Available online at www.ijpe-online.com

International Journal of Performability Engineering

vol. 15, no. 6, June 2019, pp. 1560-1569
DOI: 10.23940/ijpe.19.06.p7.15601569

Data-Driven Student Learning Performance Prediction based on RBF Neural Network

Chunqiao Mi[a,b,*]

[a]*School of Computer Science and Engineering, Huaihua University, Huaihua, 418000, China*
[b]*Key Laboratory of Intelligent Control Technology for Wuling-Mountain Ecological Agriculture in Hunan Province, Huaihua, 418000, China*

Abstract

With the expansion of college enrollment in recent years, the quality of students' learning is beginning to decline. At present, education quality governance has become the internal demand of the reform and development of higher education. Learning performance prediction is an important means to effectively resolve the academic crisis and improve the overall education quality. In this study, firstly, the current status and problems about learning performance prediction were analyzed from the perspective of basic data, evaluation indicators, and prediction methods. Secondly, driven by ten items of basic learning situation data, a learning performance prediction model based on the RBF neural network was established, which included three layers in network topology: the input layer, hidden layer, and output layer. The activation functions of the hidden layer and output layer were a Gauss radial basis function and linear function, respectively. The modeling process included three steps: forward propagation computing prediction loss, error backward propagation adjusting network parameters, and network optimization determining model hyperparameters. The obtained results showed that the trained model had small relative root mean square error values for both the training data and testing data. When comparing the original observation values and model predicted values, it was observed that most of the sample points were evenly distributed on both sides of the diagonal line of the contrast graph, which indicates that the RBF neural network model employed in this study is promising in learning performance prediction. It is of good reference significance for promoting more accurate and efficient learning performance prediction and improving the efficiency and effectiveness of education quality governance.

Keywords: learning performance prediction; RBF neural network; education quality governance

(Submitted on March 20, 2019; Revised on April 5, 2019; Accepted on June 7, 2019)

1. Introduction

At present, with the expansion of the enrollment scale of colleges and universities in our country, the situation of college students failing in course learning is becoming increasingly serious, and more and more students often fail to graduate on time [1]. Therefore, the governance of education quality has become an inherent demand of education reform and development. In this context, in 2017 and 2018, the Ministry of Education issued some related documents to promote the construction of world-class universities and strengthen undergraduate education. It was pointed out that the quality and effect of personnel training in education should be taken as the basic criterion to test all of our work as educators, and that the deep integration of modern information technology with education and the governance of learning process should be strengthened. Therefore, quality governance and early warning education have become a basic demand in higher education. Prediction and evaluation of learning performance are core basic problems in early warning education management, and they are also practical problems that need to be solved urgently. Therefore, driven by student learning situation big data, it is of good practical significance to predict student learning performance based on the RBF neural network data mining method, which can help improve the quality of personnel training in the current era's higher education.

* Corresponding author.
E-mail address: michunqiao@163.com

2. Current Research Status Analysis

2.1. The Research on Basic Data Used for Prediction

The formation of students' learning performance is related to many factors, so the selection of basic data can directly affect the quality of the predicted results. The basic data of learning situations is the precondition of learning performance prediction and early warning education management. Some related works could be traced back to the investigation of the learning situation, which rose in the 1980s [2]. At present, with the in-depth application of information technology in education, more and more learning-related data have accumulated, and the mainstream data used in the analysis and prediction of learning performance mainly comes from the learning behavior records left by students in various online learning systems such as LMS [3]. In addition, some kinds of offline survey data are also important supplements, including students' offline daily performance and test records [4]. However, there is a lack of research on students' psychology, teacher-student interaction, learning environment, and other related data, and the lack of open data sets in the education domain is also a major obstacle to learning performance analysis and early warning research. In the future, with the application of artificial intelligence technology in education, more and more intelligent sensors will be used to monitor students' external behavior, physiology, and psychological state. Therefore, it will be an important trend to integrate more types of basic learning situation data based on big data and cloud computing technologies, which will also promote the improvement of learning performance prediction.

2.2. The Research on Evaluation Indicators

The selection of evaluation indicators for learning performance prediction can directly affect the comprehensiveness of learning situation analysis and the accuracy of learning performance prediction. At present, there are two main types of indicators that are widely used. The first one is about the expression of students' personalized demographic characteristics, which is used to describe students' learning styles, attitudes, and behaviors, such as data items about students' gender, age, family support, participation of learning groups and associations, etc. [5-6]. The second type is about the description of the effectiveness of students' learning process and results, which reflects students' individual learning strengths and weaknesses, such as data items about attendance, answering questions, extracurricular reading, extracurricular activity achievements, etc. [7-8]. However, students' learning performance is the combined result of a variety of internal and external factors, but most of the current studies only focus on students' external behavior while ignoring the internal psychological status. Therefore, with the gradual mature application of intelligent sensors, intelligent wearable devices, and other artificial intelligence technologies in the field of education, it will be an inevitable trend in the future to improve the accuracy of learning performance prediction by combining curriculum learning indicators, students' background indicators, and students' psychological indicators.

