Grade Prediction Model Based on DeepCycle Neural Network Classification Algorithm

Yuanfu Mao\(^1,\ast\), Xiaoying Sun\(^1\)

\(^1\)NanChang Institute of Science and Technology, Nanchang, China, 330108

\*Corresponding author e-mail: 1722521178@qq.com

Abstract. There are many factors affecting students' grades, which make the prediction of students' grades present high dimensional and nonlinear characteristics. Therefore, the traditional method has a large error in the prediction results, which is difficult to meet the practical needs. With the rapid development of artificial neural network (Ann), the deep cycle neural network algorithm based on Ann provides a new approach for student achievement prediction. In order to further improve the accuracy of student achievement prediction, this paper proposes a performance prediction model based on deep cycle neural network algorithm. First, principal component analysis is used for data dimensionality reduction processing of the established student writing evaluation system, and the first five principal components are extracted. Then, these principal components are taken as the input of the neural network to construct a three-layer neural network prediction model. The experimental results show that, compared with the single RBF neural network and BP neural network, the prediction model under deep cycle neural network is simple in structure, fast in convergence and 21.6% higher in prediction accuracy, which verifies the effectiveness of the model proposed in this paper.

Keywords: Deep Cycle, Neural Network Algorithm, Achievement Prediction Model, Effectiveness Analysis

1. Introduction

With the development of computer technology and network technology, the popularization and promotion of student information management system and teaching management system, the teaching data in the teaching platform is growing rapidly, but currently the use of data is still limited to simple query and screening [1-2]. In addition, students' scores are often treated with simple operations such as average score, highest score, mean square error, etc., but the relationship between scores cannot be found. For example, what factors affect scores are all problems worthy of in-depth exploration [3-4].
Adaptive neurofuzzy reasoning system is mainly used in finance, engineering, material design, decision management and other fields in the past research, but rarely used in the prediction of educational learning performance. Therefore, Su Chung-ho applied rough set theory to extract the core set and generation rules of learning achievement prediction. In order to evaluate the performance of the model, Su Chung-ho collected the VCCSEGLS data set as the experimental data set and compared it with other models [5]. The aim of Leana-Taclar was to explore the relationship between resources and motivation components and to identify which are predictors of academic achievement for Turkish students (440 students, 206 fourth graders and 234 seventh graders). Leana-taclar assessed the resources described in the AMG using the Education and learning Capital Questionnaire, and assessed the motivational components using the classroom Intrinsic and extrinsic motivational orientation scales [6]. Felipe Valencia has developed a robust energy management system using model predictive control theory as a mathematical framework. Based on fuzzy prediction interval model, robust EMS model is established. The model allows the representation of nonlinear dynamic behavior and the uncertainty of available energy at NCES. In particular, the uncertainty of wind-based energy can be expressed [7].

Based on the data of students’ historical learning performance, this paper studies the behavior data of students' daily professional courses and extracurricular activities, so as to accurately predict students’ performance and provide early warning for students [8-9]. Then, the deep circulation neural network algorithm is used to explore the influencing factors of students' performance, excavate the relationship between various courses, and give targeted Suggestions for school teaching arrangements [10].

2. Performance prediction model based on deep neural network

2.1. Deep cycle neural network algorithm

The optimization method of deep cycle neural network algorithm is based on gradient. The activation function must be differentiable everywhere. In general, the hidden layer will have Sigmoid function and the output layer will have a linear activation function, so that the network output can take any value. Since the value range of Sigmoid function is (0,1), if the output layer of BP neural network adopts Sigmoid function, the output of the whole network will be in the range of (0,1):

$$J(W, b, x, y) = \frac{1}{2} \| h_{w,b}(x) - y \|^2$$

(1)

Where \( h \) is the output after forward conduction of \( X \), \( W \) is the weight of neurons, \( B \) is the bias phasor of neurons, and the overall cost function is:

$$J(W, b) = \left[ \frac{1}{m} \sum_{i=1}^{m} J(W, b, x^{(i)}, y^{(i)}) \right]$$

(2)

2.2. Achievement prediction model

Pearson correlation coefficient was used to calculate the correlation between each attribute and
academic performance, and the correlation between each attribute was ranked based on the calculated results. According to the ranking results of academic achievement correlation, redundant attributes (with small correlation) are eliminated, and the attributes that have significant influence on academic achievement are retained. The retained attributes become the influential factors of academic achievement. For the prediction of learners' academic performance, it is necessary to first use the classification algorithm to classify and predict the data subsets composed of the influencing factors and grades of learners' academic performance, and then generate the prediction model. Then, the data containing only the influencing factors of learners’ academic performance were substituted into the prediction model for classification prediction to generate the corresponding learner's academic performance prediction.

