Journal of mathematics and computer science 12 (2014), 263-270

Contents list available at JMCS

Journal of Mathematics and Computer Science

Journal Homepage: www.tjmcs.com

Forecasting Number of Students Applicant for Courses by Artificial Neural Networks

Jafar Pouramini

Department of Computer Engineering and Information Technology, Payame Noor University(PNU), Iran
pouramini@yahoo.com

Article history:
Received July 2014
Accepted August 2014
Available online August 2014

Abstract

Forecasting the number of students who are going to take a special course in next semester in Computer Engineering field at Payam Noor University is the subject. To do this, many neural network structures have been tested with MATLAB software by existing data and were compared to real data, networks like feedforward backpropagation 3 and 4-layared, RBF network, etc. To achieve a network with optimum structure, various parameters and criteria like MAE[1], MSE[2] and MSEREG[3], have been examined. At last, a 3-layered feedback neural network in the form of 20-n-1 was chosen for this problem. Comparing experiential results with real data, it is shown that the obtained model can effectively forecast enrolments of students. So it can be used for forecasting tasks especially when a forecast with high accuracy is needed.

Keywords: Artificial Neural Network, RBF network, Elman network, Hopfield network, Forecasting

1. Introduction

[1] Mean Absolute Error
[2] Mean Square Error
[3] Mean Square Error Regression

Artificial neural networks is a common method of exact forecasting that has interesting forecasting applications in engineering, economic and social problems and stock and gold price[1,2,3,4,5,6,7]. Artificial neural networks are flexible computational and approximation tools which are used in a wide range of time-series forecasting problems with high accuracy [8,9,10]. Different aspects of artificial neural networks have made them very interesting and valuable for forecasting.

One of the difficulties in managing the educational system of universities is to plan the courses. Both lacks of enough courses for applicant students and presenting courses with very few applicants can make problems for students and managers. This shows the necessity of estimating the students for each course with low error and before the courses begin. In this study, some different neural network methods, like RBF, Hopfield, feedforward BP, Elman Bp and MLP have been examined and compared for the problem of estimating the statistic of next semester.

2. Multilayer Perceptron Network (MLP)

Multilayer Perceptron network is a kind of artificial neural networks. This network is composed of an input layer, one or several hidden layers and an output layer. To teach this network, the back propagation (BP) algorithm is usually used. During teaching MLP by BP algorithm, first calculations from input to the output of the network are done and then the calculated error amounts are propagated to previous layers. At first, output calculation is done layer by layer and the output of each layer is the input of the next layer. In back propagation state, there is a desirable value and by that and also by updating rules, one can adjust weights. Although the back propagation algorithm has given many good results in solving most of problems, it is not strong for solving some special problems which can be the result of learning time to be long or not determined or due to choose the inappropriate learning factor or the distribution of weights by chance. Also, in some cases, learning process is disturbed because a local minimum exists and it makes it stop because the answer is located in flat parts of the threshold function.

The learning stages via the algorithm are:

a. To allocate an accidental weight matrix to each of the connections

b. Choosing the input vector and an appropriate output for that

c. To calculate the output of each neuron in every layer and thus to calculate the output of neurons in output layer

d. Updating the weights by method of propagating the network error to previous layers in which the aforesaid error is caused by the difference between real output and calculated output.

e. Evaluating the performance of the learned network by some defined criteria, like mean squared error (MSE), and finally getting back to stage C or to end the process of learning.

3. Feedforward BP Artificial Neural Network

Multilayer feedforward BP neural networks can use differentiable active functions to learn function approximation and to find a pattern between input data set and output data set and also to classify input data in such a way that you define. (Other types of networks can learn well though multilayer network is very useful.)

The term backpropagation is referred to the process by which the derivative of network error with respect to weights and network biases can be calculated. This process can be applied by various

optimization methods. The structure of multilayer network in not fully determined with problem. The number of neurons in input layer is determined by problem. The number of neurons in the output layer is determined by the number of outputs required by problem. But the number of middle layers and its neurons are determined by network designer.

The two layer sigmoid/linear network can present any functional-like relationship between input and output if sigmoid layer has enough neurons. Bias and one-layer sigmoid networks and linear one layer output network are suitable for every function approximation with limited discontinuity.

