# A Score Prediction Model based on Deep Learning for Teaching Quality Analysis

**Kan Ji, Bangying Wu[1], Qiang Yao[2]**

Network and Information Center, Southwest Petroleum University, Chengdu, China

[1]Corresponding Author

**Abstract:** ***Based on the increasing amount of campus data, more and more scholars are starting to explore potential patterns from it to improve teaching mode and promote teaching quality. Teaching quality analysis has gradually become a hot research area. The-state-arts-of approaches such as big data analysis, data mining, machine learning, and deep learning have begun to be widely applied in research on teaching quality, and student score is one of the crucial indicators for measuring teaching quality. In this paper, we construct a score prediction model based on deep learning to analyze the teaching quality of teacher. This model can also be applied to evaluation indicators for teaching quality analysis other than student score. The score prediction model based on neural networks solves the problem of inaccurate evaluation results of traditional teaching quality evaluation methods on nonlinear regression models. Based on known student score, a score prediction model based on deep learning can help teacher assess student' mastery level of course knowledge precisely and predict student final exam score to achieve the effect of early warning for student in the final exam.***

**Keywords:** Teaching quality analysis, deep learning, neural networks, student score, early warning

## 1. Introduction

Teacher teaching quality analysis is one of the significant tasks of teaching evaluation in universities. Teaching quality reflects the level teaching of teachers and the quality of teaching results, which is related to development of teachers and students both equally. There are many evaluation indicators for the teacher teaching quality in many different aspects, including student score, student evaluation, teaching efficiency, teaching methods, classroom atmosphere, and teaching attitude. Among them, student score is an important criterion for evaluating the teaching quality of teacher. Student score not only assesses mastery level of knowledge points at a student in a course, but also reflects level of teaching skills at a teacher. The evaluation indicators for student score and teaching quality analysis have a similar structure. For example, the evaluation indicators for student score include student attendance score, student homework score, student mid-term score, and student final exam score, etc. The evaluation indicators for teaching quality analysis consists of student score, teaching attitude, classroom management, and teaching methods, etc.

Many mainstream technologies have been applied to the field of teaching quality analysis nowadays, including teaching quality analysis based on big data technology, teaching quality analysis based on probability statistics, teaching quality analysis based on machine learning, and teaching quality analysis based on deep learning. There are many evaluation indicators for teaching quality, and the impact of various evaluation indicators also exists difference. Teaching quality analysis belongs to nonlinear regression problems. Although the above methods are widely used in research in the field of teaching quality analysis, the first three methods are difficult to exclude subjectivity and randomness in the evaluation factors of teaching quality, resulting in deviations in the accuracy of teaching quality evaluation and unable to predict future teaching quality evaluations at the mean time.

Therefore, this paper proposes a score prediction model based on deep learning called DLSPM, which uses student attendance score, homework score, and mid-term score to predict student final score. It is worth mentioning that student homework score is composed of multiple different chapters score. The main contributions of our study can be summarized as three-fold:

- We design a three-layer neural network BP model to learn the feature of student score and achieve the prediction of student final exam score. Owing to the similar structure of the evaluation indicators for student score and teaching quality analysis, this model can be applied to other evaluation indicators for teaching quality analysis as well.
- The accuracy of score prediction models based on deep learning is much higher than traditional methods such as data analysis or machine learning.
- The non-linear regression problem of teaching quality analysis can be solved more accurately. Through our model, potential information in student score can be excavated. It can provide early warning of student learning situations, and improve teacher teaching effectiveness.

## 2. Related Work

Based on the previous investigate results, in this chapter, we will briefly introduce the four commonly used methods for analyzing teaching quality of teacher.

**Teaching quality analysis based on big data technology.** Yueet al. [1] construct a football teaching quality evaluation system based on the background of big data by building a Hierarchy Model according to an evaluation system they build. Wang et al. [2] establish a three-level teaching evaluation system and develop CTQ evaluation and feedback system based on the analysis of broader data.

