R^2 = 1 - \frac{\left(1 - R^2\right)(n-1)}{n-p-1}. \tag{2}$$

[1]https://www.cnmooc.org/portal/course/2990/9620.mooc

TABLE I
MAIN STUDENT BEHAVIORAL DATA SHEET GENERATED DURING THE BLENDED LEARNING PROCESS

Features			Description	Remarks	Range
General Features	1	The time of the first access	The time of students enter the course at the first time		⩾0
	2	Study time	The time of studying from videos and other materials	When students are studying on the platform, the time of non-platform studying is not counted (measured in minutes)	⩾0
	3	The number of posts	The number of posts and replies from students in the SPOC forum		⩾0
Assignment Features	4	MCQ grades	The scores of MCQ	For each unit, students have two opportunities to submit MCQ	[0, 100]
	5	The used time for MCQ	The time it takes to answer MCQ	In the Spring class, there is time limitation for MCQ (measured in minutes)	⩾0
	6	Submission time	Submission time of MCQ		⩾0
	7	The number of submissions	The number of chances that students take to do a MCQ assignment		[0, 2]
	8	Grades of subjective questions	The scores of answering subjective questions	The scores of subjective questions are from peer review. Students are asked to score the subjective questions of at least three peers.	[0, 100]

A. Data Preprocessing

The original data of student behavior on the SPOC platform must be preprocessed before analysis. Some features in Table I were preprocessed in the following way to produce more descriptive characteristics.

1) Time of the First Access: The time of first access describes how soon students enter the course, which may be linked to the enthusiasm for learning. The interval between the time when SPOC was launched and the student's first access was calculated and defined as the time delay of study.

2) Submission Time: The submission time describes whether the student procrastinates on the assignment, which may be linked to the student's learning attitude and initiative. The time interval between the submission time of MCQ and its deadline was calculated and defined as the time delay of MCQ.

3) Number of Submissions and Scores of MCQ: Each student had two opportunities to submit an MCQ assignment. Their final MCQ score was the higher one of the two submissions. For students who did not submit their assignments, their number of submissions was recorded as 0, regardless of whether he/she used the second opportunity to make up for the assignment.

After the preprocess above, feature 1 in Table I becomes the time delay of study and feature 6 becomes the submission delay.

As mentioned in Section III, each student had multiple assignment-related records. In order to facilitate the prediction of the model, the average values of these quantities were calculated, as follows:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3}$$

where x_i can be one of the assignment features: the score, the number of submissions, the submission delay, the used time for MCQ, etc. The variable n represents the number of assignments for all units. For example, when you survey a student's MCQ scores for a semester, which has eight online tests in total, then you will get $n = 8$. Formula (3) represents the student's average score for a semester. Once a student misses an MCQ assignment, his/her final score will reduce, so the average score can give a penalty to the absence of submitting assignments.

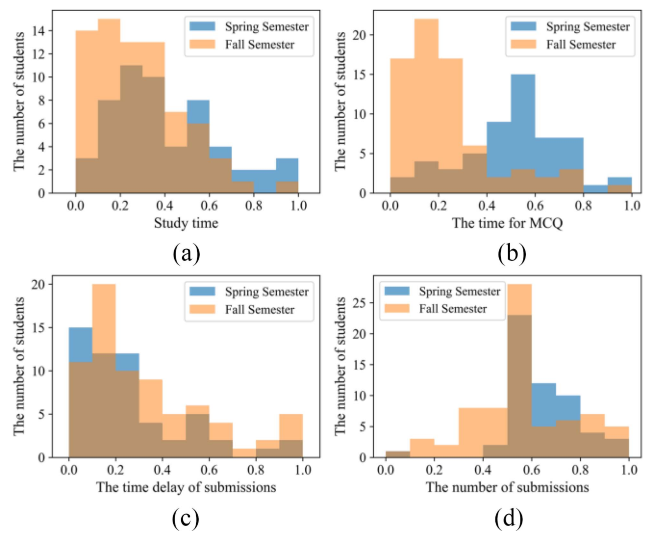

Fig. 2. Distribution of (a) study time, (b) time for MCQ, (c) submission delay, and (d) number of submissions in two semesters.

The range of each feature is shown in Table I. In the last step, each eigenvalue in Table I was normalized to the interval [0–1] by the min–max normalization to keep the same fundamental unit.