BP is the standard algorithm for gradient descent. There are many learning BP algorithm. None of them is the best for all situations and they have been examined in this study.

4. Presenting the Problem

Every academic field is composed of a number of courses. Some of them are prerequisite for others. Usually, the department of education en every university gives a program of next semester to students in which it is determined that students should take which courses in which semesters. Payam Noor University also works like this and gives a program to students which is called “9-semester program”. With existing limitations, it is not usually possible to present all courses to students but it is possible that in that very semester, there are some courses which are not taken by many students. Although we have some general information, we don’t know exactly how many students are going to take “advanced programming” next semester.

New students take the course of “Computer Basics” as soon as they get into the university. Experience has shown that each semester there exist some students who fail in this course. Therefore, either they have to take the same course for following semester or leave it for future. Those who pass the course can take advanced programming in the following semester. In addition, among those who take advance programming, some will pass and some won’t. The latter can take this course in the next semester or leave it for future semesters. This complexity is the reason why planners cannot have a good estimation of the students who are going to take advance programming in the next semester and perhaps it leads to some problems in planning the courses.

Forecasting how many students are going to take course z in the next semester depends on several things, including:

1. Number of students who are taking the prerequisite course in this semester; because only students who pass this course can take z

2. Number of students who are taking z in current semester; because students who will fail in this course are likely to take it in the following semester

3. Number of students who are in their first semester

4. For some reasons, some students who can take this course won’t take it and instead, they take another course and will take z in future semesters.

A percentage of students pass Basics (A) and a percentage of students fail Advanced (B) and need to re-take it. Number of all the people who can take this course in the next semester is Ay_n+Bz_n. Though A and B are not exactly known but it depends on students’ educational condition, ability and motivation. Moreover, experience has demonstrated that among people who can take z in the following semester, some won’t take it base on different reasons (C). So, the number of students who are taking Advanced course in the next semester (Z_{n+1}) is as follows:

$$y_{n+1}=x_n+(1-A)y_n \quad (1\text{-}1)$$

$z_{n+1}=C(Ay_n+Bz_n)$ (2-1)

where x_n is the number of new students in semester n , y_n is the number of Basics students in semester n and z_n is the number of Advanced Programming students in semester n. A, B and C are unknown and dynamic factors. A is the percentage of people who have passed z in semester n, B is the percentage of people who failed z in semester n and C is the percentage of people who can take z in semester n+1 and will take it.

5. Data of the Problem

In this study, the number of newcomer students (x_n), Basics students (y_n) and Advanced students (z_n) during 20 consecutive semesters in a Payam Noor University branch have been examined(Fig 1).

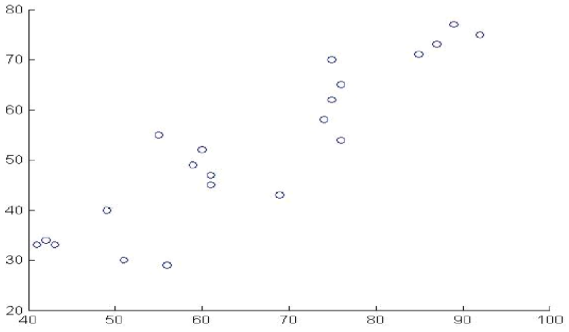

Figure 1. Graph of newcomer students and z and y students.

6. Designing Artificial Neural Network Model for Forecasting the Number of Enrollments

In this study a neural network model for "forecasting the number students in Advanced Programming course in next semester" has been developed using supervised learning. In the process of developing the model the effect of technical variables, number of neurons in input layer, number of hidden neurons and layers, functions of layer activity, different classifications of data for choosing learning set, types of improved learning algorithms and types of networks have been examined, first by discussing the nature of each one and then by doing lots of experiences. We want to find an approximation of it by neural network methods.

The estimation of parameters in a neural network model is not a systematic process [1,4] but it is done via trial-error method and repeating the experiments [7]. In this study, in order to analyze the data and estimate parameters of the model MATLAB software has been used. Every experiment has been repeated 10 times and average of performance criteria are considered as the final result of an experiment [10]. The literature shows most of researchers agree that feedforward neural networks are global approximators and have many applications for function estimation and approximation. But in this study, beside feedforward networks, Elman and GRNN have been tested too. In network learning process, MSE, MAE and MSEREG methods are used as network performance function.

