

## Comparison between Artificial Neural Network Learning Algorithms for Prediction of Student Average considering Effective Factors in Learning and Educational Progress

Saeed Ayat

Department of Computer Engineering and  
Information Technology, Payame Noor  
University, Iran  
[dr.ayat@pnu.ac.ir](mailto:dr.ayat@pnu.ac.ir)

Zabihollah Ahmad Pour

Department of Science, Islamic Azad  
University – Ayatollah Amoli Branch, Iran  
[z.ahmadpour@iauamol.ac.ir](mailto:z.ahmadpour@iauamol.ac.ir)

### Article history:

Received June 2013

Accepted July 2013

Available online July 2013

### Abstract

In this project, by using different learning algorithms in the form of 37 input parameters of network for predicting average considering effective factors in learning and educational progress, the Perceptron artificial neural network have been studied.

The requisite data have been obtained through handing out questionnaires between 400 students of Payame Noor University majoring in computer engineering, information technology and computer science.

For recognizing the best learning algorithm, 13 common algorithms considering factors such as training time, the percentage of accountability, the index of efficiency ( the mean squared errors), and the number of epoch have been studied after error propagation. Finally the LM algorithm was recognized as the best learning algorithm for prediction of average.

**Keywords:** learning algorithm, artificial neural network, MLP, prediction of average.

### 1. Introduction

The type of Perceptron artificial neural network is feed forward that is one of the most practical models of artificial neural network in prediction and estimation.

Considering this point that using the artificial neural network method in research subjects in behaving science is increasing and also few researches about improving abilities of neural network surrounding psychology especially in training psychology have been done, this article attempts to create different neural networks in term of learning algorithm with the purpose of prediction of average based on risks of factors and to prepare conditions for comparing efficiency of learning algorithms after measuring standards of efficiency of every network. With this purpose, first the effective factors in learning and educational progress have been discussed. Also, the general structure of networks, the diagram of training time, the index of efficiency, percentage of accountability, and the number of training epochs have been drawn after expressing the method for collecting numbers and choosing statistic society.

## 2. Effective factors in learning and educational progress

Learning is the process of bringing rather stable changes in an individual's behavior created based on experience [9].

Learning is a process. In every process, different factors and variants interact. The type and extremity of interaction are followed by different changes. Studying all effective factors in the process of teaching is not possible. Some of the most important factors that have an evident affect on the procedure of teaching include motivation, purpose, the situation and the environment for learning, the method of teaching, intelligence, encouragement and testimonial.

**Motivation:** everyone has a tendency towards the activity he or she likes. So it's natural for human to like some things and hate some others. Motivation can be the heart of a healthy bringing up related program in society. Motivation is necessary for learning and can also cause the main behavior that is the major condition and factor in learning. Motivation is necessary not only for the start of learning, but for continuing and improving it and solving the problems created and using learnt materials in a new situation [10].

**Purpose:** Purpose gives energy and direction to human's activity. If a human doesn't have any accessible purpose in his life, he will lose his dynamism. A valuable purpose persuades human to wish and want some things and also creates required energy for activity in him and help him to have determination and make decisions.

Purpose is one of the effective and basic conditions that persuade students into learning. When the purpose of learning is a certain instructional material, the interest in learning will increase [11].

**The situation and environment of learning:** The situation of learning and its environment is one the effective factors in learning. Environment can be physical such as light, weather, and educational facilities and equipments. Naturally, the more facilities are provided for an individual, the more learning happens. Learning will be so different in educational places that possess suitable space, library, and different scientific sources. Based on learning psychological findings, students learn educational materials better and easier by their sight and using different tools, and this is because educational facilities make education concreter, more practical, and more pleasant as a result of activating students' different senses [12].

Environment can be emotional. An arranged educational situation accompanied by affection and mutual respect affects learning more in comparison with environments lacking emotions. Emotion as a very important factor can be effective in learning process.

Anyway, the facilities of educational environment including human forces and equipments, social, cultural, and economical situation, the family's attitude towards education and school, and thousands of other environmental factors can affect the quantity and quality of learning.

