Research on Neural Network-Based Achievement Prediction of Middle School Students

Zixuan Yi[a]
Huangzhong University of Science and Technology, 1037 Luo Yu Road, Wuhan, China

Keywords: Neural Network, Grade Prediction, Linear Regression.

Abstract: With the development of information technology, so do the teaching methods. Predicting students' grades and allowing teachers and students to take appropriate countermeasures based on the predicted grades can be a great aid to teaching. This paper proposes a performance prediction method based on neural network with low data demand. The model uses students' scores in the last 3 exams. Based on a neural network with 4 nodes, which uses S-type function as the activation function, the model was established and fitted. The prediction accuracy of the model is demonstrated by a comparison chart of the predicted and real scores of 43 students. In comparison with linear regression method, the predictive power of the model is not lagging behind. The accuracy of the model is a little worse than the linear regression method's, while the model's average error is better. This shows that the neural network method in this study can be applied to practical secondary school teaching.

1 INTRODUCTION

With social progress, economic development, and the continuous progress of human society, the importance of education has also become increasingly prominent. Education work is a major event related to social development and the future of the country. Doing a good job in education not only prepares a high-quality generation for society, but also helps educated individuals achieve self-development and realize their life value. Specifically, education is beneficial for individuals to acquire knowledge, enhance their thinking and other abilities, thereby enhancing their overall quality and preparing them for future career life. In China, the content of nine-year compulsory education can enable individuals to acquire a complete, systematic, and multi-dimensional set of knowledge, laying a solid foundation for them to cope with future problems in various fields. Education can also promote social progress. It helps to cultivate talents with innovative and critical thinking in scientific thinking, which are necessary for scientific research, cultural innovation, and other fields. Therefore, education plays a role in promoting technological progress and social civilization development.

At present, the evaluation indicators for education are not complex, and grades (scores) are the most widely referenced and directly meaningful indicators for evaluating educational achievements. The higher the score a student obtains in an exam, the better it can basically represent his level of knowledge acquisition and mastery. Therefore, the significance of predicting fractions is self-evident.

Domestic and foreign scholars have conducted relevant research in the field of performance prediction. Zhang and Li used association rule algorithms to mine student grades, found the connections between courses, and then established a model of student grades using linear regression. They analyzed the impact of basic course learning on related professional course learning, providing a scientific theoretical basis for teaching management (Zhang and Li, 2020). Lin et al. used linear regression and deep neural network DNN to predict learners' grades. Experiments have shown that DNN can better fit the correlation between learning behavior and grades, achieving more accurate predictions of grades (Lin et al., 2019). Zhou et al. (2018) proposed a student performance prediction method based on BP neural network to predict the final grades of a certain grade C language course at Nanjing University of

[a] https://orcid.org/0009-0001-9914-149X

Yi, Z.
Research on Neural Network-Based Achievement Prediction of Middle School Students.
DOI: 10.5220/0012886800004508

In *Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence (EMITI 2024)*, pages 5-9
ISBN: 978-989-758-713-9

Posts and Telecommunications, as well as the college entrance examination scores of high school seniors. The effectiveness of this method was verified.

However, these methods did not consider the impact of differences between samples of different categories and the similarity of samples within categories on the prediction performance. Therefore, Wang et al. proposed a K-Means based student performance prediction model, combining the K-Means algorithm with multiple linear regression to predict student performance more targeted, providing more accurate reference information for teaching and improving the quality of school teaching (Wang et al., 2023). They also proposed a feature selection based score prediction method, which first uses sequence forward selection algorithm to perform feature selection on sample data, thereby selecting the optimal feature subset to construct a multiple linear regression prediction model. Then, the model is used to predict grades. The experimental results show that this method can improve prediction accuracy (Liu et al., 2023).

Deep learning methods have also been widely used in grade prediction (Bakhshinategh et al., 2019). Li et al. (2020) proposed a student grade prediction method based on dual attention mechanism, which achieves more comprehensive and accurate utilization of student attributes, thereby ensuring accurate prediction of student grades. However, machine learning and deep learning based methods usually require large-scale training data, Therefore, it is difficult to achieve performance prediction in the early stages of academic studies with insufficient data, so it is particularly important to supplement sparse data with information. Therefore, Liu et al. (2023) proposed a score prediction model based on multi-layer feature fusion to address the above issues. They constructed a two-layer historical score modeling module, which achieved synchronous feature extraction of the temporal dependence of grade information and course relevance; A similar student network was constructed based on co-occurrence frequency, integrating similar student characteristics for information complementarity to achieve timely prediction.