2.3. The Research on Prediction Algorithms

Using data mining technology to predict students' learning performance can be traced back to the beginning of the 21st century, and related research can be divided into two categories. The first category involves statistical methods. For example, discriminant analysis [9], factor analysis [10], logistic regression analysis [11], principal component analysis, and multiple linear regression analysis [12] were all used to study the relevant factors affecting students' performance and to predict their learning performance. However, most of these attempts were based on the analysis of learning-related static result data, rather than the modeling of dynamic learning process data. Thus, the results obtained were mostly descriptive rather than predictive, and the accuracy was usually not very high. The other category is machine learning approaches. For example, decision tree and association rules [13], support vector machine [14], Bayesian algorithm [15], BP artificial neural network [16], genetic algorithm [17], and fuzzy comprehensive evaluation [18] were all applied in the prediction of students' learning performance and identification of at-risk students with learning difficulties. From these, it can be noted that research on learning performance prediction methods has shown a trend of algorithmization and automation. However, the application of the RBF (radial basis function) neural network in solving such problems is still rare. The RBF network uses RBF as the "base" of the hidden unit to form the hidden layer space, so that input vectors can be mapped directly into the hidden space without the need of weights connection. This can greatly speed up the learning speed and avoid the local minimal problem. In this study, learning performance prediction based on the RBF neural network is studied, which can provide a new method reference for this field.

3. Materials and Methods

3.1. Source Data Collection, Pre-Processing, and Calculation

Student-centred learning situation source data is the basic starting point of learning performance prediction. Thus, in this

study, basic data related to students' curriculum learning were first collected comprehensively from relevant information systems such as the student roll system, educational administration system, personnel management system, and teaching assistant system. Additionally, some students' personal and background data were also supplemented by questionnaires, personal dialogues, and interviews. Then, all of the data were pre-processed by unified methods of quantization, normalization, desensitization, and cleaning. Some noise data, such as outliers and missing values, were also detected, marked, repaired, or eliminated. Finally, all kinds of data were integrated into a database with the student ID number as a unique identification to form student-centred learning situation big data, which could be managed in a unified way and lay a foundation for learning performance prediction.

The formation of learning performance is influenced by many factors. In this study, aiming at comprehensively reflecting the influencing factors of learning performance, based on the formation mechanism of learning performance and the behavior data of learning process, four categories of influence factor were identified according to the relevant standards of education industry and the relevant specifications of student performance management in our university. They are shown in Table 1. The first category, personal background information, convers a student's gender, a student's age when he or she entered university, and a student's major discipline, which can reflect the basic demographic characteristics of a student. The second category is basic learning potential, which includes a student's high school status, reflecting whether he or she attended a key high school or not before entering university. It also includes a student's obtained college entrance examination score, which can indicate his or her basic learning ability. Learning activity participation is the third category; it can imply the enthusiasm of a student in learning activities and include his or her leadership role and attendance record in learning activities. The last category is daily earned score, which can represent the daily efforts a student has paid during the learning process, and it includes a student's assignment cumulative earned score, student's online quiz cumulative earned score, and student's offline test cumulative earned score. The concrete evaluation indicators used in this study are the corresponding independent variables X_1 to X_{10}, which are shown in the second column of Table 1. Y, the dependent variable, represents a student's final performance and is our prediction goal. The data type and description details of the variables are also shown in Table 1.

Table 1. The information of source data used in this study

Influence factor category	Evaluation indicator name	Data type	Description
Personal Background Information	X_1	string	student's gender, values in ('male' or 'female')
	X_2	number	age when student entered university
	X_3	string	student's major discipline, values in ('engineering' or 'liberal arts')
Basic Learning Potential	X_4	string	status of a student's high school where he or she attended before entered university, values in ('key high school' or 'ordinary high school')
	X_5	number	student's college entrance examination score
Learning Activity Participation	X_6	number	student's course attendance score
	X_7	string	student's leadership status, indicating whether a student was a learning group or class leader during course studying, values in ('yes' or 'no')
Daily Earned Score	X_8	number	student's assignment cumulative earned score
	X_9	number	student's online quiz cumulative earned score
	X_{10}	number	student's offline test cumulative earned score
Final Performance	Y	number	student's final score