**Method of teaching:** Human forces, especially professors are definitely one of the most important factors in constituting educational environment [13]. Professional growth of teachers is the most important factor affecting any educational system. If professors know themselves as guides and creators of desired condition for learning and teach students methods of gaining experience rather than transfer information, students will be more active in dealing with problems, use different sources, obtain required data, and organize and then analyze them to solve problems. By using this method, students will not only learn scientific facts, but be familiar with scientific ways of obtaining knowledge, learn how to work with different sources and how to use educational materials to solve problems and act their creativity in learning process. Various factors such as used teaching patterns of teachers play a role to make teaching effective and fulfill learning by teaching. Teaching patterns affect educational progress more in comparison with traditional teaching methods (lectures) [14].

**Intelligence:** Intelligence is one of the other effective factors in learning. There isn't still a comprehensive and acceptable definition of intelligence for all psychologists, though much research has been done on intelligence and its importance in learning materials [11].

**Encouragement and testimonial:** in stimuli-reaction theories, when there is a reinforcing factor or testimonial, connection is made between stimuli and reaction. This testimonial is mostly eats for children and animals and other things for the adult. In this kind of theories, if the reaction made as a result of stimuli is not reinforced by a testimonial or a reinforcing factor, it will be lost soon and it's not reasonable to expect a behavior [11]. So encouragement and testimonial can be an effective factor in learning.

### 3. Artificial neural networks

Artificial neural network is a model for mathematics process [1]. This network was first created by formulizing the abilities of human brain by Warren Mack Cloth and Walter Pits in 1943[2]. This network is basically a kind of parallel and massive distributing processor constituted of simple process units [16]. The problem with the prior artificial neural network was lacking capability in solving linear problems. The researchers of artificial neural networks presented the model of multilayer Perceptron network to resolve these problems [3]. The method of artificial neural network is a new method in solving complicated problems in which there isn't any certain algorithm for solving them or there are traditional long solutions [16]. Perceptron neural networks have capability to learn through empirical knowledge presented in its own inter communicative units and can use the learnt materials [15]. These models are used not only for studying available phenomenon, but for predicting them in different times [2]. Neurons are the smallest units of artificial neural networks. Every Perceptron neural network is constituted of one input layer, one output layer, and one or more hidden layers. The neurons of every layer are joined to the next layer neurons by some weights [17]. Through the process of training network, these weights and some stable values so called "Bias" are added to them change continuously in order to reach minimum error that means the root of mean square error is reduced to minimum. Furthermore, transfer functions are used to transfer output of every layer to the next layers [4].

#### 3.1. Effective factors in training neural networks

The purpose of training a network is to minimize errors among output and the reference values [1]. Success in training artificial neural networks based on the type of problem in the first step depends on model architecture (the number of neurons of output and input layers, the number of layers and the neurons of the hidden layer) and in the second step depends on used algorithm for training network [5]. Every neuron consists of some inputs. The purpose of input layer is transferring the received values to the neurons of the hidden layer. The number of inputs of available neurons in the hidden layers is equal to the number of outputs of the previous layer [6].

**3.1.1. The number of neurons of the input and output layers:** The number of neurons in input and output layers is mostly determined by problem in using artificial neural networks in different problems [18]. The large number of neurons in input and output layers makes the size of model larger and may jeopardize the learning quality and time [5]. In sciences like behavior science that various factors have a role in a behavior, it's usually difficult to limit inputs of neural network. Furthermore, it's difficult to not consider the effect of some factors in a special behavior just for a better learning by network. But it's possible to improve the learning quality by reducing output variants.

**3.1.2. The number of hidden layers and their neurons:** Hidden layers have to be in low numbers as much as possible. It's been proved that every function can be approximated to at most three hidden layers. First, the network is trained by a hidden layer and then increases the number of hidden layers in case of unsuitable function.

For a good-sized neural network, the number of hidden neurons is chosen with respect to small proportion of the number of inputs. If the Perceptron multilayer network doesn't converge into the desired account, the number of neurons of hidden layer will be increased. If the network converges and enjoys good extension, if possible, it will train a less number of hidden neurons. An exact choice of layers and hidden neurons is not possible for various reasons such as uncertain and casual nature of many learning procedures and also freedom of systems under investigation and so hidden layer is mostly measured empirically [18].