**Teaching quality analysis based on probability statistics.** Zhang et al. [3] provide an efficient statistical analysis method to evaluate online teaching quality by establishing a comprehensive evaluation model based on factor analysis theory. A novel evaluation method for online teaching based on the analytical hierarchy process (AHP) and Dombi weighted partitioned Muirhead Mean (PMM) operator under Fermatean fuzzy (FF) environment is put forward by Zeng et al [4].

**Teaching quality analysis based on machine learning.** Qi et al. [5] build a SVMs classifier on a dynamic student model and makes the customized learning resource suggestion. A classifier based on weighted naive Bayes is analyzed and designed for teaching evaluation in the study of Qiao [6].

**Teaching quality analysis based on deep learning.** Zhu et al. [7] employ the Ada Boost's multicore neural network learning algorithm to learn several weak classifiers and combine them into a single strong classifier. Zhuang et al. [8] proposed an integrated evaluation model based on deep learning technology incorporating YOLOX model, Retinaface model, and SCN model.

## 3. System Model

The implementation of the DLSPM will be described in detail in this chapter.

### 3.1 The evaluation system of teaching quality

There are several factors including student score, teaching attitude, classroom management, and teaching methods influence the teaching quality. Figure 1 outlines the structure of teaching quality evaluation systems. The student score is core element among these evaluation indicators.