B. Empirical Study

The Matplotlib module [42] was used to visualize data. In Fig. 2, the distribution of study time, the used time for MCQ, submission delay, and the number of submissions in the spring and fall classes are shown.

In Fig. 2(a), the distribution of spring class in study time is more uniform than that of the fall class. The study time of the fall class is shorter than that of the spring class, which is consistent with the fact that there were fewer requirements for fall class to study online than that for the spring class. In

TABLE II
PEARSON COEFFICIENT WITH P-VALUE AND VIF BETWEEN THE VARIABLES AND THE OFFLINE FINAL GRADE

	Spring class						Fall class					
features	r	p of r	VIF	coefficient	std err	p of coeff	r	p of r	VIF	coefficient	std err	p of coeff
Study time	.495	<0.001	1.414	0.3593	0.120	0.005	.256	0.033	1.764	\	\	\
The number of posts	.442	<0.001	1.407	0.2536	0.147	0.091	.328	<0.01	1.161	0.0827	0.037	0.027
Time delay of study	-0.258	0.0624	2.859	\	\	\	-0.386	0.001	1.455	\	\	\
MCQ grade	.300	<0.05	3.822	-0.3705	0.244	0.136	.679	<0.001	2.191	0.2196	0.067	0.002
Used time for MCQ	.406	<0.01	1.321	0.2916	0.129	0.029	.249	<0.05	1.715	\	\	\
Submission delay	.497	<0.001	1.580	0.4198	0.129	0.002	.387	<0.001	1.732	\	\	\
The number of submissions	.284	<0.05	1.955	\	\	\	.342	<0.01	1.391	\	\	\
Grade of peer review	.437	<0.01	2.526	0.3823	0.207	0.072	.672	<0.001	2.331	0.2672	0.084	0.002
Constant	\	\	\	0.1504	0.141	0.290	\	\	\	0.4654	0.048	0.000
R^2: 0.579 Adjusted R^2: 0.524							R^2: 0.567 Adjusted R^2: 0.547					

* r: Person coefficient
* std err: the standard error of the coefficient.
* coefficient, coeff: linear model coefficient
* p values: Generally, p < 0.05 means a statistical difference, p < 0.01 means a significant statistical difference, p < 0.001 means an extremely significant statistical difference.

Fig. 2(b), the used time for MCQ in the spring class is positively skewed, and the used time of most students is shorter, while in the fall class, the distribution is more like the normal distribution, and the total answer time is longer than that in the spring class because there was no limitation of the answer time in the fall class. The distribution of submission delays between the two classes is similar, and some students like to submit assignments before the deadline. In Fig. 2(d), the distribution of the spring class is relatively narrow, and the number of submissions is concentrated in 0.5–0.8 (note that this is the normalized value, so it is less than 1), while the variance of the fall class distribution is more extensive.

Overall, the spring class seems to be more active online than the fall class, and the results are consistent with the empirical cognition. The online learning requirements of the spring class were more than those of the fall class.

C. Prediction of Student Performance

The linear analysis was used to discover the correlation between student behavior and performance. Python module Statsmodels [43] and SciPy [44] were used for the analysis of learning behavior.

1) Correlation Analysis: The Pearson coefficient describes the correlation between dependent variables and independent variables. In Table II, the Pearson coefficient with p-value was calculated.

Pearson coefficients between the time delay of study and the final grades of the two classes were -0.258 and -0.386, respectively. This feature describes how soon students enter the course, and the result indicates that the longer the delay, the worse his/her academic performance is likely to be. It can be observed that the submission delay had a positive effect on student performance in the two classes, indicating that the earlier the homework is submitted, the better student academic performance may be. Except for the time delay of study, other variables were positively correlated with the final grade, from weak correlation [0.2–0.4), medium correlation [0.4, 0.6), to strong correlation [0.6, 0.8). The correlation between the MCQ grade and the final grade was 0.3 in the spring class but was 0.679 in the fall class. The MCQ grade is more correlated with student performance in the fall class. In addition to the grade of MCQ and peer review, the correlation between most of the online behavior and the final grade in the fall class is weaker than that of the spring class, which was in line with the empirical cognition that their online learning requirements were fewer than those for the spring semester and assignment-related activities became their major activities.

Through correlation analysis, teachers can understand which behaviors are correlated with student performance and screen learning behaviors for the prediction of student performance.