**3.1.3. The algorithm of training artificial neural network:** The method of learning Perceptron multilayer network happens after error propagation, because output error of the model is returned to back and inside of the system in order that resumption and arrangement of weights happen [19]. This algorithm has been improved along time by researchers' attempt [3]. And various kinds of algorithms were presented for instructing network after error propagation. In using artificial neural networks, the choice of suitable learning algorithm is done based on mental inner judgments and the methods of examination and error. Using one algorithm of learning for increasing accuracy of designing model is what is highlighted in using artificial neural network in sciences like behavior science [20]. There are 12 algorithms for training BP in Matlab toolbox that are: br, lm, oss, bfg, scg, cgb, cgp, cgf, rp, gdx, gdm, gd.

### **3.2. Transfer functions**

Neurons can use different stimulus functions for producing output. The most common of them are log-sigmoid, tan-sigmoid, and pure-line functions [7].

## **4. The method of collecting data and statistic society**

The requisite data were obtained through handing out 400 standard questionnaires consisting of effective factors in learning. The statistic society has been chosen from students of the Payamenour University of Isfahan and Najaf Abad majoring in computer engineering, information data, and computer science. The sample consisted of both genders including 117 male and 283 female.

## **5. Neural networks architecture under investigation**

A network must have at least input, hidden, and output layers after error propagation. For choosing the number of hidden layers and their neurons properly, experiments are needed [8].

One empirical way for determining this number is considering proportion of the number of inputs of network. If the network doesn't converge to a desired account, the number of neurons will increase. And if the network converges to the same number of neurons and enjoys potential for extension, if possible, the less number of neurons will be experimented [18]. In this project, the model of Perceptron neural network was used for implementation of networks. The number of neurons considered in the first and second layers are 20 and 13 respectively. The type of stimulus functions or transferring hidden layers is tan-sigmoid.

All used data in input layer of artificial neural network were risk of effective factors in determining the extent of educational progress by predicting average. So, 37 parameters were considered as input of the network.

The purpose of implementation of neural networks in the present research has been predicting educational progress based on 37 parameters. The standards of educational progress of students are their scores in different courses and their averages. So the average of student has been used for measuring this extent. Stimulus function or transferring input layer is regarded as pure line function.

## 6. Results

The following presented results have been obtained by the Matlab software, R2010a version, using a system with core i5 2.20 GHz processor, and 2GB memory under XP SP3 operating system. It's necessary to mention that biases and prior weights assumed equal for all networks. And the procedure of accountability was studied based on the reference value. In fact, this project, contrary to other done researches, studies the effective indexes related to every network per Bias and equal weights and defaults of the software Matlab in various training courses, not indexes such as training time, MSE, the number of Epoch, and the percentage of accountability. In every educational course, Biases and weights are calculated based on used algorithm, and in the next course these Biases and weights are considered as prior Biases and weights to examine if improvement in training happens in every course or not.

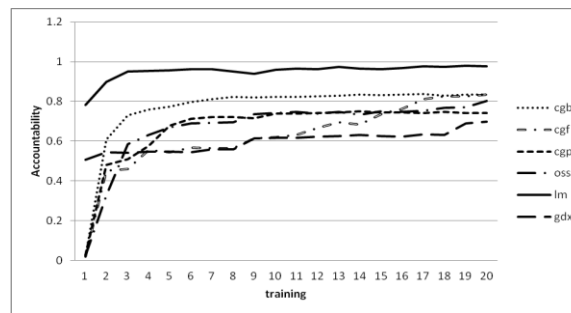


Figure 1. Comparing 6 training algorithms of MLP with respect to percentage of accountability

First 12 learning algorithms of post-propagation type, in the form of network, with 37-20-13-1 architecture, with 400 data, and 70 courses, were trained. Just 6 algorithms presented an acceptable percentage of accountability. The percentage of accountability has been showed through Regression diagram.

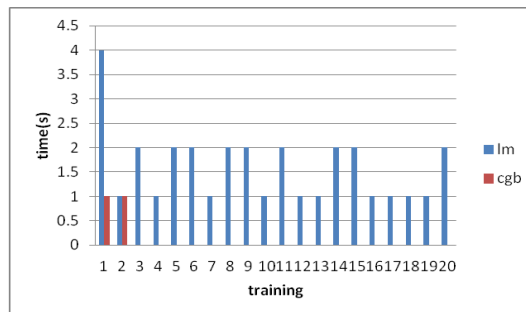


Figure 2. Diagram of training time of lm and cgb algorithms

Then, algorithms were trained in 3 stages, in 20 courses, for recognizing the best learning algorithm. The mean of the percentage of accountability that is explained in chart 1 was obtained.