At present, the development of information technology has improved the level and efficiency of education, but at the same time, there are still problems waiting to be solved. The situation between students is different, and due to the differences in personality and needs, not all technologies and education methods are needed by everyone (Yan, 2009). Therefore, predicting students' grades can better and earlier formulate special training plans for each student, helping them unleash their subjective initiative earlier.

2 METHODS

2.1 Data Sources and Descriptions

The data in this paper comes from the four test scores of students in the first class of the second grade of Huaihua No. 2 Middle School, and the data files are .csv files. According to the time when the tests occur, they have been named class1_March, class1_April, class1_June, and class1_November. Then the data file structure will be introduced. And the file is in tabular form, using 10 rows, as shown in table 1. The indexes in the table correspond to: A-name, B- total score of the 1st exam, C- total score of the 2nd exam, D- total score of the 3rd exam, E- total score of the 4th exam.

Table 1: Part of the data used.

A	B	C	D	E
Zhou Yuhang	495	631	615	609
Zhou Yixiang	646	689	685	693
Zhou Jiahui	578	590	607	566
⋮	⋮	⋮	⋮	⋮
Zeng Zixuan	239	182	267	237
Chen Wenbo	311	321	373	358

2.2 Indicator Selection and Description

Learning rate, which is used to regulate the speed of the training process. Through the test, the learning rate was finally identified as 0.1. MSE (Mean Variance) Loss, which is a metric used to measure whether the current neural network is “good” or “bad”, the formula for MSE is as follows:

$$MSE = \frac{\sum_{i=1}^{n}(y_{true}-y_{pred})^2}{n} \quad (1)$$

where n is the number of samples, which in this case refers to the number of classmates in the class. The word y_{true} is the 4th exam score, i.e., the correct data used for validation, and y_{pred} is the prediction data, i.e., the output of the model. The smaller the MSE, the more accurate the prediction result of the model.

2.2.1 Data Pre-Processing

Due to the change in the full scores of Chinese, mathematics and English in the first and second grades of junior high school, the data is visualized first, and the students' scores are divided by the full

scores, that is, the normalized score rate is used as the data for calculation.

2.3 Methodology Introduction

Neural network: neural network is a machine learning technology that imitates the way of signal transmission and interaction between biological neurons, so as to achieve artificial intelligence and achieve the purpose of learning experience, the basic elements of the neural network are neurons, neurons have input and output, the output of one layer of neurons is used as the input of the next layer of neurons, and the output of the penultimate layer passes through the last layer of neurons to become the output of the model, that is, the output of the entire neural network, as shown in Figure 1 of the neural network used in this training. As shown in Figure 1, there are a total of 4 neurons, all of which function as (2):

$$y_j = \sum_{i=1}^{3} x_i * w_i + b_j \quad (2)$$

where y_j is the output result of node j, $bias$ is the offset, w_i (i = 1,2,3) are the weights, and x_j (j = 1,2,3) are the inputs.

Activation Function: The hidden layer and the output layer select the S-type function as the activation function, as shown in (3):

$$f(x) = \frac{1}{1+e^{-\theta x}} \quad (3)$$

Stochastic gradient descent method: used to optimize parameters, parameter optimization formula such as (4):

$$w_i = w_i - \eta \frac{\partial L}{\partial w_i} \quad (4)$$

where w_i represents the parameter to be optimized, η represents the learning rate, and L represents the loss function.

Backpropagation algorithms: used to calculate $\frac{\partial L}{\partial w_i}$. For example, the principle to calculate w_1 is shown in (5):

$$\frac{dL}{dw_1} = \frac{dL}{dy_{pred}} * \frac{dy_{pred}}{dh_1} * \frac{dh_1}{dw_1} \quad (5)$$

3 RESULTS AND DISCUSSION

3.1 Parameter Selection

The model was used to predict the results of the total scores of middle school students, and the results are as follows:

According to figure 1 and function (2), there are in total 12 weights and 4 offsets.

After iterating the results of 500 students for 100 times, the parameters take the values shown in table 2 in whose indexes correspond to: P- parameter, V-value (Table 2).