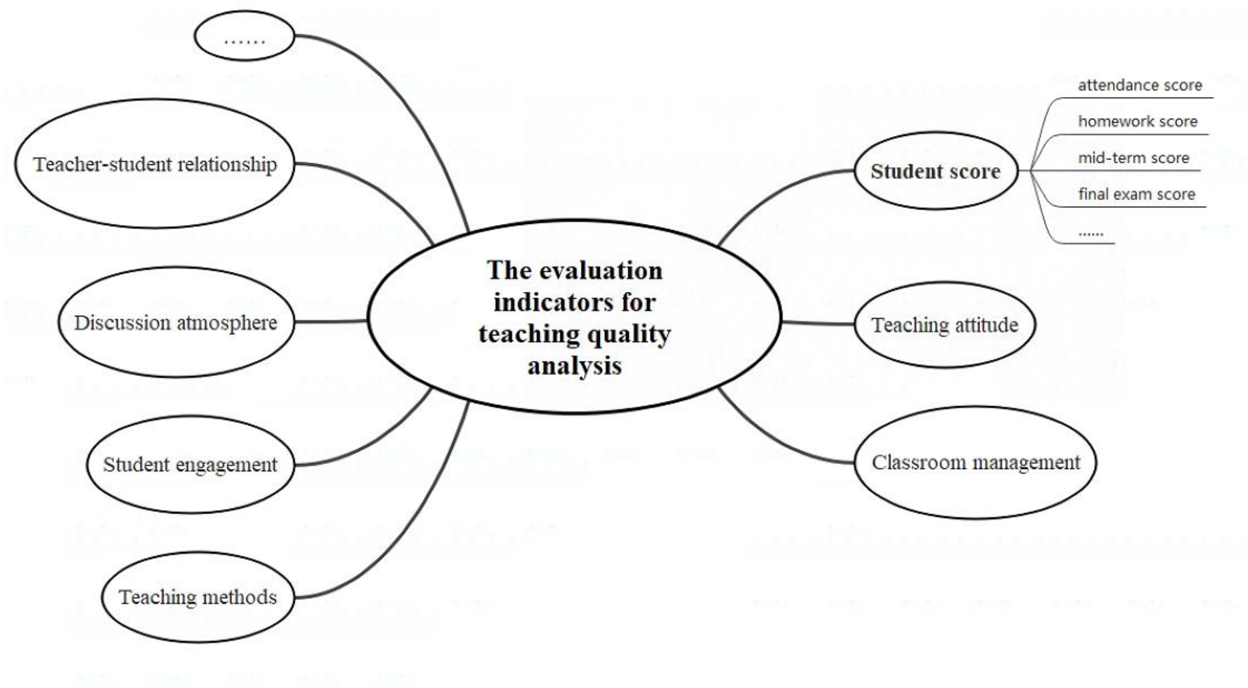



**Figure 1:** The structure of teaching quality evaluation systems.

### 3.2 The architecture of DLSPM

The DLSPM we build is a three-layer BP neural network as shown in Figure 2. The first layer is input layer receive the input feature vectors. There are 64 neurons in the input layer. And the number of neurons in each layer can be flexibly adjusted. The middle layer, also known as the hidden layer, serves as the second layer. The number of hidden layer and neurons are set to 1 and 32 respectively in our model. And the number of hidden layers and the number of neurons can also be flexibly changed. Output layer is the last layer in the neural network.

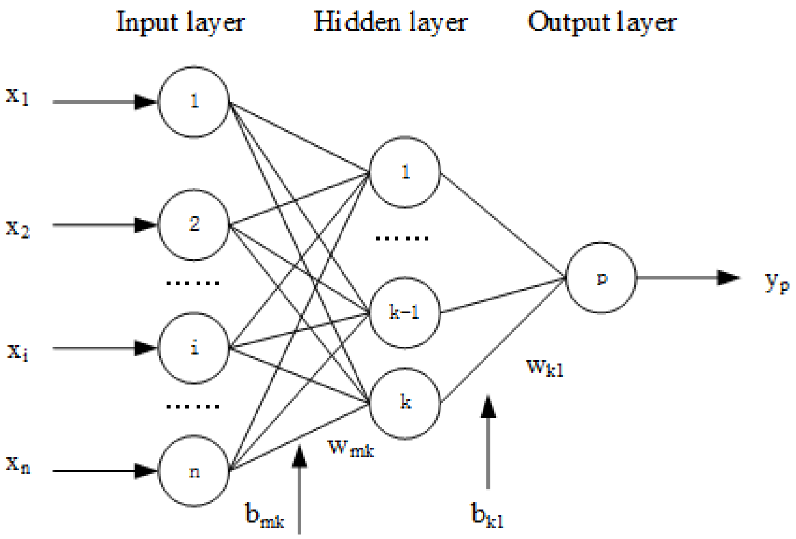



**Figure 2:** A three-layer BP neural network in our method.

The results generated by output layer will be sent into an activation function Relu that we choose in our model. The biggest advantage of using the Relu activation function is that when $x > 0$, the gradient will not disappear, which enables the backpropagation to proceed smoothly. The Figure 3 displays the Relu activation function. The Relu activation function is defined as:

$$f(x) = \max\left(0, \mathbf{w}^{\mathrm{T}}\mathbf{x} + b\right) \tag{1}$$

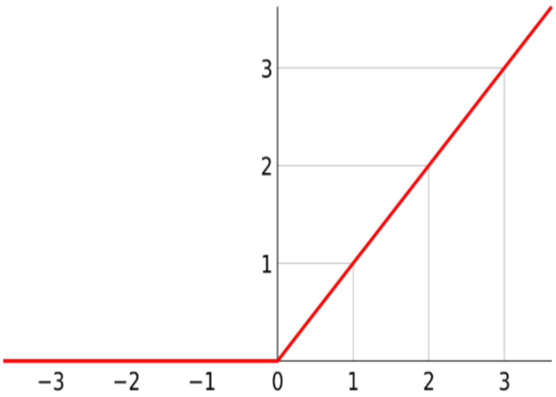



**Figure 3:** Relu activation function.

Subsequently, the optimization function will be employed to optimize the parameters of the model. We take a random gradient descent algorithm called SGD to complete backpropagation. The SGD algorithm is defined as:

$$\boldsymbol{W} \leftarrow \boldsymbol{W} - \eta \frac{\partial L}{\partial \boldsymbol{W}} \tag{2}$$

where $\boldsymbol{W}$ represents the weight parameters that need to be updated, gradient of loss function $L$ with respect to $\boldsymbol{W}$ denoted as $\frac{\partial L}{\partial \boldsymbol{W}}$, and $\eta$ stands for learning rate.

