Analysis and Prediction of Satisfaction Index of Online Learning

Jian Wang¹, Yanmei Chai¹, Wei Zhang¹, Yuanyuan Zhang¹

¹School of information, Central university of finance and economics, Beijing, China
Wanderingful@126.com

Abstract. In the process of learning with the help of online platform, students' learning satisfaction is an important factor that constitutes the effect of online teaching. On the basis of the existing research, this study proposes three factors, namely, cognitive level, technological environment and social environment, to constitute online learning satisfaction. At the same time, through statistical analysis, the key factors affecting learning satisfaction are sorted out, the selection model of identifying the key factors affecting learning satisfaction is established, and online learning satisfaction is fitted by prediction. The study shows that self-efficacy evaluation, teaching support evaluation, platform use evaluation and student-teacher interaction evaluation have a great impact on students' learning satisfaction in the process of online learning. At the same time, the 93% prediction accuracy of the prediction model based on the above-mentioned research can be achieved. In addition, this paper also puts forward suggestions on how to strengthen and improve learning satisfaction.

1. Introduction
In the context of online teaching, the students’ satisfaction is affected by the teaching environment as well as traditional teaching. In addition, we also need to focus on students’ own physical and mental experience and self-evaluation, which are the main factors that can affect the student-led learning process, including the ability of acquisition, processing and feedback of learning information, as well as students' ability to control their own emotions. After these factors are carried out index analysis, we can judge the degree of students' sense of acquisition in the learning process and predict their satisfaction with online learning[1~4]. In particular, judging from the perspective of students can help better capture the individual needs of students, understand the relationship between the elements depicted by students' learning behavior, and excavate the key elements affecting students' final judgment, which has a strong guiding significance for improving online teaching quality and defining the direction of improvement[5~6].

Educational Data Mining (EDM) and Learning Analytics (LA) [7]are commonly used teaching analytics in recent years. They utilize various data generated in the process of education and teaching, such as the data in learning management system and adaptive learning system, and are based on different analytics to establish data models, to interpret and analyze students’ learning behaviors, and explore and summarize students' learning rules. In this paper, through learning analysis, we will establish a model of students' satisfaction index, and predict their online learning satisfaction.

Online learning satisfaction is the overall response of college students to the experience of online learning. Through the activities and experiences, students produce different cognitive experiences and make different emotional reactions. Therefore, it is a complex concept that integrates multiple learning processes and is influenced by multiple features. Its determinants involve many aspects. The evaluation
index of learning can be quantified. In this study, the feature elements based on LM neural network are used for modeling. By establishing the index system of students' online learning satisfaction and prediction, we’ll provide a basis for better personalized education in the era of intelligent education.

2. Research Background
The research mainly carries out data analysis around the subject of learning satisfaction through questionnaire survey. The object is the students who took online course of Computer Application Basis in our school. As a school-wide computer public basic course, Computer Application Basis plays an important role in training students' computational Thinking and improving students' digital literacy. Online teaching creates a platform for students to learn independently and provides space for students' innovative spirit.

In this study, all the students who took online course of Computer Application Basis in 2016 were taken as the object. There were 300 participants, at the end of the study, the questionnaire was provided to the students to answer online. A total of 297 valid questionnaires were collected.

From the content of survey, there are three main aspects. First of all, we paid attention to students’ self-characteristic elements, such as gender, grade, computer proficiency and whether they have participated in online learning, etc. In the main part of the questionnaire, the index is designed according to three dimensions[8]: cognitive level, technological environment and social environment. Cognitive level includes two secondary factors: self-efficacy evaluation and expectation achievement evaluation; technological environment includes two factors: teaching support evaluation and platform use evaluation; social environment mainly focuses on interaction, and three secondary factors are designed: curriculum interaction evaluation, student-teacher interaction evaluation and group interaction evaluation. Under each secondary index, several third-level factors are given as support. The third part of questionnaire is the evaluation of students' overall satisfaction. The second & third parts are shown in the table.

In order to facilitate data statistics and regression analysis, this study adopted five-point option scoring method of Likert scale. In the second and third parts, the topic options are set exactly the same, a score of 5 representing that the evaluator agrees with the question very much (very good, very satisfied, etc.); a score of 4 representing agreement (good, satisfied, etc.); a score of 3 representing it’s ok (ok); a score of 2 representing disagreement (not very good, not very satisfied, etc.); and a score of 1 representing very disagreement (very bad, very dissatisfied, etc.).