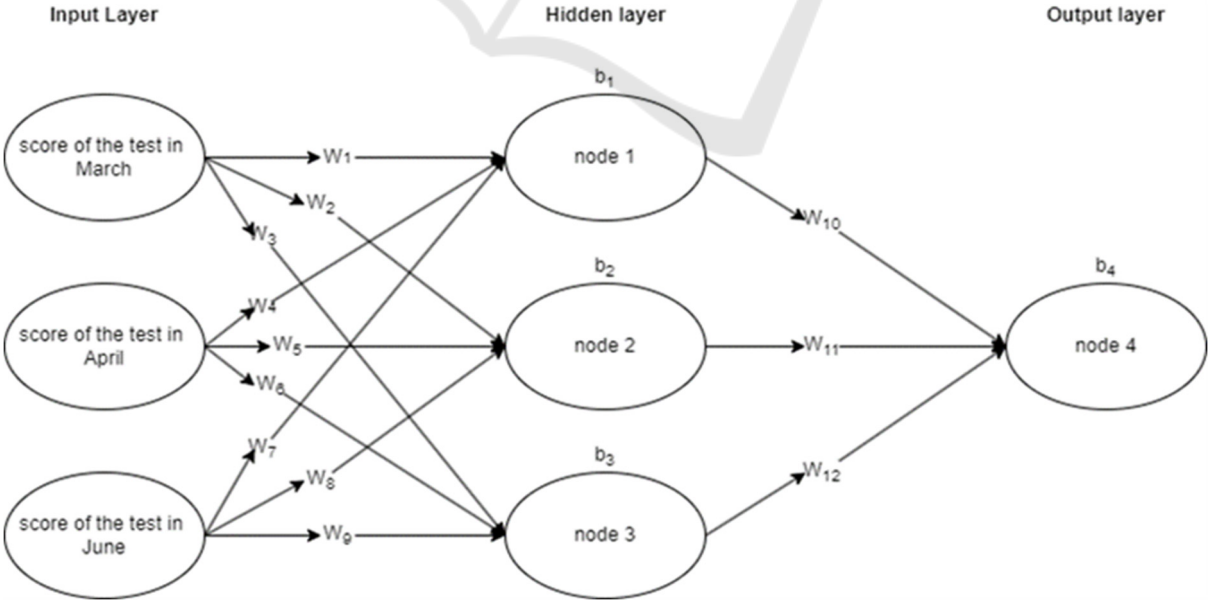

Figure 1: The structure of the neural network used.

Table 2: Parameters' values (leave three decimal places).

P	V	P	V	P	V	P	V
w_1	-0.049	w_5	0.166	w_9	-1.119	b_1	1.088
w_2	-0.936	w_6	2.494	w_{10}	-2.251	b_2	-2.842
w_3	-1.320	w_7	-0.452	w_{11}	4.315	b_3	0.746
w_4	0.377	w_8	-0.784	w_{12}	-1.734	b_4	0.879

3.2 Testing Results

When the learning rate is 0.1, the change trend of the loss function during the training process is shown in figure 2. With the iteration rounds increasing, the value of the loss function is declining, too. So, the training process is effective to improve the model's accuracy.

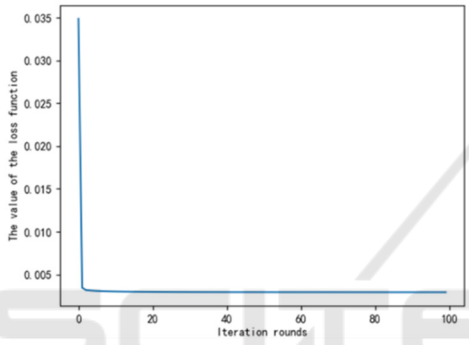

Figure 2: Changes of the loss function.

Then test the model with scores of another 43 students. The figure 3 is obtained by comparing the predicted Chinese results with the correct Chinese results, in which the red dots represent the true grades while the blue dots are the predicted one.

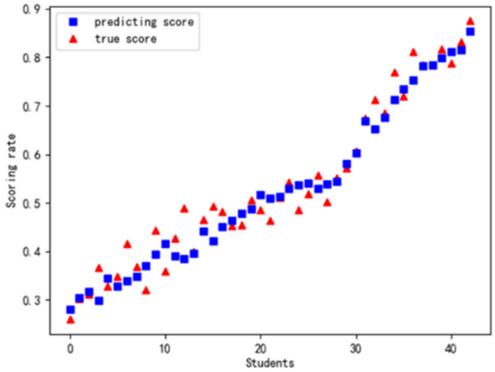

Figure 3: Comparing the predicted total scores with the correct total scores using neural network.

The predicting scoring rate of the model is basically consistent with the true ones, and the predicting accuracy also varies with the change of the score rate, and the accuracy reaches the highest in the range of about 0.5 to 0.65 in the true score rate, and there will be some deviation when the true score rate is high or low, because most of the students' scoring rates are between 0.5 and 0.65. When the score is lower or higher, the number of students decreases, and the training sample is less, which makes deviation easier to occur.

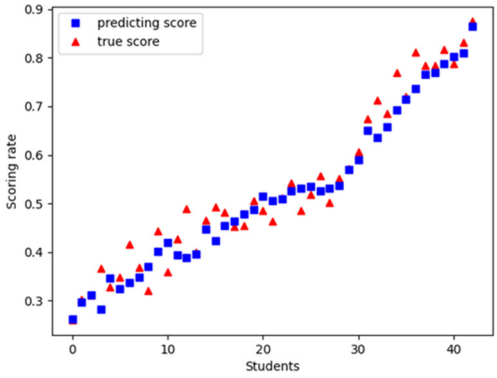

Figure 4: Comparing the predicted total scores with the correct total scores using linear regression.