3. Design of the experiment

3.1. Experimental background

In the assessment of students, student performance assessment is an important index to evaluate teaching quality. If students' past performance and other information can accurately predict the future trend of students, it can improve the way of cultivating students and promote the improvement of teaching quality. However, in practice, the value hidden behind these teaching data has not been fully exploited. The emergence of neural network algorithm alleviates the problem that the data in the database is not fully utilized. The grade prediction model USES the deep cycle neural network classification algorithm to analyze the data generated in the process of education and serve for teaching.

3.2. Experimental design

This experiment by deep circulation neural network classification algorithm, through the student achievement and behavior data mining analysis, explore curriculum and curriculum, the relationship between curriculum and student behavior, use SPSS Modeler19.0 software, according to the experiment set minimum support and minimum confidence, find meet the requirements of frequent itemsets, provide effective basis for teaching activities. The experimental sample data were from students majoring in Internet of Things and computer science in a university. The main public courses of these two majors were taken as factors affecting students' scores. Through data processing, 400 pieces of student score data and behavioral information data were obtained. The experimental results are shown in Table 1.

| The consequent   | In the preceding paragraph | Percentage of support | Percentage of confidence |
|------------------|----------------------------|------------------------|--------------------------|
| Higher mathematics | Linear algebra             | 76.471                 | 82.051                   |
4. Discussion of results prediction results

4.1. Analysis of grade prediction model based on deep cycle neural network classification algorithm

As shown in Figure 1, the final analysis it is concluded that the confidence level between various subjects were 82.051%, 78.571%, 78.049%, 75%, illustrates the link between the higher mathematics, linear algebra, probability theory relationship is very high, good students general linear algebra, advanced mathematics is also good, the same general higher mathematics good students, linear algebra is better. However, under the premise of good linear algebra, the confidence of higher mathematics is higher than that of higher mathematics. According to these, schools can take linear algebra as the prerequisite course of higher mathematics, which is beneficial to improve students' scores. After standardizing the raw data, the covariance coefficient matrix of the standardized data is calculated. It can be seen that the covariance coefficient between the initial index X1 and the index X2-X11 is large, the covariance coefficient between the index X2 and the index X1, X3-X11 is large, and the covariance coefficient between X3 and the index X1, X4-x11 is also large, etc. This indicates that there is serious information overlap between the original indicators, leading to information interference between each indicator. If these original indicators are directly used as the input of Rbf network, it will inevitably affect the convergence speed of the network model and lead to the reduction of the prediction accuracy of the network model. Therefore, it is necessary to reduce the dimension of the original data.
The distribution density A value has a great influence on the performance of RBFNN. From the influence of different A values on the training results, it can be seen that when $A = 3.00$, the prediction error is minimized and the number of neurons activated in the hidden layer is also reduced to the minimum. At this time, the network model structure is the most simplified and the performance of the model is the most optimized. However, when 4.00 or 5.00, the number of neurons in the hidden layer increased and the convergence rate slowed down, but the training error of RBFNN was not significantly improved, which was obviously not the optimal value of A. In general, if the value of A is too small, it is prone to overfitting. The network complexity is high, the network convergence speed is slow, the prediction accuracy is low, and the generalization ability is poor. If a value is too large, there is no overfitting phenomenon, the network complexity increases, the convergence rate decreases significantly, and the prediction accuracy does not improve significantly. To sum up, this paper comprehensively considers the complexity, convergence speed, prediction accuracy and generalization ability of the network model, and finally determines the value of A as 3.00. At this point, the model can take into account the above parameters and achieve the optimal performance.

As shown in Figure 2, the relative error for the contrast experiment, red (error1) represent only take into consideration of the student's result of relative error, black (error2) on behalf of the student performance and behavior prediction relative error information is taken into consideration, can be seen from the diagram, most black is lower than red, only a small amount in more than the range of red, illustrates the student performance and behavior information are taken into consideration the relative prediction error under the condition of the vast majority of less than the relative error of student achievement, the experimental results show that the behavior of the students have a certain influence to the result, It is of great significance to encourage students to go to the library more often, get up early on time and keep good study and living habits. Artificial neural network in this experiment the number of hidden layer upon layer, using a conventional algorithm, namely the number of hidden layer is equal to the number of input layer and output layer divided by the number of 2, according to the experiment requirement, can be divided into two experiments, the first experiment will normally only 10 courses scores as factors that influence the performance when implicit layer set 5; In the second experiment, the scores of 10 courses and the behaviors of 2 students were taken as factors affecting the scores, and the hidden layer was set as 6.
The accuracy of the deep cycle neural network classification algorithm is 93.8%, which is much higher than the traditional prediction algorithm, such as decision tree algorithm, with the accuracy of 80.3% and 80.84%, respectively. Decision tree algorithm is the most commonly used method in data mining classification algorithm. Compared with other classification techniques, decision tree algorithm is simple and easy to understand. Using greed strategy to decision tree variance, finding the right starting point of the tree will affect the final result. In other words, small changes in the early stage will have a big impact later on. If you change the value a little bit, the decision tree will change and be unstable. The classifier we want is robust to noise, so this is not what we expect, there is the problem of low classification accuracy. The deep cycle neural network algorithm USES feedback propagation and has the ability of associating and remembering external stimuli and input information.