In order to eliminate the influence of different units used by different variables on the model, all the continuous number raw data, including X_2, X_5, X_6, X_8, X_9, X_{10}, and Y, were normalized before they were fed into our model. The method used is shown in Equation (1), where X is the original raw data and X^* is the normalized data. X_{MIN} and X_{MAX} are the maximum and minimum values of the original X. For discrete string raw data, including X_1, X_3, X_4, and X_7, binary processing was carried out, making their values of 1 or 0. Taking X_1 (namely, the student's gender) as an example, in our binary processing the string value of 'male' was transferred into 1 and 'female' into 0, and the others were treated similarly.

$$X^* = \frac{X - X_{\mathrm{MIN}}}{X_{\mathrm{MAX}} - X_{\mathrm{MIN}}} \tag{1}$$

After all pre-processing and computation, a sample data set including 418 students was obtained for this study. It was then randomly resampled into a training set and a testing set by ratio of 3:1. Therefore, our training data set and testing data set finally included 313 and 105 students respectively, and each student had ten features, X_1 to X_{10}, and one observed final performance value, Y.

3.2. The RBF Neural Network

The radial basis function (RBF) method was proposed by Powell in 1985. It is a scalar function with radial symmetry. Any function satisfying the characteristic of $\Phi(x) = \Phi(\|x\|)$ can be called a radial basis function. Standard usage generally involves the Euclidean distance (also known as the Euclidean radial basis function), although other distance functions are also possible. Therefore, the radial basis function is a real value function whose value depends only on the distance from the origin or the distance from any point c, i.e., $\Phi(x, c) = \Phi(\|x-c\|)$, where the point c is called the centre of the radial basis function. The effect of the radial basis function is often local, that is, when x is far from the centre point, the function value is very small. The most commonly used radial basis function is the Gauss kernel function in the form of $\Phi(\|x-c\|) = \exp\{-\|x-c\|^2/(2*\sigma^2)\}$, where c is the centre of the kernel function and σ is the width parameter that controls the radial range of the function. In 1988, Moody and Darken proposed a RBF neural network structure based on the radial basis function, which can approximate an arbitrary continuous function with arbitrary precision. It is especially suitable for solving the problem of non-linear prediction, such as in learning performance prediction.

The formation of learning performance is influenced by many subjective and objective factors and has great uncertainty. The neural network method can simulate the non-linear relationship affected by multi-factors and improve quantitative prediction. The RBF neural network can map low-dimensional input vectors into high-dimensional space, and it can transform linear inseparable problems in low-dimensional space into separable problems in high-dimensional space. Therefore, it can approximate arbitrary non-linear functions, mine the regularity of complex relationships, and has good generalization ability and fast learning speed, which is suitable for solving uncertain problems of learning performance prediction.

3.3. Information Forward Propagation Computing Prediction Loss

The RBF neural network is a three-layer neural network, which includes an input layer, a hidden layer, and an output layer. The transformation from the input layer to the hidden layer is non-linear, while that from the hidden layer to the output layer is linear. The topology of the RBF neural network constructed in this study is shown in Figure 1. The number of nodes in the input layer is assumed to be m (in this study, m = 10 corresponding to the ten independent variables in Table 1), the hidden layer node number is assumed to be h (needs to be determined and optimized in the network optimization process), the output layer has one node (corresponding to the dependent variable in Table 1), the activation function of the hidden layer node is a Gauss radial basis function, and that of the output layer node is a pure linear function.

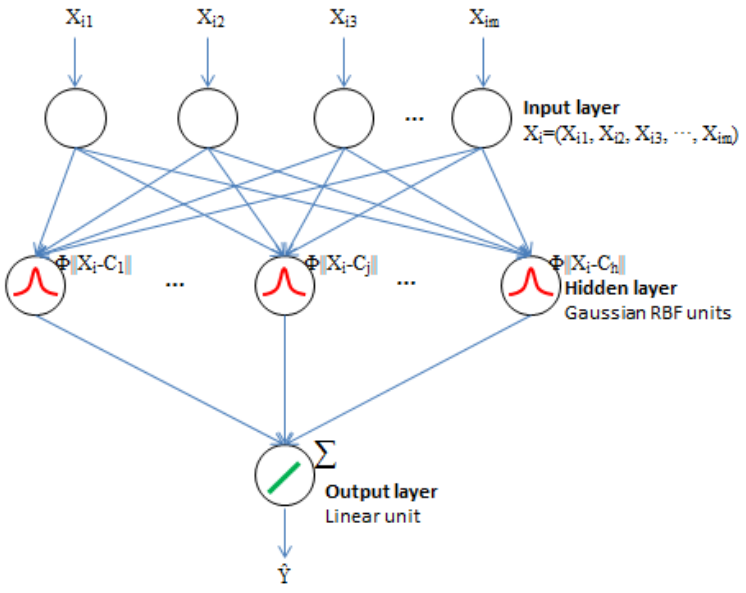

Figure 1. The construction of our RBF neural network model

Suppose there are n samples (in this study n = 418) in the training data set X, namely $X = \{x_1, x_2, \cdots, x_n\}$, where the i^{th} ($1 \le i \le n$) training sample is $x_i = \{x_{i1}, x_{i2}, \cdots, x_{im}\}$, that is, the attribute number of each sample is m (corresponding to the number of nodes in the input layer). Then, for the training sample x_i, the output of the j^{th} node in hidden layer can be

calculated by Equation (2), where $c_j = \{c_{j1}, c_{j2}, \cdots, c_{jm}\}$ and σ_j are the centre and width parameters of the j^{th} node activation function, respectively ($1 \le j \le h$).