In order to reflect the feasibility of the model, the linear regression method was used to predict the performance on the same sample, and the prediction results shown in figure 4.

Table 3: Comparing with the linear regression.

Model	Accuracy	Average error (scoring rate)
Neural network	76.744%	0.0290
Linear regression	81.395%	0.0311

Define the average error as the average difference between the predicted and true values. Define the proportion of samples with precision within 0.05 of the difference between the predicted value and the true value to the total number of samples. Compare the neural network's these two indexes with the linear regression's, the result is in the table 3. Though the accuracy is worse, the neural network has better average error. This means the way neural network is effective in predicting scores.

3.3 Discussion

As can be seen from Figure 2, the accuracy of performance prediction with neural networks is very high, which can meet the needs of predicting scores in actual teaching. Moreover, the neural network model training in this paper only uses the scores of

more than 500 students in the last three exams. This shows that the prediction method using neural networks does not need to obtain other characteristics of students, such as records of consumption behavior, and then calculate the similarity between students. It does not rely on diverse and hard-to-obtain data, nor does it require a large amount of data, making it ideal for use in secondary school teaching. In practice, teachers can take different actions on students based on predicted scores. For example, for a student with a low predicted score, the teacher can respond in advance to make the student study more seriously, and can also ask other students to help him answer questions and help him learn. Similarly, teachers can also pay more attention to the overall predicted performance of the class, and if the overall predicted performance is not satisfactory, the teacher needs to discuss with other teachers to discuss their own teaching shortcomings.

However, in this study, only changes in students' total scores were concerned. In the course of secondary education, there will be changes in the subjects that students learn, e.g. Physics will be added in Secondary 2, and Chemistry will be added but Biology and Geography will be added in Secondary 3. This results in a change in the total score of the exam. Since it is difficult to get a score in different subjects, for example, it is difficult to get a high score in a liberal arts subject but it is easy to get a certain score, while a low score in a single digit is common for a science subject, so if the proportion of liberal arts subjects increases, then the overall score of most students will increase. But in reality, the students' learning situation has not changed, only the subjects have changed. This cannot be found in a single study of the total score, so the foothold of future research can be refined from the total score and focus on the results of each subject.

4 CONCLUSION

Student achievement prediction has always been a very practical direction. With the development of information technology, the methods of statistical analysis of student performance are becoming more and more advanced. The application of computer technology to teaching is an unstoppable trend. In this paper, a performance prediction method based on neural network is proposed, and its operating principle, composition structure and computational function are introduced. The feasibility was tested by a prediction test of the performance of 43 students. However, the forecasting methods in this paper have their drawbacks. The frequency of exams for junior high school students is not high, and there are only two mid-term and final exams in a semester on average. The time span to obtain the results of the three exams is long, and it may take more than half a year. This may not allow teachers to take action sooner.

So, the method is more suitable for high schools where the test is more frequent. Alternatively, in future research, with regard to the acquisition of experimental data, the number of exams referred to can be reduced, and the data of other dimensions can be appropriately increased.

REFERENCES

Bakhshinategh, B., Zaiane, O., Elatia, S., 2018. Educational data mining applications and tasks: A survey of the last 10 years. *In Education & Information Technologies.*

Hao, Z., 2019. Deep learning review and discussion of its future development. *In MATEC Web of Conferences.*

Lin, P., He, X., Chen, T. et al., 2019. Prediction of loss and teaching intervention for learners in MOOC from perspective of deep learning. *In Computer Engineering and Applications.*

Liu, X., et al., 2023. Research on student scores prediction method based on feature selection. *In Information Technology.*

Li, M., Wang, X., Ruan, S. et al., 2020. Student performance prediction model based on two-way attention mechanism. *In Journal of Computer Research and Development.*

Liu, T., Qi, H.R., Ni, W., 2023. Multilayer feature fusion based model for predicting student academic performance. *In Computer Engineering and Design.*

Wang, G., Liu, H., Li, J. et al., 2023. Research on student score prediction method based on K-Means. *In Information Technology.*

Yan, S., 2009. The alienation and its alleviation of modern informational technology. *In Modern Educational Technology.*

Zhang, X., Li X., 2020. Analysis of students' achievements based on multiple linear regression. *In Computer and Digital Engineering.*

Zhou, J., Xue, J., Han, C. et al., 2018. Research on student performance prediction method based on BP neural network. *In Computer Era.*