Figure 1 shows the diagram of the mean of percentage of accountability by these 6 algorithms. This figure shows that cgb and lm as two algorithms have better efficiency in reaction to the given training with respect to the other four algorithms.

In the last stage, for recognizing the best algorithm between lm and cgb, some information about training time (the duration of time in which training ends by one of the ending conditions), index for efficiency of network (the most important parameter for evaluating network; the more the index approaches to zero, the more training is desirable), the percentage of accountability (representing the extent of accountability of network for the collected data), and the number of steps (representing the number of repeating training samples; the more it is small, the more it's desirable) were studied to reach a suitable index of efficiency between these two algorithms presented in table 2.

The figures 2, 3, 4 and 5 show training time, index of efficiency, the percentage of accountability, and the number of epochs respectively in the process of network related training.

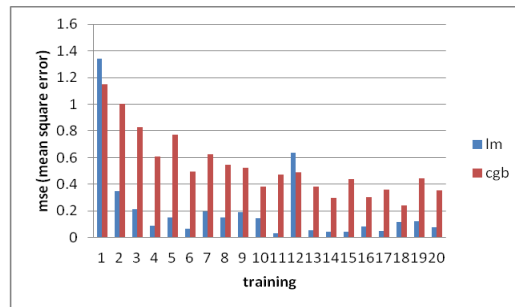


Figure 3. Diagram of performance index of lm cgb algorithms

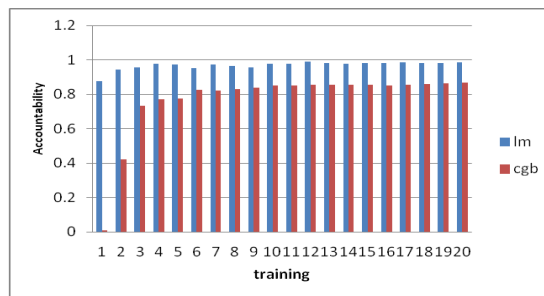


Figure 4. Diagram of response time of lm and cgb algorithms

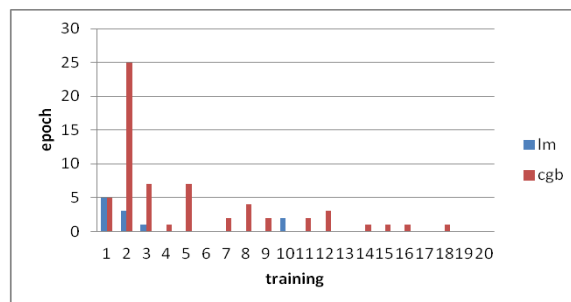


Figure 5. Diagram of number of epoch in training network

Table 1. Average percentage of responsibility of Algorithms in 20 courses

Algorithm Training course	Gdx	lm	oss	cgb	cgf	cgb
1	0.022639	0.019799	0.026911	0.024467	0.783477	0.50714
2	0.60806	0.450523	0.481157	0.326317	0.89875	0.544407
3	0.729507	0.4614	0.50894	0.58615	0.95117	0.540157
4	0.75784	0.55533	0.576427	0.632017	0.952877	0.548197
5	0.772377	0.542993	0.677197	0.670437	0.95633	0.54628
6	0.797353	0.566693	0.712467	0.688587	0.9613	0.542873
7	0.811523	0.563597	0.722597	0.692243	0.96288	0.55967
8	0.822403	0.564433	0.720373	0.694177	0.95156	0.559827
9	0.819113	0.61417	0.716627	0.73542	0.938973	0.614347
10	0.822967	0.61861	0.740027	0.74035	0.958833	0.616233
11	0.823813	0.629913	0.737977	0.748197	0.964817	0.61823
12	0.82707	0.665437	0.741827	0.737993	0.962823	0.62222
13	0.82969	0.693937	0.743837	0.748473	0.97313	0.625697
14	0.83389	0.684897	0.750567	0.734077	0.965907	0.631773
15	0.83049	0.736833	0.745763	0.74951	0.962923	0.62445
16	0.834697	0.763023	0.74495	0.748113	0.96828	0.623917
17	0.83653	0.81179	0.740927	0.75316	0.977213	0.632883