$$\Phi(x_i, c_j) = \exp\left(-\frac{\|x_i - c_j\|}{2(\sigma_j)^2}\right) \tag{2}$$

The output of the hidden layer is then mapped as the input of the output layer by the weight matrix between the hidden layer and output layer. The activation function of the output layer node is a pure linear function, so the output of the output layer can be obtained by Equation (3), where $\hat{y}_i$ is the predicted value by the output layer node corresponding to the sample x_i, w_j is the weight between the j^{th} hidden layer node and the output layer node, and b is the bias of the output layer node.

$$\hat{y}_i = \sum_{j=1}^{h} w_j \Phi(x_i, c_j) + b \tag{3}$$

Then, after all samples are fed into the network, the loss function of the RBF neural network can be defined as Equation (4), where L is the loss function value and e_i is the error between the predicted value ($\hat{y}_i$) and the observed value (y_i) of the i^{th} sample.

$$L = \frac{1}{2n}\sum_{i=1}^{n}(e_i)^2 = \frac{1}{2n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{2n}\sum_{i=1}^{n}\left(y_i - \left(\sum_{j=1}^{h} w_j \Phi(x_i, c_j) + b\right)\right)^2 \tag{4}$$

3.4. Error Backward Propagation Adjusting Network Parameters

The objective of model training is to make the value of the loss function as small as possible. From Equation (4), it can be seen that the value of the loss function depends on the values of unknown parameters $\{c_j\}$, $\{\sigma_j\}$, and $\{w_j\}$ ($1 \le j \le h$). Therefore, the gradient descent method is used to adjust the values of each parameter so as to reduce the value of the loss function. The specific calculation method is shown in Equations (5) and (6), where α is the number of iterations and η is learning rate, which is a real number between 0 and 1.

$$\begin{cases} c_j(\alpha+1) = c_j(\alpha) + \Delta c_j(\alpha) \\ \sigma_j(\alpha+1) = \sigma_j(\alpha) + \Delta\sigma_j(\alpha) \\ w_j(\alpha+1) = w_j(\alpha) + \Delta w_j(\alpha) \\ b(\alpha+1) = b(\alpha) + \Delta b(\alpha) \end{cases} \tag{5}$$

$$\begin{cases} \Delta c_j(\alpha) = -\eta\dfrac{\partial L}{\partial c_j} = \eta\dfrac{1}{n}\dfrac{w_j}{(\sigma_j)^2}\displaystyle\sum_{i=1}^{n} e_i\Phi(x_i, c_j)(x_i - c_j) \\ \Delta\sigma_j(\alpha) = -\eta\dfrac{\partial L}{\partial \sigma_j} = \eta\dfrac{1}{n}\dfrac{w_j}{(\sigma_j)^3}\displaystyle\sum_{i=1}^{n} e_i\Phi(x_i, c_j)\|x_i - c_j\|^2 \\ \Delta w_j(\alpha) = -\eta\dfrac{\partial L}{\partial w_j} = \eta\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} e_i\Phi(x_i, c_j) \\ \Delta b(\alpha) = -\eta\dfrac{\partial L}{\partial b} = \eta\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} e_i \end{cases} \tag{6}$$

3.5. Network Optimization Determining Model Hyperparameters

The accuracy of prediction results by the RBF neural network model is related not only to the values of network parameters such as weights, but also to hyperparameters including the number of hidden layer nodes, learning rate, and iteration times. Because there is no unified method to determine these hyperparameters, they can only be determined through a trial-and-error method based on empirical knowledge. In order to get relatively optimal values for these hyperparameters, the relative

root mean square error (RRMSE) as shown in Equation (7) is used to compare the results of different trials. The smaller the RRMSE value, the better the result.

$$\text{RRMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \Big/ \left(\frac{1}{n}\sum_{i=1}^{n} y_i\right) \tag{7}$$

4. Result and Discussion

In order to implement our RBF neural network model, the Python (v3.5) programming language and TensorFlow (v1.1) APIs were used in this study. They can effectively construct the network computation graph and realize the process of model training, parameter optimization, and result prediction and testing. Using these tools, the model hyperparameters, which include the hidden layer node number indicated by h, learning rate denoted by η, and training iteration number represented by a, were determined using a trial-and-error method based on domain knowledge experience. The processes and results are shown in Table 2.