4.2. Suggestions for grade prediction models based on deep cycle neural network classification algorithm

Before the experiment, the non-numerical data can be converted into numerical data by numerical method. There are three most commonly used numerical methods: thermal coding and Map function. Singleton encoding applies to smaller features, such as gender, where only "male" and "female" properties are optional, which are converted to "01" and "10" binary Numbers, respectively. The latter two can be thought of as one way to convert a string to a number according to the type of string in the attribute. According to the characteristics of the attributes in this paper, the latter can be selected to numeric the data.

In numerical value, the data are converted into digital data set, but the different attributes, result in have a quite different dimension between the attributes, such as the learning time (study_lastTime) numerical value after "1707191201" and gender larger dimension difference exists between the "1", in order to eliminate the dimension differences, the need for normalization processing data. Normalization processing is carried out on the basis of numeralization, which is used to eliminate the differences caused by different dimensions between data and map all data to the Numbers in the range of [0,1].
With the rapid development of deep learning cycle especially neural network, on the issue of sequence to sequence, encoder and decoder architecture because of its flexible and efficient characteristics has been widely used, almost became a common framework to solve the problem of the class, but, no matter how long the input sequence, the encoder will input data encoded into fixed length of dense vector, provided to the decoder to decode, and fixed length coding vector can contain information is limited, and as a result of the existence of the problem such as gradient disappeared, calculation process of each step will cause the loss of information, As a result, although a relatively reasonable encoding vector can be obtained when the input sequence length is short, the ability of encoding vector to express data characteristics reaches a bottleneck when the sequence length is long, and the model effect is greatly reduced. The proposal of attention mechanism can alleviate this situation to some extent.

5. Conclusion

Classification algorithms, based on the deep circulation neural network will be a university student's result as the predicted sample, constructs the input and output are independent of each other's major input and of the low dimensional vector neural network, and minimize the neural network input vector is the mutual mapping between the interference, simplifies the design and structure of the neural network, which greatly reduce the cost of the neural network and the amount of calculation, effectively improves the calculation speed of the neural network convergence and the accuracy of prediction. The results show that the prediction model system can realize and complete the task of student achievement calculation and prediction well, and has strong theoretical and practical operation.

Acknowledgments

1. This work was supported by the Science and technology project of the Jiangxi Provincial Education Departmen(No.GJJ191095);

2. The Science and technology project of the Jiangxi Provincial Education Departmen(No.GJJ191099).

References

[1] Wendt H , Kasper D , Trendtel M . Assuming measurement invariance of background indicators in international comparative educational achievement studies: a challenge for the interpretation of achievement differences[J]. Large Scale Assessments in Education, 2017, 5(1):1-34.

[2] Xiao Xiang , Wang Laixiang , Xie Hongwen , et al. Study on Railway Locomotive Demand Prediction Model[J]. Management & Engineering, 2015(18):24-28.

[3] Chuang C C , Chen T T , Chen C C . Application of Grey Theory in the Construction of Impact Criteria and Prediction Model of Players' Salary Structure[J]. Mathematical Problems in Engineering, 2018, 2018(PT.2):7365615.1-7365615.9.

[4] Baena-Extremera A , Granero-Gallegos A . Prediction Model of Satisfaction with Physical Education and School[J]. Revista De Psicodidáctica, 2015, 20(1):177-192.
[5] Su, Chung-Ho. A Novel Hybrid Learning Achievement Prediction Model: A Case Study in Gamification Education Applications (APPS)[J]. International Journal of Information Technology & Decision Making, 2017, 16(02):515-543.

[6] Leana-Taclar, Marilena Z . The Actiotope Model of Giftedness: Its Relationship With Motivation, and the Prediction of Academic Achievement Among Turkish Students[J]. The Australian Educational and Developmental Psychologist, 2015, 32(01):41-55.

[7] Valencia F , Collado J , Doris Sáez, et al. Robust Energy Management System for a Microgrid Based on a Fuzzy Prediction Interval Model[J]. Smart Grid, IEEE Transactions on, 2016, 7(3):1486-1494.

[8] Tahir M , Hayat M . Machine learning based identification of protein-protein interactions using derived features of physiochemical properties and evolutionary profiles[J]. Artificial intelligence in medicine, 2017, 78(may):61-71.

[9] Mueller M , Aoki T . New High Performance GPGPU Code Transformation Framework Applied to Large Production Weather Prediction Code[J]. ACM Transactions on Parallel Computing, 2018, 5(2):7.1-7.42.

[10] Malkan U Y , Gunes G , Isik A , et al. Rebound Thrombocytosis following Induction Chemotherapy Is an Independent Predictor of a Good Prognosis in Acute Myeloid Leukemia Patients Attaining First Complete Remission[J]. Acta Haematologica, 2015, 134(1):32-37.