2) Prediction: The variance inflation factor (VIF) quantifies the severity of multicollinearity. A variable with a VIF larger than 5 indicates that the variable is collinear with other variables. The VIF was calculated to detect multiple collinear variables before using multiple linear regression.

In Table II, the VIF of each variable was lower than 5, indicating no severe multicollinearity between the variables. Since there was no multicollinearity between the variables and there existed correlations between the variables and student performance, all features were considered in the prediction. Multiple linear regression with forward selection used ordinary least squares (OLSs) to learn the coefficient of each variable, whose significance was tested by student t-test. When fitting the model, forward selection would select the variable whose inclusion improves the fit significantly. The coefficients,

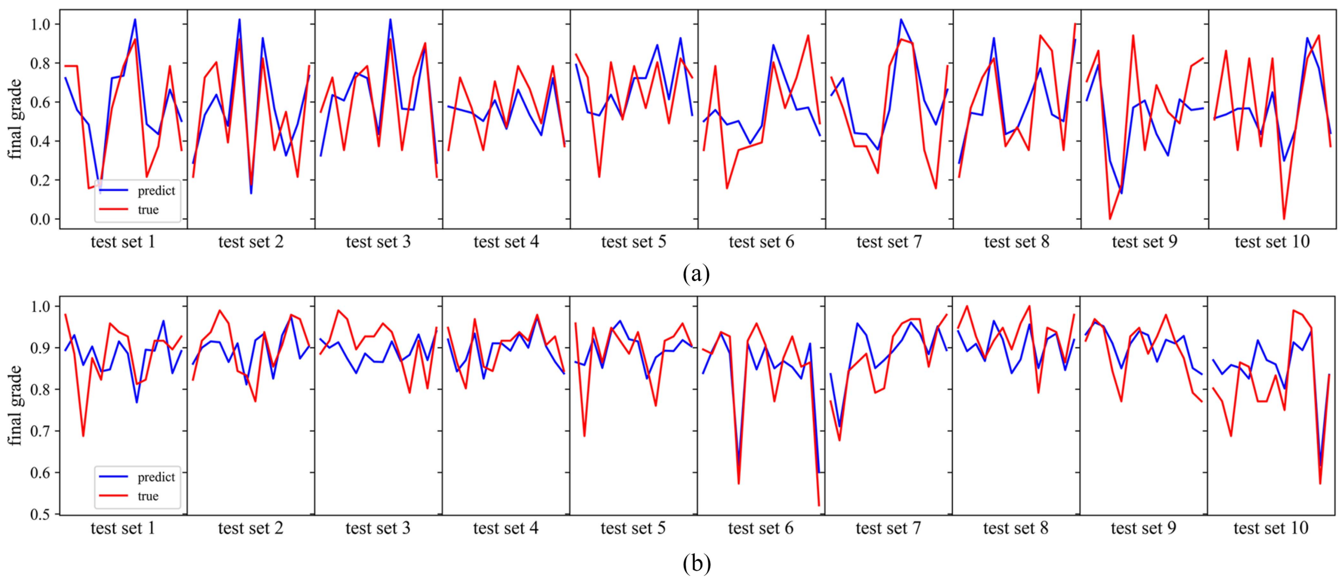

Fig. 3. Predicted results and ground-truth values in the (a) spring class and (b) fall class.

standard errors, and *p*-values of the variables in the model are shown in Table II.

The R^2 and adjusted R^2 of the spring class were 0.579 and 0.524, respectively, and were 0.567 and 0.547 in the fall class, respectively. Compared with the previous prediction models that had a reasonable amount of explained variance ranging from 0.22 to 0.52 [30], the results in this experiment show that online learning behavior is predictive for student performance.

It can be seen from the "coefficient" and "*p* of coeff" columns in Table II that the study time, the number of posts, MCQ grade, the used time for MCQ, the submission delay, and the grade of peer review are predictors in the spring class, but the coefficient of the MCQ grade was not significant, indicating that the other features are sufficient enough to predict the final grade. The higher the grade of peer review, the longer the study time, and the sooner the assignments can be submitted, the better their final academic performance may be. In the fall class, the number of posts, the MCQ grade, and the grade of peer review become predictors. Although only three features played a role in the prediction model of the fall class, the prediction effect was not weak. Assignment-related features such as the MCQ grade, the grade of peer review seems to be important predictors as they appeared in the prediction models of both classes.