18	0.82771	0.826983	0.74635	0.76738	0.97528	0.632537
19	0.83471	0.825573	0.742907	0.77135	0.981067	0.68968
20	0.835877	0.834333	0.74288	0.802397	0.976007	0.698033

Table 2: Values of time, number of epoch, accountability and mse of lm and cgb algorithm

Algorithm Training course	TIME		Accountability		EPOCH		MSE	
	lm	cgb	lm	cgb	lm	cgb	lm	cgb
1	4	1	0.87457	0.011375	5	5	1.3388	1.1487
2	1	1	0.94523	0.42225	3	25	0.34997	1.0021
3	2	0	0.95734	0.73268	1	7	0.21274	0.82576
4	1	0	0.97856	0.76982	0	1	0.09044	0.60843
5	2	0	0.97223	0.7745	0	7	0.15233	0.77177
6	2	0	0.95275	0.8269	0	0	0.069041	0.4949
7	1	0	0.9743	0.82218	0	2	0.197	0.62433
8	2	0	0.96344	0.83174	0	4	0.15254	0.54557
9	2	0	0.95844	0.84049	0	2	0.1901	0.52135
10	1	0	0.97915	0.85009	2	0	0.14432	0.37963
11	2	0	0.97596	0.85068	0	2	0.032024	0.47296
12	1	0	0.98872	0.85758	0	3	0.63695	0.49223
13	1	0	0.98036	0.85395	0	0	0.05352	0.38408
14	2	0	0.97629	0.85445	0	1	0.045709	0.299



15	2	0	0.98326	0.85607	0	1	0.043639	0.43599
16	1	0	0.9836	0.84943	0	1	0.084147	0.30275
17	1	0	0.9848	0.8567	0	0	0.051243	0.36203
18	1	0	0.97994	0.85797	0	1	0.11524	0.23899
19	1	0	0.98051	0.86379	0	0	0.12526	0.44646
20	2	0	0.98448	0.86614	0	0	0.075678	0.35288

Finally, two algorithms that selected with training time, index of network performance, number of epoch and percentage of accountability are evaluated and compared. With respect to figure 2, training time of lm algorithm is more than cgb algorithm in every course. Figure 3 shows that index of network performance in lm algorithm is better than cgb algorithm. In fact, mean square error of network related to lm algorithm is more converge to zero. With respect to figure 4 results that lm algorithm comaried to cgb algorithm has better response time. Furthermore, lm algorithm has more convergence to target function. Figure 5 shows number of epoch in every course. Approximately, lm algorithm represent less epoch than cgb algorithm in every course.

By simulating network with input data, simulated output resulted. If figure 6 and 7 considered, we conclude that lm algorithm present the most convergence to target function and represent better predict of average.

## 7. Discussion and conclusion

12 algorithms concerned with Perceptron multilayer neural networks were studied. 6 algorithms were omitted as a result of unsuitable efficiency regarding time and the extent of percentage of accountability during 70 training courses. Just 6 algorithms lm, oss, cgb, cgp, cgf, and gdx were studied and compared again. The algorithms were trained in 3 stages and in 20 courses. According to figure 1, lm and cgb showed the most accountability to training. Cgf had the best convergence as the number of training courses increased. Gdx showed the least extent of convergence or the least percentage of accountability. Gdx, cgp, cgb, and cgf didn't present a suitable percentage of accountability in the first training course, but presented a better efficiency after the second course. All of the algorithms had almost stable procedures after the third training course. Sensible changes are not seen in accountability to training after this course. Gdx and cgp had similar behaviors to training courses.

Finally two chosen algorithms were evaluated and compared by use of 4 indexes including the training time, the index of the efficiency of network, the number of epochs, and the percentage of accountability. According to figure 2, in every course training time of lm is more than cgb. The figure 3 represents that the index of the efficiency of network (mean square errors of network) of lm is better than of cgb. In fact the mean square errors of network related to lm approaches more to zero. It's concluded from figure 4 that lm enjoys a better percentage of accountability with respect to cgb. In other

words, lm has a better convergence to objective function. Figure 5 shows the number of epochs in every training course in which lm has almost less epochs with respect to cgb. The simulated output is obtained by simulating networks by input data. According to figures 6 and 7 it can be concluded that lm has the most convergence to objective function and predicts average better.