Finally, the training set in the preprocessed data is used to train the model DLSPM. The preprocessed procedure will be explained in next chapter in detail. During the training process, we check the training performance of the model through several commonly used regression model evaluation indicators and adjust some key parameters such as the number of neurons, layers, training cycles, batch, and learning rate as necessary. To this extent, the overall architecture of DLSPM can be outlined in Figure 4.

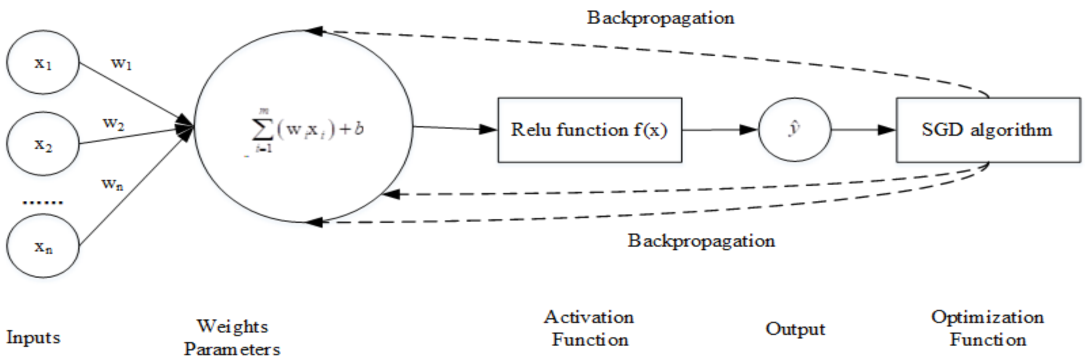



**Figure 4:** The architecture of DLSPM

### 3.3 Model implementation process

The implementation steps of the model can be summarized in Figure5.

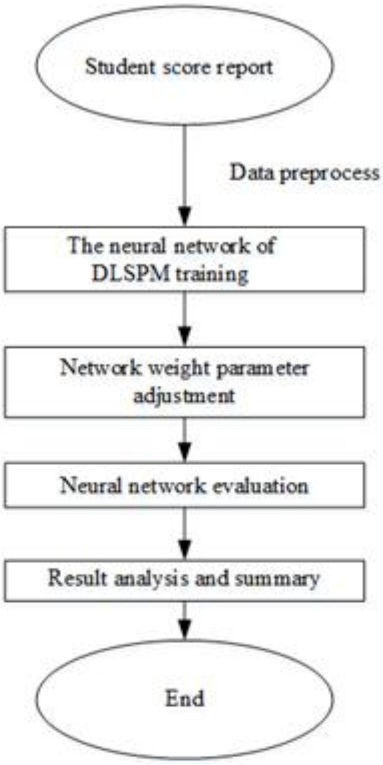



**Figure 5:** The implementation steps of DLSPM.

## 4. Experiment and analysis

In this part, we utilize the a real-word dataset which exports from educational management system on the model DLSPM. And then present the evaluation results of our model.

### 4.1 Dataset

The dataset used in our study is a score report of the entire 2022 grade students on Computer Think course. There are 5332 records of different students in all. And each record has many attributes such as serial number, school, student number, course number and so on. The training set, validation set, and test set are allocated in 3:1:1. The original dataset cannot be implied in model DLSPM directly. We need to accomplish the data process. Before introducing the experiment procedure, we summarize primary notations used in this paper in Table 1.

### 4.2 Data Process

Each attribute in a record can be regarded as a feature. Some of the attributes are useless in helping with our model training. Even worse, the useless attributes become noisy factors that deteriorate the outcome of parameter optimization. To solve the negative effects introduced by these useless attributes, we execute the steps as shown in the following.