7. Learning and Experiment Data Ratio

Usually, time-series data are divided to two parts, 'learning' and 'experiment'. Learning set, which is composed of 'input- output goal' pairs, is used to teach a special task to network. After learning is stopped, data set of the experiment is used to examine the performance of the network and to see

how much it has learned this special task. Model's answer for data of the experiment is compared with real data. Allocating the data to these two sets are often done according to one of the following ratios: 90%-10%, 80%-20%, 70%-30%.

8. Number of Network Inputs

The most critical decision variable is to determine the number of amounts in order to enter to the network because this vector includes important information about data structure. In general, there is no accepted systematic method to determine the optimized number of inputs[7,9].

After the number of network inputs, number of hidden layers is another critical variable. The correlation factor between network output and real goals is a measure of how much outputs meet the goals. If it is 1, then a complete correlation between output and goals exists. It has been done for 3-layered networks with gradient descent learning command for network in the form of 20-n-1. The results were very discouraging and performance was about 4000.

It has been done with the data and Elman method for 3-layered networks in the form of 20-n-1. The best performances were between 10 and 20. The required number of epochs was 10000.

Learning method	20-5-1	20-6-1	20-8-1	20-9-1
Elman-MSEREG	0.745	1.893	1.163	1.863
Epoch	1008	1032	1008	1107

Table 1. Learning method and Epoch of Elman

It has been done with learning command of Levenberg-Marquaret for 3-layered networks in the form of 20-n-1. The results were very encouraging and the performance was less than 1 and the number of epochs was less than 10.

Learning method	20-5-1	20-8-1	20-14-1	20-20-1
Levenberg-Marquaret MSEREG	0.64	0.465	0.304	0.23
Epoch	12	7	9	6

Table 2. Learning method and Epoch of Levenberg_marquaret

It has been done with learning command of levenberg-marquaret for 3-layered networks in the form of 20-n-1 but we decided the criterion of performance to be MAE. The criterion of performance changed considerably and reached to 10^{-14}. The bests are as follows. The epochs are less than 10.

Learning method	20-5-1	20-8-1	20-10-1	20-13-1	20-15-1	20-19-1
Levenberg-Marquaret MAE	2.007e13	4.17e-11	2.13e-14	1.62e-14	2.67e.14	2.84e-14
Epoch	9	6	6	8	7	8

Table 3. Learning Method and Epoch of Levenberg_marquaret

It has been done with learning command of Levenberg-Marquaret for 3-layered networks in the form of 20-n-1 but we decided the criterion of performance to be MSE.

$$F = mse = \frac{1}{N}\sum_{i=1}^{N}(e_i)^2 = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2$$

The criterion of performance changed considerably and reached to 10^{-27}. The bests are as follows. The epochs are less than 50.

Learning method	20-3-1	20-4-1	20-9-1	20-15-1	20-18-1	20-19-1
Levenberg-Marquaret MSE	1.496e-26	2.398e-28	5.469e-28	5.065e-27	1.427e-22	1.241e-27
Epoch	27	13	7	9	6	6

Table 4. Learning method and Epoch of Levenberg_marquaret

It has been done with learning command of Gradient descent with momentum for 3-layered networks in the form of 20-n-1 but we decided the criterion of performance to be MAE. The criterion of performance is not good and is about 10. The bests are as follows. The epochs are 390 and 1590. It is suitable for forecasting of the data p=[30 51] and the items 20 and 6.

Learning method	20-6-1	20-20-1
Gradient descent with momrntum-MAE	4.130	10.900
Epoch	390	1513

Table 5. Learning method and Epoch of Levenberg_marquaret

However, when the network was evaluated with MSE criterion, the results were very unfavorable. Even when it was examined with MSEREG criterion the results were unfavorable though at some points it gave a few good results.