## Acknowledgment

This research is supported by a research grant from Payame Noor University, the authors also wish to thank Monireh Ghasemi, a student from Najafabad Payame Noor University who help during this research.

## Reference

- [1] M. R. Mustafa, R. Rezaur, S. Saiedi, M. He. Isa, River Suspended Sediment Prediction Using Various Multilayer Perceptron Neural Network Training Algorithms\_A case Study Malaysia, *Water Resour manage.* 26, (2012)1879-1897.
- [2] N. I. Daliakopoulos, P. Couliblay, H. K. Tsanis, Ground water level forecasting using artificial neural networks, *J. Hydrol.*, 309, (2005) 229-240.
- [3] M. T. Hagan, H. B. Demuth, M. H. Beal, *Neural Network Design*, PWS, Boston, 2005.
- [4] B. D. Ripley, R. M. Ripley, Neural networks as statistical methods in survival analysis, *Journal of statistics*, (2006) 409-456.
- [5] G. L. Dee, N. Bakhary, A. Abdul Rahman, B. Hisham Ahmad, A Comparison of Artificial Neural Network Learning Algorithms for Vibration- Based Damage Detection, *Journal of Advanced Materials Research.* 163-167, (2011)2756-60.
- [6] A. Fbrizzio, L. Edna, D. Christian, A. Gilberto, C. Antonio, V. Paschoarelli, Recursive diameter Prediction And Volume Calculation of Eucalyptus Trees Using Multilayer Perceptron Networks, *Applied Soft Computing*12, (2012) 2030-2039.
- [7] H. Demuth, M. Beale, *Neural Network Toolbox User's Guide*, Copyright 1992-2002, By The Math Works. Inc. Version4, (2000)480.
- [8] k. Abhishek, M. P. Singh, S. Ghosh, A. Anand, Weather Forecasting Model Using Artificial Neural Network, *Procedia Technology*4, (2012)311-318.
- [9] A. Lakdashti, R. Yousefi, Kh. Khatiri, Effect of Simulated Education Software on Student Learning and Compare it With Traditional Method of Teaching, *Journal of Information and communication technology in education*, No. 1, Vol 3, (1390) 5-21, in Persian.
- [10] A. Akbari, *Methods of Study*, Faiz Kashani, Tehran, 1382, in Persian.
- [11] M. Parsa, *Psychology of Learning*, Fourth Edition, Payame Noor, Tehrn, 1385, in Persian.
- [12] S. Sattari, E. Jaafar Nezhad, Factors affecting the utilization of visual aids in the teaching - learning from the perspectives of teachers in Mazandaran, *Journal of Information and communication technology in education*, No. 1, Vol 2, (1389) 6-20, in Persian.
- [13] E. Khalkhali, Z. Shakibae, M. Andesh, Meta-analysis of the effects of ICT on professional development of teachers, *Journal of Information and communication technology in education*, No. 1, Vol. 3, (1390) 165-183, in Persian.
- [14] M. Behrangi, *Teaching models*, Kamal Tarbiat, Tehran, 1387, in Persian.
- [15] M. Kantardzic, *Data Mining: Concepts, Models, Methods and Algorithms*, Second Edition, John Wiley, 2011.
- [16] M. Akbari, A. Ranaee, H. Namaghi, Assessment of desertification sensitivity parameters using artificial neural network, *Journal of Soil and Water (Agricultural Sciences and Technology)*, No. 2, Vol. 25, (1390) 398-410, in Persian.
- [17] T. Sabzevari, R. Mohammad Pour, M. E. Zemorrodian, Evaluation parameters of earth dam failures using neural networks, *First National Congress on Civil Engineering*, Sharif University, 12-14 May, 2003, in Persian.

- [18] M. B. Menhaj, Foundations of Neural Networks, Vol. 1, Amir Kabir University, Tehran, 1387, in Persian.
- [19] S. Shokr Beygi, A. Afkari Sayyah, H. Shayeghi Moghanloo, E. Shokouhian, Effects of temperature and Cushioned surface on the apple contusion volume Created by Tension and Predict it using artificial neural network, Journal of Food Science and Technology, No. 8, Vol. 31, (1390)85-94, in Persian.
- [20] A. Ali Pour, S. Kalantarian, The relationship between handedness with academic achievement of middle school students, Journal of School Psychology, No. 1, Vol. 1, (1391) 7-26, in Persian.