- We remove non-numeric attributes include school, student name, teacher name and course status.
- Some numeric attributes which meaningless in our model training are dropped, such as serial number, student number and course number.
- The rest of these attributes will be used for our experiment after some attributes which without values are all set to 0.

**Table 1:** Notations used in our research

| Symbol | Description |
|---|---|
| $\hat{y}_i$ | The degree of difference between the estimator/predict result. |
| $y_i$ | The estimated quantity/true value. |
| $\overline{y}$ | The mean value of error of true value. |

### 4.3 Experimental Indicators

Mean square error (known as MSE), Root mean square error(known as RMSE) and $R^2$ score are three standard methods used in regression model.

**MSE:** A measure that reflects the degree of difference between the estimator/ predict result $\hat{y}_i$ and the estimated quantity/true value $y_i$, defined as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{3}$$

where $n$ is the number of the estimators.

**RMSE:** RMSE is a common indicator in the field of regression. It is root on the result of MSE that can narrow down the gap between the estimator and the estimated quantity at the same dimensional level, defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4}$$

***$R^2$* score:** $R^2$ score utilize the mean value of error of true value as bench mark denoted as $\overline{y}$. $R^2$ score reflects the difference between $\hat{y}_i$ and $\overline{y}$. There are three kinds of outcomes of $R^2$ as shown in Table 2.

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

**Table 2:** The implication of several different $R^2$ score results

| Results of $R^2$ score | Description |
|---|---|
| =1 | The maximalvalueof$R^2$score. In this situation, $\hat{y}_i = y_i$ and loss=0. Achieve the most ideal predict model. |
| =0 | $\hat{y}_i = \overline{y}$ in this scenario. $\hat{y}_i$ can be directly calculated by $\overline{y}$ or the predict model. |
| <0 | The predict model has no significance in existence. |

### 4.4 Experimental Results

We conduct our experiment applying the student score report dataset which are processed into model DLSPM and then demonstrate the evaluation indicators of DLSPM by compare with several baselines we set.

The first baseline called BSL1 has the same experimental procedures as DLSPM except the input data without the mid-term exam scores feature.

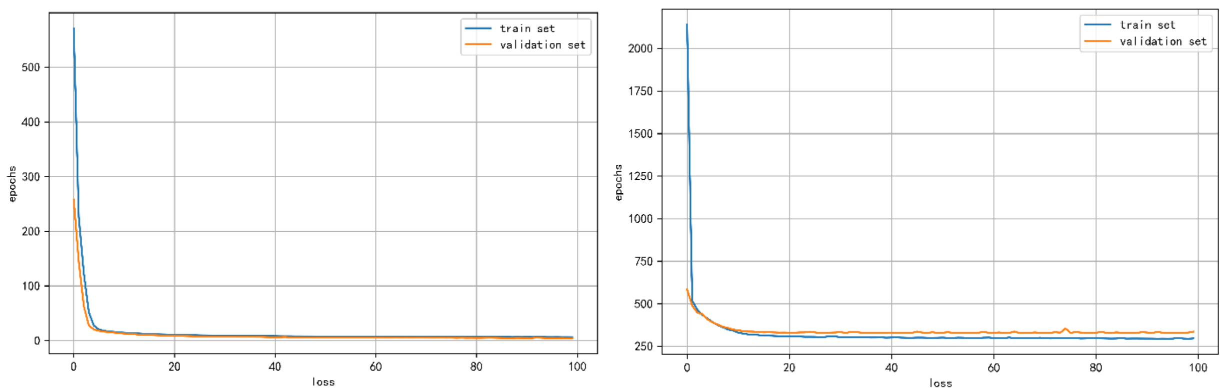
Random forest is denoted as BSL2 which represents one of the machine learning algorithms as the second baseline.

The Table 3 presents the performance comparison among these methods mentioned before. The results suggest that the final exam score is highly influenced by the mid-term score.