Table 2. The processes and results of determining our model hyperparameters

Parameter	Determining process	Optimal result
Hidden layer node number (h)	Keep η = 0.5 and a = 5000, test h with values in interval of [5,25] by step = 1	16
Learning rate (η)	Keep h = 16 and a = 5000, test η with values in interval of (0,1) by step = 0.05	0.4
Training iteration number (a)	Keep h = 16 and η = 0.4, test the effect of iteration number a	70000

Firstly, we kept η = 0.5 and a = 5000, and we let h vary in the interval of [5, 25] with step = 1. Then, different RRMSE values were obtained, as shown in Figure 2. From this figure, it can be seen that RRMSE reaches the lowest value when h = 16. Therefore, we used 16 as the optimal value of h.

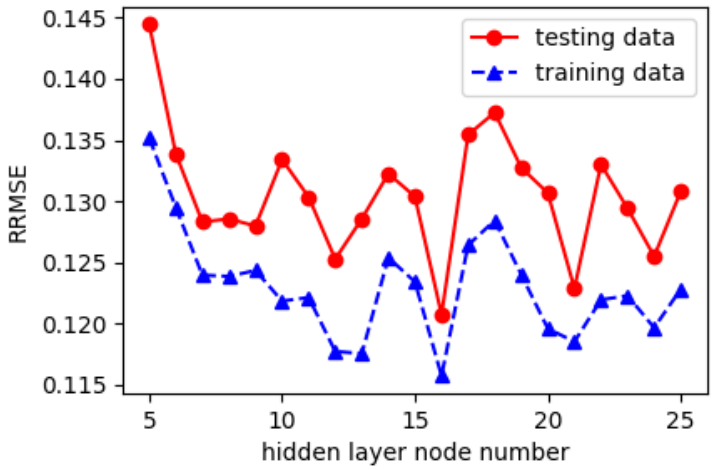

Figure 2. Investigation of the effects of different node numbers in the hidden layer

Secondly, with the optimal value of hidden layer node number being determined, we set h = 16 and a = 5000 as constant, and the values in the interval of (0,1) with step = 0.05 were used to explore the optimal value of learning rate η. The errors at various trials are shown in Figure 3. The lowest RRMSE was obtained at η = 0.4, which means the optimal learning rate in our study was 0.4.

Thirdly, in conjunction with h = 16 and η = 0.4, the effect of different iterations was investigated, and it is represented in Figure 4. The stop criterion we adopted was to obtain the critical point where the RRMSEs of both our training data and testing data became relatively stable. After that, the RRMSE of our testing data slightly increased while that of our training data slightly decreased; this indicated an overfitting occurrence. From Figure 4, it can be seen that when the training process reached the point of 7000 iterations, the RRMSE of our training data and that of our testing data were both relatively stable, but afterwards the latter began to increase slightly while the former continued to decrease. The overtraining phenomenon occurred. Therefore, it could be determined that, for our model in this study, the optimal iteration number was 7000.