What interesting is in the spring class, the MCQ grade had a positive correlation with student performance, but its coefficient was negative. This does not mean the correlation between the MCQ grade and student performance has changed from positive to negative. Correlation analysis is like single factor analysis, as it can provide point-to-point relationships for teachers to understand which behaviors are positive and which are negative. But when different behaviors were combined, they may influence each other and have different effects on student performance. The prediction allows teachers to understand the comprehensive effect of the combination and the performance of each student.

Fig. 3 shows the predicted grade and the true grade intuitively. The horizontal axis identifies ten tests, each of which randomly selects 20% of the original data as the predicted samples. The vertical axis represents the normalized value of the final grade. As shown in Fig. 3, most of the predicted grades are close to the true values, which means student performance can be well predicted by multiple linear regression.

In the experiment, the eight students with the lowest grades and the eight students with the highest grades were selected to further illustrate the fitting results of the model. In the spring class, predicted grades and true grades of the worst eight were both lower than 0.5 (note that the score was normalized between 0 and 1). The last and the fourth of the worst students got the predicted grades closest to their true grades. The true grade of the worst student was 0.1 while his/her predicted grade was 0.0, and the true grade of the fourth-worst student was 0.17 while his/her predicted grade was 0.176. Top eight in the spring class had high predicted grades, but the gaps between the predicted and true values were larger than those of the worst students. Overall, the prediction model of the spring class had a good prediction for the worst students, which means it can better discover at-risk students. In the fall class, the predicted grades of the worst and the top eight were both close to their true grades, with the largest gap of 0.162 and a minimum of 0.004. Students with final grades no more than 0.5 also had predicted grades that no more than 0.5. The prediction model of the fall class had a good prediction for both poor and good students.

The experimental results show that online learning behavior can predict student performance. Teachers can learn about students' learning situations and discover at-risk students through the prediction. If the prediction can be realized earlier, it may be more helpful for personalized guidance.

D. Prediction in the Learning Process

To verify the predictability of student performance in the early stage, the behavioral data generated in the 1/4 semester, midsemester, and 3/4 semesters were used to predict the final exam grade. It is essential to state that in the 1/4 semester, only

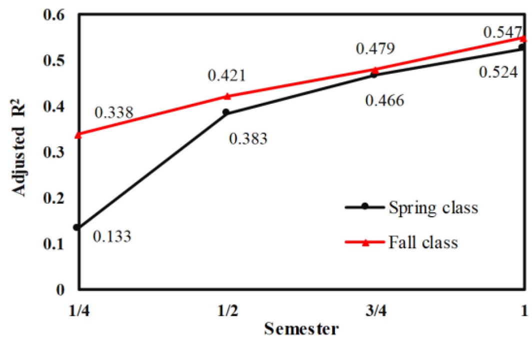

Fig. 4. Adjusted R^2 in the 1/4, 1/2, 3/4 semesters, and the whole semester.

part of the records was ready for analysis. The MCQ assignments in the first and second units as well as the grades of the first peer review. In the midsemester, the records of MCQ assignments from unit one to unit four and the grades of twice peer review were used. In 3/4 semesters, MCQ assignments from unit one to unit six and the grades of all peer review were accessible. In this process, the records of the study time and the numbers of posts were not accessible. The number of features may be different in the learning stage. Since the adjusted R^2 takes the number of features into account, the experiment only examines this indicator.

Similarly, the forward stepwise regression was run to select predictors and produce prediction formulas. For comparison, the adjusted R^2 of multiple linear regressions in the 1/4 semester, midsemester, 3/4 semester, and the whole semester are plotted in Fig. 4. In the 1/4 semester, the adjusted R^2 of the spring class was only 0.133, and the fall class was 0.338, which was higher. In the midsemester, the adjusted R^2 of the spring class was increased by about 0.25 compared with its previous result, whereas, the fall class slightly increased by about 0.08. As the course advanced, the adjusted R^2 of the two classes rose to 0.466 and 0.479, respectively. Though only part of the data were used in the 3/4 semester, its results were close to the whole semester.

The result shows that student performance can be predicted at the early stage, but the earlier the stage, the more unstable the prediction results are. In the middle stage, student performance can be preliminarily predicted to help teachers adjust teaching as soon as possible, and intervention can be performed to achieve personalized guidance and teach students in accordance with their aptitude.