3. Predictive Algorithm

3.1. Adaptive-Lasso method
Through index analysis, we can see that the impact of different factors on the final results is different, and even some indexes will cause deviation. Therefore, before using neural network to model, this study first uses adaptive-Lasso variable selection method to analyze the factors affecting the satisfaction of online learning[9].

Adaptive-Lasso method is evolved on the basis of ordinary Lasso method. Lasso method uses regularization method in logistic regression, which satisfies simultaneous parameter estimation and variable selection. The definition of its parameter estimation is shown in Formula 1:

$$\hat{\beta}\text{(lasso)} = \arg\min_{\beta} \left\| y - \sum_{j=1}^{p} x_j \beta_j \right\|^2 + \lambda \sum_{j=1}^{p} \left| \beta_j \right|$$

In which, $\lambda$ is a non-negative regular parameter, $\lambda \sum_{j=1}^{p} \left| \beta_j \right|$ called a penalty term.

Lasso method can effectively solve the shortcomings of least squares method and stepwise regression local optimal estimation, but the premise is that its own needs meet certain strict conditions. An
An improved Lasso method is proposed, which adds different weights to different coefficients. It is called Adaptive-Lasso method. The definition is shown in Formula 2:

$$
\hat{\beta}^{(n)} = \arg\min_{\beta} \left\{ y - \sum_{j=1}^{p} x_j \beta_j \right\}^2 + \lambda_n \sum_{j=1}^{p} \hat{w}_j |\beta_j| 
$$

In which, weight $\hat{w}_j = \sqrt{\frac{1}{|\beta_j|}} \ (\gamma > 0), \ j = 1, 2, \cdots, p$ , $\hat{\beta}_j$ is a coefficient obtained by the ordinary least squares method.

In this study, taking students' satisfaction with online teaching as dependent variable and the selected index as independent variable, we can construct a prediction model between behavior indicators and learning outcomes.

### Table 1. Satisfaction Index of Online Learning

| Secondary index | Three-level index |
|-----------------|-------------------|
| Self-efficacy evaluation (x1) | Online learning improves my self-study ability. |
| | In the process of online learning, I can actively try different ways to solve problems. |
| | Online learning enables me to master learning time flexibly. |
| | In the process, I can keep a high level of enthusiasm and interest in learning. |
| Expectation achievement evaluation (x2) | I can arrange my study in a planned way with the help of online learning platform. |
| | The learning results of the course are in line with my expectations. |
| | I look forward to the online learning model for future courses. |
| Teaching support evaluation (x3) | The content of the course (such as courseware, teaching materials etc.) can meet my learning needs. |
| | Teaching AIDS (such as homework reviews, learning manuals, etc.) can help me learn better. |
| | Course assignments are moderately difficult and can help me learn step by step. |
| | The design of curriculum system is reasonable, which can meet my practical application needs. |
| Platform use evaluation (x4) | Access and operation of online learning platform are convenient. |
| | Online courseware can be viewed smoothly and course columns can be accessed easily. |
| | Functional modules of online courses can meet the needs of daily learning. |
| Curriculum interaction evaluation (x5) | My total length of online learning (browsing teaching resources) meets the requirements. |
| | In the process of learning, I frequently log on to the teaching platform for learning. |
| Student-teacher interaction evaluation (x6) | Teachers can answer questions in time when learning encounters problems. |
| | Submitted assignments can receive timely feedback from teachers. |
| | Thematic counseling and answering questions can help me better carry out my study. |
| Group interaction evaluation (x7) | I would like to communicate with other students on learning problems. |
| | I often discuss problems with other students. |
| | I think online teaching platform provides a good channel for communication. |
| Overall satisfaction (y) | Your overall satisfaction evaluation of the course |

### 3.2 LM neural network

LM neural network is an improvement based on the traditional BP algorithm. LM algorithm belongs to the optimization algorithm and is the most widely used non-linear least squares algorithm. In the process of searching, it has the advantages of gradient method and Newton method, so that it can’t only follow the direction of negative gradient in each iteration process, but also combine them effectively, which
greatly speeds up the convergence speed and allows errors to search along the upward direction, so that the target is not easy to fall into the local optimal solution [10].