**Table 3:** The performance comparison between DLSPM and other baselines

| Methods\Indicators | MSE | RMSE | R2 score |
|---|---|---|---|
| DLSPM | 1.334 | 1.155 | 0.998 |
| BSL1 | 371.021 | 19.262 | 0.397 |
| BSL2 | 6.072 | 2.464 | 0.990 |

The loss function curves of the first two methods in Table 3 are shown in Figure 6(a) and Figure 6(b). The results support the viewpoint that significant of mid-term score in score predict model.




(a) (b)

**Figure 6:** The loss function curves of DLSPMandBSL1.

Mathematical statistics method is also utilized for further affirming the best behavior our model has. We calculate the mean value student final score and mid-term score respectively. And compare the absolute value of difference of these two mean value with the RMSE of DLSPM. We find that the value of the former one 8.1 is much greater than that of the latter 3. This discovery indicates that the predict result our model DLSPM generated is more precise than traditional mathematical statistics method.

## 5. Conclusion

In this paper, we study four state-of-art methods which are commonly used in teaching quality analysis and argue that student score is the most essential indicators in teaching quality analysis. We develop a novel student score predict model based on neural network. The score report dataset worked pretty well on this model. And the experiment results declare the score predict model based on neural network has the best performance than other methods. Through analysis on the result, we observe that the student final exam score is highly influenced by the mid-term score.

Due to the same architecture the score report and the indicators of teaching quality have, this student score predict model based on neural network can be also used as an ANN model in teaching quality evaluation system. Specifically, some attributes we removed from dataset can also have potential influence on the performance of predict model. For instance, teacher name attribute can relate to a new table includes teacher professional titles, how many years have teacher worked and so on. Taking these multi-source data together to dig out the latten pattern of teaching quality is worth pursing in future work.

## References

[1] L. Yue, Z. Wei. Evaluation of Football Teaching

Quality Based on Big Data. Computational and Mathematical Methods in Medicine. Volume 2022, Article ID 7174246, 10 pages.

[2] J.W. Wang, C.S. Liu, W. G. Teaching Quality Evaluation and Feedback Analysis Based on Big Data Mining. Mobile Information Systems. Volume 2022, Article ID 7122846, 8 pages.

[3] S.M. Zhang, Q.J. Tong. An efficient statistical method for evaluating online teaching quality. The Teaching Reform Project of Xi'an University of Post and Telecommunications 2021.

[4] S.Z. Zeng, P. Yan, H.H. Jin. Online Teaching Quality Evaluation of Business Statistics Course Utilizing Fermatean Fuzzy Analytical Hierarchy Process with Aggregation Operator. Systems 2022, 10, 63.

[5] S. Qi, L. Liu, B.S. Kumar, A. Prathik. An English teaching quality evaluation model based on Gaussian process machine learning. Expert Systems, 2022, Vol.39, No.6, e12861, ISSN 0266-4720.

[6] L.J. Qiao Teaching Quality Evaluation of Ideological and Political Courses in Colleges and Universities Based on Machine Learning. Journal of Mathematics. Volume 2022, Article ID 2835029, 10 pages.

[7] Y. Zhu. Research on Education Teaching Quality Analysis Based on the Neural Network Model. Security and Communication Networks. Volume 2022, Article ID 2644032, 8 pages.

[8] W. Zhuang, F.N. Xing, J.L. Fan, C.M. Gao, Y.H. Zhang. An Integrated Model for On-Site Teaching Quality Evaluation Based on Deep Learning. Wireless Communications and Mobile Computing. Volume 2022, Article ID 9027907, 13 pages.

[9] Z. Nin, F. Wang. Quality Assessment of Vocational Education Teaching Reform Based on Deep Learning. Computational and Mathematical Methods in Medicine. Volume 2022, Article ID 1499420, 11 pages.

[10] S.W. Li, M.F. Shi. Optimization and Quality Evaluation of Online Teaching Courses Based on Machine Learning. Journal of Sensors. Volume 2022, Article ID 5081505, 12 pages.