E. Comparing With Nonlinear Models

Based on the related work, this article uses GBDT, DT, and KNN to verify the generalization of the prediction. R^2 is a common indicator to exam the goodness of fit of nonlinear prediction models, and therefore, the experiment only compared R^2 of these nonlinear models with the multiple linear regression. The nonlinear models were implemented through the machine learning tool [41] and fitted with default parameters.

In the spring class, R^2 of GBDT reached 0.999, DT reached 1.00, and for KNN with the number of neighbors set to 5, R^2 was 0.540. In the fall class, R^2 of GBDT reached 0.994, DT reached 1.00, and for KNN with the number of neighbors set to 5, R^2 was 0.499. R^2 of GBDT and DT was high, and R^2 of KNN was close to that of linear regression. GBDT and DT are more robust, they can perfectly fit the nonlinear relationship between learning behavior and student performance, but KNN cannot fit such nonlinear relationship as well as GBDT and DT. The performance of KNN may be affected by the number of neighbors [45]. These results verify the generalization of the prediction. Although the performance of the nonlinear model is better than that of linear regression, the interpretability is poor as it utilizes nonlinear combinations of behaviors and their nonlinear relationships with student performance.

V. Discussion

This article implemented blended learning in two different semesters to explore the predictability of student performance and the possibility of early intervention.

The experiment found that online learning data involving part of the learning activities of blended learning can be used to predict student performance, and assignment-related features were potentially important predictors, which means teachers can save time and effort in collecting offline data. They can use general online learning behavior, especially assignment-related behavior to learn about students' learning situations.

Student performance can also be predicted at the early stage, but the earlier the stage, the more unstable the prediction results are. The impact of learning behavior on student performance is different from that in the correlation analysis, which means the combination of learning behaviors has different effects on student performance. Such effects are complex as they may be related to the learning attitude, learning method, learning process, accidental factors, etc. The correlation analysis provides candidate factors for prediction and allows teachers to learn about point-to-point relationships between behavior and performance. The prediction of student performance allows teachers to know who would have poor performance and who would have good performance so that teachers have the opportunity to learn about students' situations in advance and provide personalized intervention and guidance in offline face-to-face teaching.

VI. Future Work

There is still some work to explore in the future. The experiments can be developed to discover abnormal students so as to provide more abundant information for intervention. In terms of early intervention, the relationship between intervention time and predictive stability can be further explored through a large number of experiments. It would be possible to recommend intervention time intelligently based on the relationship.

It is believed that with the publicity of educational information, more and more teachers can collect educational data and participate in the research, making more contributions to improve the quality of higher education.

REFERENCES

[1] R. Zeng, L.-R. Xiang, J. Zeng, and C. Zuo, "Applying team-based learning of diagnostics for undergraduate students: Assessing teaching effectiveness by a randomized controlled trial study," *Adv. Med. Educ. Pract.*, vol. 8, pp. 211–218, Mar. 2017.

[2] M. Anshari, M. N. Almunawar, M. Shahrill, D. K. Wicaksono, and M. Huda, "Smartphones usage in the classrooms: Learning aid or interference?" *Educ. Inf. Technol.*, vol. 22, no. 6, pp. 3063–3079, 2017.

[3] J. S. Mendoza, B. C. Pody, S. Lee, M. Kim, and I. M. McDonough, "The effect of cellphones on attention and learning: The influences of time, distraction, and nomophobia," *Comput. Hum. Behav.*, vol. 86, pp. 52–60, Sep. 2018.

[4] J. Yang and L.-M. Yao, "The investigation analysis on college class size," *J. Higher Educ.*, vol. 7, pp. 86–93, Jul. 2012.

[5] M. Dillon and E. C. Kokkelenberg, "The effects of class size on student achievement in higher education: Applying an earnings function," in *Proc. 42nd Annu. AIR Forum*, Tallahassee, FL, USA, 2002.

[6] J. Monks and R. M. Schmidt, "The impact of class size on outcomes in higher education," *BE J. Econ. Anal. Policy*, vol. 11, no. 1, p. 62, 2011.

[7] B. N. Matta, J. M. Guzman, S. K. Stockly, and B. Widner, "Class size effects on student performance in a Hispanic-serving institution," *Rev. Black Polit. Econ.*, vol. 42, no. 4, pp. 443–457, 2015.