LM is used to modify the gradient descent method of BP algorithm. In the combination of BP neural network, the modified formula of weight is as follows:

$$W_{ij}(t+1) = W_{ij}(t) - (J^T J + \mu I)^{-1} J^T e$$

(3)

In which, $W_{ij}(t+1)$ is the weight of the network of $t+1$ iteration; $W_{ij}(t)$ is the weight of the network of $t$ iteration; $J$ is Jacobian matrix of error to weight differentiation; $\mu$ is the constant when the value is greater than zero and can be adjusted adaptively, when the value is very large, formula is close to the gradient method, and when it is very small, the formula is close to Gauss-Newton method; $I$ is the unit matrix; $e$ is the error vector.

4. Simulation Experiment and Result Analysis

4.1 Variable selection based on Adaptive-Lasso

In this study, LARS algorithm is used to solve the adaptive-lasso estimation in the formula. For each given $\gamma$, this algorithm will find an optimal $\lambda_n, \gamma = 1$, and the analysis results are shown in Table 2.

| Factor Coefficient | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|--------------------|---|---|---|---|---|---|---|
| $x_1$              | 0.2734 | 0 | 0.1976 | 0.0456 | 0 | 0.1453 | 0 |

We can see from the table 1 that the coefficient of factors such as expectation achievement, curriculum interaction and group interaction, etc., is 0, which indicates that these variables are excluded in the process of modeling. At present, the curriculum is open for students of Grade One, and they have not yet been able to have the characteristics of autonomous learning. Therefore, there is no value judgment and tendency basis for the realization of self-goal, so that the ability of self-restraint is relatively poor. This is why the expectation achievement and curriculum interaction are excluded. At the same time, it is also because students’ learning mode continues the way of student-teacher interaction in senior high school, and generally needs to adapt to the process of group interaction, so the differences in group interaction between students are not fully demonstrated.

In order to verify the accuracy of the Adaptive-Lass algorithm, the collinearity test of the four selected factors is carried out, as shown in Table 3. If Tolerance is less than 0.1 or VIF is greater than 10, the collinearity exists. In this case, Tolerance is much greater than 0.1 and VIF is less than 10, so there is no multiple collinearity. It shows that when Adaptive-lasso method is used to construct the model, the variables with collinearity can be eliminated to reflect the advantages of Adaptive-lasso in multi-index modeling.

| Table 3 Collinearity Statistics |
|-------------------------------|
| Model | Tolerance | VIF |
|------|-----------|-----|
| 1 | $x_1$ | .956 | 1.046 |
|    | $x_2$ | .887 | 1.127 |
|    | $x_3$ | .974 | 1.027 |
|    | $x_4$ | .865 | 1.156 |

Dependent Variable: $y$

4.2 Prediction results of fuzzy neural network model

In this study, MATLAB software is used to complete the operation of LM neural network. In order to verify the accuracy of the prediction model, 247 students’ data were used as training set and 50 students’ answers were used as test set names. The settings are as follows: input node of LM neural network is 3,
output node is 2 and hidden layer node is 10, the number of display intervals is 25, the maximum number of cycles is 1000, the target error is 0.0, the initial $\mu$ is 0.001, the maximum number of check failures is 6 and the minimum error gradient is $10^{-7}$. The confusion matrix of training sample modeling is shown in the figure. In this study, in order to make a comparative experiment and predict student’s online learning satisfaction by logistic regression model, a regression model is established by using the variables selected in Section 4.1. Its confusion matrix is shown in the figure 1 and figure 2.

Figure 1. Confusion Matrix of LM Neural Network

Figure 2. Confusion Matrix of LR

The accuracy of the training samples shown in Figure 1 and Figure 2 reached 93.1% and 92.7% respectively. In order to further evaluate the prediction performance of the model, LM neural network model and contrast model are evaluated by the test samples. After ROC curves are established respectively and by contrast, we can see that the ROC curve of model LM is closer to the upper left corner and has better prediction performance than that of logistic regression model.