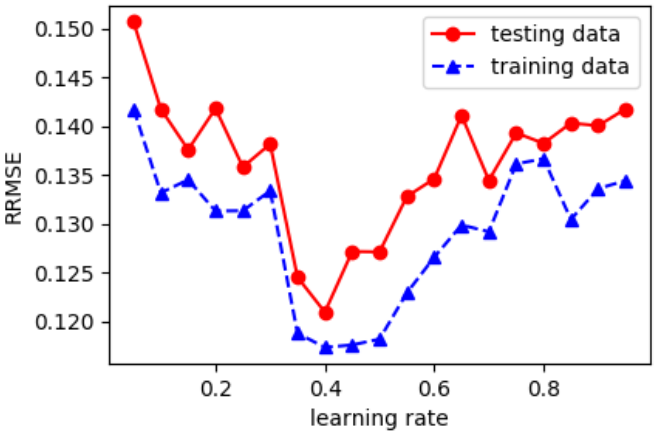

Figure 3. Investigation of the effects of different learning rate values

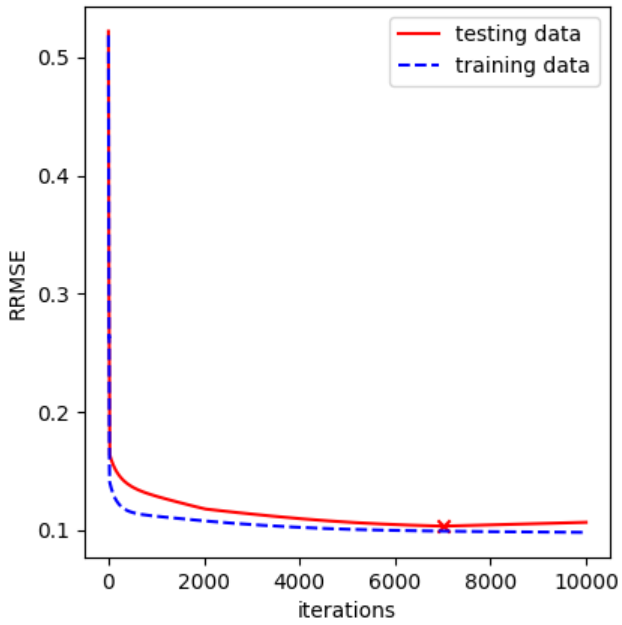

Figure 4. Investigation of the effects of different training iteration times

After the hyperparameters were determined, we used them again to train the model and obtained the final values of the network parameters, namely the values of $\{c_j\}$, $\{\sigma_j\}$, and $\{w_j\}$ $(1 \le j \le h)$. Once the hyperparameters and network parameters were all determined, the final model of our RBF neural network was determined. At this time, the corresponding prediction value could be obtained by feeding new sample data into the network. For the training sample set and testing sample set, the RRMSE values of our final model obtained in this study were 0.1055 and 0.1163, respectively. The accuracy of training data was slightly better than that of testing data, which meant that our trained RBF neural network had slightly better fitting ability in learning performance prediction compared to its generalization ability. Generally speaking, these results were relatively very ideal in the field of education practice. This also showed that our model learned the relevant patterns in the sample data effectively. Thus, the model obtained is promising and can be used in the practice of student learning performance prediction.

Additionally, in order to better demonstrate the predictive ability of the model obtained, we further compared the original observation values with the predicted values of our model, as shown in Figure 5. By observation, we notice that in Figure 5, for the samples whose observation values were less than 0.6, the predicted values of our model tended to be higher than expected, as shown by the points in the circle located at the lower left corner of the figure. Meanwhile, when the observation values were greater than 0.8, the predicted values of the model contrarily tended to be lower than expected, as shown by the points in the circle located at the upper right corner of the figure. Thus, it can be seen that the model obtained in this study has a certain degree of optimistic estimation for students with very poor performance and a certain degree of pessimistic estimation for students with very good performance.

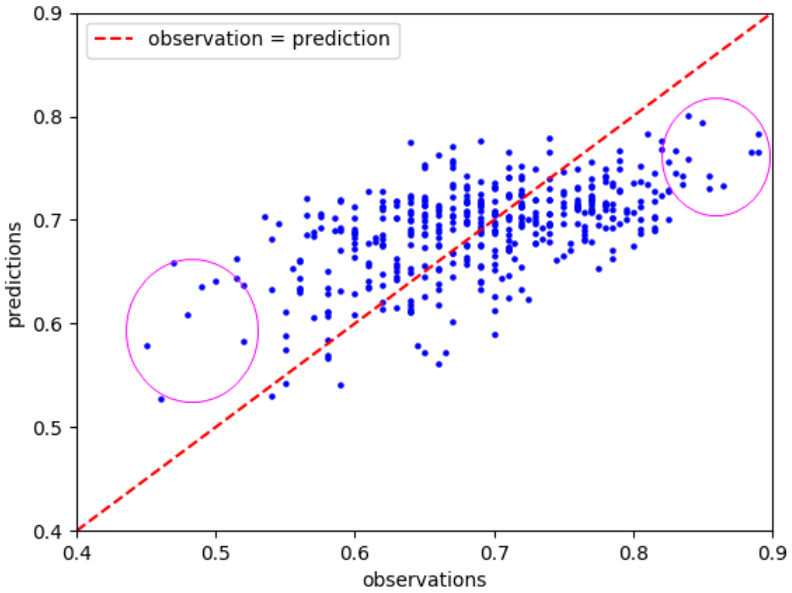

Figure 5. Comparison between model predictions and sample observations

One possible reason is the imbalanced distribution of our sample data points. That is to say, in educational practice, only a very small number of students perform particularly poorly or excellently, while most of them usually perform at about an average level or so. The same is true for the sample data in this study, whose specific distribution is shown in Figure 6. In a total of 418 samples, there are only 53 samples with an observed performance value less than 0.6 and 41 samples with an observed performance value greater than 0.8, while there are 324 samples with an observed performance value between 0.6 and 0.8. Therefore, the number of sample points distributed at both end regions (less than 0.6 or greater than 0.8) is far smaller than that in the middle region (between 0.6 and 0.8). In this situation of imbalanced sample data, more attention is usually paid to the majority ones, while the minority ones are overlooked during the process of algorithm modeling. This could result in a decrease in prediction accuracy of the model in both end regions of the sample distribution interval. Another possible reason may be related to the smoothing effect of the model itself, that is, a model often shows a certain smoothing effect and a certain low accuracy in the two end regions compared to the middle region. This reminds us that in practice, we should be alert to the prediction at both end regions. However, in general, we can see from Figure 5 that most of the sample points are evenly distributed on both sides of the diagonal line. This means that the predicted values are consistent with the observed values, so the model in this study has good predictive ability as a whole.