[8] P. Guo, "MOOC and SPOC, which one is better," *Eurasia J. Math. Sci. Technol. Educ.*, vol. 13, no. 8, pp. 5961–5967, 2017.

[9] D. Gao, "A coolheaded response to Hot MOOCs: Reflections on the six problems of MOOCs," *J. Distance Educ.*, vol. 2, pp. 39–47, Feb. 2014.

[10] C.-M. Liao, "Research on the SPOC online course of higher education," in *DEStech Transactions on Economics, Business and Management*, ICEM, Vélizy-Villacoublay, France, 2019.

[11] R. Mendoza-Gonzalez, "User-centered design strategies for massive open online courses (MOOCs)," *Int. J. e-Collaboration*, vol. 13, no. 1, pp. 188–190, 2017.

[12] Y. Kang, "An analysis on SPOC: Post-MOOC era of online education," *Tsinghua J. Educ.*, vol. 35, no. 1, pp. 86–93, 2014.

[13] S. Ziebarth and H. U. Hoppe, "Moodle4SPOC: A resource-intensive blended learning course," in *Proc. Eur. Conf. Technol. Enhanced Learn.*, 2014, pp. 359–372.

[14] M. Kris, "Integrating a small private online course (SPOC) component in an undergraduate engineering management subject as a blended learning approach: A case in Hong Kong," in *Proc. 4th Int. Conf. Distance Educ. Learn.*, 2019, pp. 16–19.

[15] F. J. D. Cepeda, *Small Private Online Research: A Proposal for a Numerical Methods Course Based on Technology Use and Blended Learning*, Int. Assoc. Develop. Inf. Soc., Lisbon, Portugal, 2017.

[16] H. BydŽovská, "A comparative analysis of techniques for predicting student performance," in *Proc. 9th Int. Conf. Educ. Data Min.*, 2016, pp. 306–311.

[17] R. Raga and J. Raga, "Early prediction of student performance in blended learning courses using deep neural networks," in *Proc. IEEE Int. Symp. Educ. Technol. (ISET)*, Hradec Kralove, Czechia, 2019, pp. 39–43.

[18] J. H. Kim, S. Seodaemun-gu, Y. Park, J. Song, and I.-H. Jo, "Predicting students' learning performance by using online behavior patterns in blended learning environments: Comparison of two cases on linear and non-linear model," *Korea*, vol. 120, p. 750, 2014.

[19] A. Figueira, "Predicting grades by principal component analysis: A data mining approach to learning analyics," in *Proc. 16th Int. Conf. Adv. Learn. Technol. (ICALT)*, Austin, TX, USA, 2016, pp. 465–467.

[20] H. Singh, "Building effective blended learning programs," in *Educational Technology*, vol. 43. Englewood Cliffs NJ, USA: Saddle Brook, 2003, pp. 51–54.

[21] D. Gaševiæ, S. Dawson, T. Rogers, and D. Gasevic, "Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success," *Internet Higher Educ.*, vol. 28, pp. 68–84, Jan. 2016.

[22] Z. Ren, H. Rangwala, and A. Johri, "Predicting performance on MOOC assessments using multi-regression models," 2016. [Online]. Available: arXiv:1605.02269.

[23] C. G. Brinton, S. Buccapatnam, M. Chiang, and H. V. Poor, "Mining MOOC clickstreams: Video-watching behavior vs. in-video quiz performance," *IEEE Trans. Signal Process.*, vol. 64, no. 14, pp. 3677–3692, Jul. 2016.

[24] O. H. Lu, A. Y. Huang, J. C. Huang, A. J. Lin, H. Ogata, and S. J. Yang, "Applying learning analytics for the early prediction of students' academic performance in blended learning," *J. Educ. Technol. Soc.*, vol. 21, no. 2, pp. 220–232, 2018.

[25] P. M. Moreno-Marcos *et al.*, "Predicting admission test success using SPOC interactions," in *Proc. 9th Int. Conf. Learn. Anal. Knowl.*, Tempe, AZ, USA, 2019, pp. 4–8.