Learning method	20-5-1	20-6-1	20-8-1	20-9-1
Resilient backpropagation -MAE	0.745	1.893	1.163	1.863
Epoch	1008	1032	1008	1107

Table 6. Learning method and Epoch of Resilient Backpropagation

It has been done with learning command of Resilient backpropagation for 3-layered networks in the form of 20-n-1 but we decided the criterion of performance to be MAE. The criterion of performance changed considerably and reached less than 1. The bests are as follows. The epochs are more than

1000. It is suitable for forecasting of the data p=[30 51] and the items 5, 6, 8, 9, with low performance and close approximation.

It has been done with the data and Feedforward BP method with learning command of gradient descent for networks in the form of 3-n-1. All of them had bad performance and their performance was more than 10300 even with number of epochs more than 1000.

RBF networks are used. The approximation accuracy is very high and near 10-29 with epoch number of 20. Although it has approximated the given points well but can it approximate well the experiment points. In other words, can it perform forecast as well as performing the interpolation? Its approximation is not that good, with relative error of ~33%.

The Levenberg- Marquardt Algorithm was used to update the artificial neural network weights. It is a high usage algorithm because it performs network learning very fast and minimizes the level of error. In fact, this algorithm was designed to increase learning speed of network and is based on Hesin matrix.

9. Conclusion

In this study, forecasting the number of students who are going to take a special course in next semester in Computer Engineering field at Payam Noor University has been examined as a problem.

With supervised learning, a neural network model for “forecasting the number of students in Advanced Programming course in next semester” was developed. In the process of developing the model the effect of technical variables, number of neurons in input layer, number of hidden neurons and layers, functions of layer activity, different classifications of data for choosing learning set, types of improved learning algorithms and types of networks have been examined, first by discussing the nature of each one and then by doing lots of experiences. Finally, a feedback 3-layered neural network in the form of 20-n-1 with the learning command of levenberg-marquaret and performance criterion of MSE near to 10^{-27} and performance criterion of MAE near to 10^{-14} were chosen as the best models for forecasting the results.

10. Acknowledgment

This research is supported by a research grant from Payam Noor University.

References

[1] Saeed Ayat, Zabihollah Ahmad Pour, Comparison between artificial neural network learning

algorithms for prediction of student average considering effective factors in Learning and educational progress, Journal of mathematics and computer science 8 (2014), 215 - 225

[2] Ali Ghezelbash, Predicting changes in stock index and gold prices to neural network approach, The Journal of mathematics and computer science Vol. 4 No. 2(2012)227-236

[3] Mahnaz Bagheri, The bankruptcy prediction in Tehran share holding using neural and it’s comparision with logistic regression, The Journal of mathematics and computer science Vol. 5 No. 3(2012)219-228

[4] Mehdi Sotoudeh,Elahe Farshad, Application of neural network for forecasting gas price in America, The Journal of mathematics and computer science Vol.4 No.2 (2012)216-226

[5] Mehdi Sotoudeh1, Elahe Farshad2 ,Application of neural network for forecasting gas price in America, Journal of mathematics and computer science Vol. 4 No. 2 (2012) 216- 226

[6] k. Abhishek, M. P. Singh, S. Ghosh, A. Anand, Weather forecasting model using artificial neural network, procedia technology4, (2012)311-318.
[7] Qi, M., & Zhang, G.P., An Investigation of model selection criteria for neural network time series forecasting, european journal of operational research, Vol. 132, 2001, pp. 666-680.

[8] Ufuk Yolcu, Erol Egrioglu, Cagdas H. Aladag, A new linear & nonlinear artificial neural network model for time series forecasting, Decision Support Systems, Volume 54, Issue 3, February 2013, Pages 1340-1347
[9] Juan Peralta Donate, Paulo Cortez, Germán Gutiérrez Sánchez, Araceli Sanchis de Miguel, Time series forecasting using a weighted cross-validation evolutionary artificial neural network ensemble, Neurocomputing, Volume 109, 3 June 2013, Pages 27-32
[10] Kaastra, I., & Boyd, M., Designing a neural network for forecasting financial and economic time series, neurocomputing, Vol. 10, 1996, pp. 215-236.
[11] Azoff, E., "Neural network time series forecasting of financial markets", John Wiley & Sons , 1994.