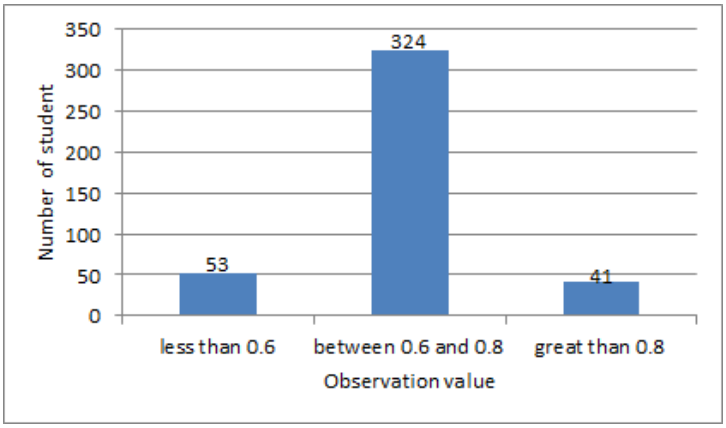

Figure 6. Distribution of sample observations

5. Conclusions

Student learning performance prediction has always been a difficult problem because there are many influencing factors that can add uncertainty and affect the predicted result. The RBF neural network is suitable to deal with this type of problem. The data flow direction in the RBF neural network is from training samples to RBF hidden layer to weight matrix to output layer. Between the input layer and the hidden layer, RBF is used as the "base" of the hidden unit to form the hidden layer

space, so that the input vector can be mapped directly into the hidden space without the need for a weight connection. Therefore, the mapping of the RBF neural network from input to output is non-linear, while the output of the RBF neural network is linear for adjustable parameters, which can greatly speed up the learning speed and avoid the local minimum problem. The results of this study showed that the data-driven learning performance prediction model based on the RBF neural network could accurately model the impact of multiple factors on learning performance, reducing the uncertainty of learning performance prediction. Specifically, this study used ten basic items of learning situation data to model the formation of student learning performance, and 418 samples were used to investigate and model the hidden relationship. By training the adopted RBF neural network, the parameters were finally optimized as follows: hidden layer node number h=16, learning rate η=0.4, and training iteration number α=7000. For the training sample set and testing sample set, the final RRMSE values of the model were 0.1055 and 0.1163, respectively. By displaying and comparing the original observations and model predictions in a scatter graph, it was shown that most of the sample points were evenly distributed on both sides of the graph diagonal line, showing consistency between observations and predictions. In summary, the RBF neural network model employed in this study is promising in student learning performance prediction, and it is of good reference significance for education quality governance in practice.

However, the results of this study also show that in the front end of the sample data interval, that is, for those students who perform poorly, the prediction of the model is somewhat higher. Meanwhile, for those who perform well in the back end of the sample data interval, the prediction of the model is somewhat lower. This tells us that in the practice of education quality governance, we should keep a pessimistic attitude towards the prediction of the front end of the sample interval and an optimistic attitude towards the prediction of the back end, which is more in line with the actual student performance situation and more conducive to the governance of education quality in practice. Further analysis reveals that this may be mainly due to the fact that the number of samples at both ends is much smaller than that in the middle region and the smoothing effect of the model itself. Therefore, in the next step, more feature data will be added into the model to reflect the factors affecting learning performance more comprehensively. Additionally, learning situation analysis and performance prediction modeling in the context of high-dimensional and imbalanced educational big data will be further explored, so as to improve the prediction accuracy and adaptability of the model.

Acknowledgments

This study is supported by the project "Data-driven study on risk assessment and early warning of learning situation in Hunan Local Universities" (No. 17YBQ087), granted by the Hunan Provincial Philosophy and Social Sciences Foundation.