[26] P. Hill, *Emerging Student Patterns in MOOCs: A Graphical View*, WordPress, e-Literate, Oct. 2013. [Online]. Available: https://eliterate.us/emerging_student_patterns_in_moocs_graphical_view/

[27] S. Jiang, A. Williams, K. Schenke, M. Warschauer, amd D. O'Dowd, "Predicting MOOC performance with week 1 behavior," in *Proc. 7th Int. Conf. Educ. Data Mining*, London, U.K., Jul. 2014, pp. 273–275.

[28] C. Yu *et al.*, "SPOC-MFLP: A multi-feature learning prediction model for SPOC students using machine learning," *J. Appl. Sci. Eng.*, vol. 21, no. 11, pp. 279–290, 2018. [Online]. Available: https://doi.org/10.1787/f8d7880d-en

[29] H. Wan, J. Ding, X. Gao, Q. Yu, and K. Liu, "Predicting performance in a small private online course," in *Proc. Educ. Data Mining*, 2017, pp. 384–385.

[30] R. Conijn, C. Snijders, A. Kleingeld, and U. Matzat, "Predicting student performance from LMS data: A comparison of 17 blended courses using moodle LMS," *IEEE Trans. Learn. Technol.*, vol. 10, no. 1, pp. 17–29, Jan.–Mar. 2017.

[31] R. Conijn, A. Van den Beemt, and P. Cuijpers, "Predicting student performance in a blended MOOC," *J. Comput. Assist. Learn.*, vol. 34, pp. 615–628, Oct. 2018.

[32] A. S. Lan, C. G. Brinton, T.-Y. Yang, and M. Chiang, "Behavior-based latent variable model for learner engagement," in *Proc. Int. Conf. Educ. Data Min.*, 2017, pp. 64–71.

[33] J. Zhuoxuan, Z. Yan, and L. Xiaoming, "Learning behavior analysis and prediction based on MOOC data," *J. Comput. Res. Develop.*, vol. 52, no. 3, pp. 614–628, 2015.

[34] S. Ameri, M. J. Fard, R. B. Chinnam, and C. K. Reddy, "Survival analysis based framework for early prediction of student dropouts," in *Proc. 25th ACM Int. Conf. Inf. Knowl. Manag.*, 2016, pp. 903–912.

[35] W. Li, M. Gao, H. Li, Q. Xiong, J. Wen, and Z. Wu, "Dropout prediction in MOOCs using behavior features and multi-view semi-supervised learning," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Vancouver, BC, Canada, 2016, pp. 3130–3137.

[36] H. Wan, Q. Yu, J. Ding, and K. Liu, "Students' behavior analysis under the Sakai LMS," in *Proc. IEEE 6th Int. Conf. Teach. Assess. Learn. Eng. (TALE)*, Hong Kong, China, 2017, pp. 250–255.

[37] H. Al-Shehri *et al.*, "Student performance prediction using support vector machine and k-nearest neighbor," in *Proc. IEEE 30th Can. Conf. Elect. Comput. Eng. (CCECE)*, Windsor, ON, Canada, 2017, pp. 1–4.

[38] W. Punlumjeak and N. Rachburee, "A comparative study of feature selection techniques for classify student performance," in *Proc. 7th Int. Conf. Inf. Technol. Elect. Eng. (ICITEE)*, Chiang Mai, Thailand, 2015, pp. 425–429.

[39] Q. A. Al-Radaideh, E. M. Al-Shawakfa, and M. I. Al-Najjar, "Mining student data using decision trees," in *Proc. Int. Arab Conf. Inf. Technol. (ACIT'2006)*, Irbid, Jordan, 2006.

[40] R. Kabra and R. Bichkar, "Performance prediction of engineering students using decision trees," *Int. J. Comput. Appl.*, vol. 36, no. 11, pp. 8–12, 2011.

[41] M. Pandey and V. K. Sharma, "A decision tree algorithm pertaining to the student performance analysis and prediction," *Int. J. Comput. Appl.*, vol. 61, no. 13, pp. 1–5, 2013.

[42] *Matplotlib*. Accessed: Feb. 13, 2019. [Online]. Available: https://matplotlib.org/

[43] *Statsmodels*. Accessed: Feb. 13, 2019. [Online]. Available: https://www.statsmodels.org/stable/index.html

[44] *SciPy*. Accessed: Sep. 30, 2018. [Online]. Available: https://docs.scipy.org/doc/scipy/reference/stats.html

[45] A. Ng. (2017). *Machine Learning Yearning*. [Online]. Available: http://www.mylearning.org/