Figure 3. ROC Curve of LM Neural Network Is Constructed by Test Samples

Figure 4. ROC Curve of LR by Test Samples

5. Research Conclusion

In China, with the rapid development of Internet technology, online education is becoming more and more popular, and the means of teaching is becoming more and more mature. More and more people begin to get used to acquiring new knowledge in this way. However, in the process of learning, its influencing factors include many aspects, especially the learning experience including personal satisfaction, which is more and more attracted by teaching managers and organizers. How to improve students' learning experience is an urgent problem to be solved.

On the basis of literature research, this study summarizes five factors that affect college students' online learning satisfaction, and carries out data analysis. From the analysis process, it can be seen that students' perception of their own individuals, satisfaction of resource needs and interaction between teachers and students are the important factors that determine whether students are satisfied with online
learning. From this study, we can see that curriculum resources are an important part of building a popular online teaching course, and providing effective, comprehensive and attractive curriculum materials, and strengthening students' attention is the first thing to do.

In this study, curriculum interaction and student-teacher interaction are regarded as influencing factors, which is not different from traditional teaching mode, while the interaction between learners and learners is excluded from the influencing factors. However, we still believe that the focus of online teaching design should not ignore the interaction between students, which is also very important for the shaping of students' self-learning ability. In addition, through the analysis, we can see that students can evaluate the performance of the learning process objectively. Although the expectation achievement is not clear, they have basically mastered the skills of self-regulated learning, which can also enhance students' good experience in the process of online learning. Students with positive self-efficacy evaluation tend to get higher satisfaction with online learning.

In conclusion, in order to improve the satisfaction of college students in online learning, in the process of designing and implementing the course, we need to proceed from the students' own reality to improve the organizational links and teaching contents in a targeted way, and design better online learning courses. At present, our research is only aimed at the analysis of a single course, and in the future, we also need to carry out multi-course and multi-stage research.

Acknowledgments
This work was supported by National Natural Science Foundation of China (71874215, 71571191, 61671030), Beijing Natural Science Foundation (9182016, 9194031), MOE (Ministry of Education in China) Project of Humanities and Social Sciences (15YJCZH081, 17YJAZH120, 19YJCZH253), Fundamental Research Funds for the Central Universities (SKZZY2015021) and Program for Innovation Research in Central University of Finance and Economics.

References
[1] Klaas S. Factors Influencing Learners' Satisfaction in an Open E-Learning Environment[J]//Journal of Teacher Education for Sustainability.2016,13(1):29-42.
[2] Y.C. Kuo, A.E. Walker, B.R. Bell and K.E.E. Schroder. A Predictive Study of Student Satisfaction in Online Education Programs[J]//International Review of Research in Open & Distance Learning.2013,14(14),107-127.
[3] Y. C. Kuo, A.E. Walker, K.E.E. Schroder et al. Interaction, Internet Self-efficacy, and Self-regulated Learning as Predictors of Student Satisfaction in Online Education Courses[M]//Ann Arbor Science Publishers.1978.
[4] J. Costley, C. Lange. The Effects of Instructor Control of Online Learning Environments on Satisfaction and Perceived Learning [J]//Electronic Journal of e-Learning.2016,14(3):169-180.
[5] J. M. Simpson. Student Perceptions of Quality and Satisfaction in Online Education[J]//Proquest Llc.2012:132.
[6] J. Kranzow. Faculty Leadership in Online Education: Structuring Courses to Impact Student Satisfaction and Persistence[J]//Journal of Online Learning & Teaching,2013,9(1):131-139.
[7] Learning Environments: Evaluation from a Social Constructivist Perspective and implications for their Design[J].International Journal on E-Learning ,2006,5(5):593-622.
[8] J.H. Wu, R.D. Tennyson & T.L.A. Hsia. Study of Student Satisfaction in a Blended e-Learning System Environment[J]//Computers & Education,2010,55(1):155-164.
[9] H. Zou, The Adaptive Lasso and Its Oracle Properties[J]//Journal of Industrial and Management Optimization (JIMO), 2006, 101(476):1418-1429.
[10] R. Behling, A. Fischer, M. Herrich et al. A Levenberg-Marquardt method with approximate projections [J]//Computational Optimization and Applications.2014,59(1):5−26.