References

1. C. Mi, Q. Deng, X. Peng, D. Yin, and Y. Liu, “The Business Process Optimization of Early Warning Education in Colleges and Universities —Taking H College for Example (in Chinese),” *Modern Educational Technology*, Vol. 28, No. 3, pp. 92-98, 2018
2. S. Zheng, “An Investigation and Analysis of College English Students' Learning Situation in Our College (in Chinese),” *Journal of Huizhou University*, No. 3, pp. 86-92, 1988
3. L. Razzaq, J. Patvarczki, S. F. Almeida, M. Vartak, M. Feng, N. T. Heffernan, et al., “The Assistment Builder: Supporting the Life Cycle of Tutoring System Content Creation,” *IEEE Transactions on Learning Technologies*, Vol. 2, No. 2, pp. 157-166, 2009
4. S. K. Yadav, B. Bharadwaj, and S. Pal, “Data Mining Applications: A Comparative Study for Predicting Students' Performance,” *International Journal of Innovative Technology & Creative Engineering*, Vol. 1, No. 12, pp. 13-19, 2011
5. T. M. Christian and M. Ayub, “Exploration of Classification using NB Tree for Predicting Students' Performance,” in *Proceedings of the International Conference on Data and Software Engineering*, pp. 1-6, Bandung, Indonesia, 2014
6. C. Romero, M. I. López, J. M. Luna, and S. Ventura, “Predicting Students' Final Performance from Participation in On-Line Discussion Forums,” *Computers & Education*, Vol. 68, pp. 458-472, 2013
7. S. T. Jishan, R. I. Rashu, N. Haque, and R. M. Rahman, “Improving Accuracy of Students' Final Grade Prediction Model using Optimal Equal Width Binning and Synthetic Minority over-Sampling Technique,” *Decision Analytics*, Vol. 2, No. 1, pp. 1-25, 2015
8. U. B. Mat, N. Buniyamin, P. M. Arsad, and R. Kassim, “An Overview of using Academic Analytics to Predict and Improve Students' Achievement: A Proposed Proactive Intelligent Intervention,” in *Proceedings of the IEEE 5th International Conference on Engineering Education*, pp. 126-130, Selangor, Malaysia, 2013
9. B. B. Minaei and W. Punch, “Using Genetic Algorithms for Data Mining Optimization in an Educational Web-based System,” in *Proceedings of Genetic and Evolutionary Computational Conference*, pp. 2252-2263, Chicago, Illinois, USA, 2003
10. J. P. Campbell, “Utilizing Student Data within the Course Management System to Determine Undergraduate Student Academic Success: An Exploratory Study,” Doctoral Dissertation, pp. 31-61, Purdue University, 2007
11. R. S. Baker, D. Lindrum, M. J. Lindrum, and D. Perkowski, “Analyzing Early at-Risk Factors in Higher Education E-Learning Courses,” in *Proceedings of the 8th International Conference on Educational Data Mining*, pp. 150-155, National University for Distance Education, Madrid, Spain, 2015
12. S. J. H. Yang, O. H. T. Lu, A. Y. Q. Huang, J. C. H. Huang, H. Ogata, and A. J.Q. Lin, “Predicting Students' Academic

Performance using Multiple Linear Regression and Principal Component Analysis," *Journal of Information Processing*, Vol. 26, pp. 170-176, 2018
13. J. Bravo, S. Sosnovsky, and A. Ortigosa, "Detecting Symptoms of Low Performance using Prediction Rules," in *Proceedings of the 2nd Educational Data Mining Conference*, pp. 31-40, Universidad de Cordoba, Cordoba, Spain, 2009
14. S. M. Jayaprakash, E. W. Moody, E. J. M. Lauría, J. R. Regan, and J. D. Baron, "Early Alert of Academically at-Risk Students: an Open Source Analytics Initiative," *Journal of Learning Analytics*, Vol. 1, No. 1, pp. 6-47, 2014
15. A. K. Hamoud, A. M. Humadi, W. A. Awadh, and A. S. Hashim, "Students' Success Prediction based on Bayes Algorithms," *International Journal of Computer Applications*, Vol. 178, No. 7, pp. 6-12, 2017
16. C. Mi, X. Peng, and Q. Deng, "An Artificial Neural Network Approach to Student Study Failure Risk Early Warning Prediction based on TensorFlow," *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Vol. 219, pp. 326-333, 2018
17. W. Xing, R. Guo, E. Petakovic, and S. Goggins, "Participation-based Student Final Performance Prediction Model Through Interpretable Genetic Programming: Integrating Learning Analytics, Educational Data Mining and Theory," *Computers in Human Behavior*, Vol. 47, pp. 168-181, 2015
18. C. Mi, Q. Deng, J. Lin, and X. Deng, "A Dynamic Early Warning Method of Student Study Failure Risk based on Fuzzy Synthetic Evaluation," *International Journal of Performability Engineering*, Vol. 14, No. 4, pp. 639-646, 2018

Chunqiao Mi received his Ph.D. degree from the College of Information and Electrical Engineering at China Agricultural University, China in 2012. He is currently an associate professor in the School of Computer Science and Engineering at Huaihua University, China. His research interests include data science and educational information technology.