Controlling Cost and Time of Construction Projects Using Neural Network

Li Ping Lo

Faculty of Computer Science and Engineering Beijing University China

Abstract

In order to achieve optimized management, there is need to define a criterion in the projects, through which one can calculate cost and time of the project, so that through comparing the criteria in every project one can make decision about desirability and undesirability of the project due to optimized conditions. However, the important issue in proper management is claiming a model for prediction of amount of claims that may be existed in the project due to presence or lack of presence of every factors of appearance of these claims, so that one can have certain amount of probable criterion of claim in a specific project. At the present study, based on characteristics of different cases of projects based on time and cost and other criteria, suitable formulation between time and cost ratio would be presented using classic methods (regression) and smart methods (artificial neural networks), so that the model can be applied in next projects based on previous information and optimized status can be selected.

Key words: construction project management, minimizing completion time of projects, minimizing completion cost of project, artificial neural networks'

Introduction

One of the most important issues in management of construction projects is choosing the best possible option for conducting every activity in order to establish the projects, so that final cost and time of project include the lowest amount. Hence, the issue includes two objectives of minimizing completion time and minimizing completion cost of the project, which should be optimized simultaneously. On the other hand, in real construction projects, usually predicted costs for these activities are along with uncertainty, which can result in many changes in final cost of the project. Main issues that currently constructional projects of Iran are faced to and can result in delay of project completion time can be as follows: lack of proper recognition of capacity of activities of projects; inaccurate estimations of required capacity of resources; weakness in programming cost and budget of the project; weakness in controlling cost performance of the project,; and in general, weakness in management system of project's cost. Clearly, going along the time and using modern and updated technologies of the world can guarantee success in optimal management of cost of projects.

Using available empirical information and using intelligent and classic expanded methods through combining input and output data, characteristics of data relationships can be explored. Intelligent networks are nonlinear and can be applied in complicated relations between input and output variables of a system through using no basic knowledge. As a result, the models have been evaluated by regression analyses.

Success indicators of the project

Success in the project is depended on providing suitable balance among three important factors including completion time, applied resources, and results of work in regard with presenting an adequate level of services to the customers. Project tries to provide customer satisfaction and hence, three important factors should be considered and suitable balance should be provided among them. The individuals tend to complete their project in the predicted time, to observe results of required characteristics and the results can be exactly what they wanted, and finally the project can have desirable final price, which is the price that has been predicted and has been desired. The mentioned three factors present three basic factors of time, quality, and cost, which a balance is existed among them. In other words, when one of them is changed, other would be also changed following it. For example, in order

to achieve a very high quality, costs (and in some cases time) would be enhanced. In order to reduce implementation time of the project, quality should be decreased, or costs would be increased. In addition, it is clear that reduction of cost can result in reduction of quality and increase in implementation time of the project.
The ratio among the three factors is not linear and sometimes it is possible that reduction of half of implementation time can lead to enhancement of cost of using resources to 4times. Decision makers, supporters, and owners of the project should before beginning of implementation time of project make decision about the ratio among the three factors and provide desirable proportion among them. Most of the time it is assumed that only costs of construction period should be considered; although some additional factors are required and have key role in completion time of the project and induced costs by its delay time including considering incomes and application costs in diagram of cash flow. Thus, if the target is saving costs, it is necessary to realize amount of cost reduction in all lifecycle of the product, not just reduction of costs of construction phase. If the results are implemented in a well-designed feasibility study and indicate that the project should be progressed, one can move to programming and implementing phase confidently. If the results are disappointing, the information should be applied for redesigning the project and conducting another feasibility study.

Moreover, while implementing a project, management of the project needs cooperation of authority of cost controlling in order to estimate cost of a project, so that he can evaluate announced prices by different constructors based on latest information of the market and project budget and help project manager to make desirable decision. Presence of proposed database is a suitable instrument for the project in order to estimate the mentioned costs while implementation of the project. If one considers all responsibilities related to supply, leadership, guidance, organization, backup, controlling, and programming as components of the project, then successful project management would certainly lead to success of the project and the only remained factor that can affect success of the project is environmental factor. Based on desirable definition of the success, it can be defined as realization of objectives in the domain of sciences. Not only realization of these objectives, but also implementation of them in presented framework based on the desired standard is necessary for the success of the project.

Budget and cost estimation

Budget is the first limitation of large and small projects. Even in cultural projects, in which individuals work for no wage, shortage of workforce and credit for requirements can be considered as limitations too. Project budgeting is one of the most important responsibilities of project programmers and also quality of supplying credits for realizing project targets is important responsibility of project managers. Specifically, in large projects, in addition to supplying credits for activities, organization of activities and determining their prerequisites are effective for efficient completion of the project based on probable time of allocation of financial credits. For example, it is impossible to complete whole infrastructure of a highway during a single year and wait for the budget in the next year without providing credit for its asphalt coverage. In fact, snowfall and rainfall can have destructive effect on the infrastructure, so that a part of the budget should be consumed on renewal and reconstruction of the infrastructure. In this case, the highway can be completed in separated parts and these parts should be assembled gradually. In order to phase the work and determine optimal size of every section, budget and how to supply it are the most important factors. Managers should be informed exactly about gaining budget and quality of financial commitment of suppliers and provide them for programmers and follow budget allocation from the resources during implementation of the project.

Therefore, estimation of costs in two sections including before making contract in order to provided proposed price and finally project budget after making contract and while implementation of the project is required in

order to control project costs with approved budget. In this regard, integration of cost estimation process in different sections is necessary and for this purpose, suggestion of providing a database has been presented, which it can not only provide updated information about cost rates, but also it can provide a series of basic information and learned things from previously implemented projects in the company, which using them in next projects can be useful.

One if the methods of enhancement of profitability is using pedagogical experiences and studies out of the project in order to improve project management. The first step is reviewing existed documents related to project difficulties, which has remained from previous projects. The next step is expansion of framework of knowledge and science of individuals. In fact, if one can't organize previous information for others in order to use it, in fact the company has not gained benefit of previously learned things.

Artificial neural network

Artificial neural network is a modern name in engineering sciences, which initially has been introduced to the world in 1962 by Frank Rosen Blat and it was seriously introduced in 1986 by Rommel Hart and McKland through innovating improved model of Perceptron. The method uses a neuron and intelligent structure with suitable benchmarking of existed neurons in human's brain and tries to simulate neurons of the brain through different defined mathematical functions of intercellular behavior. The model also tries to model synaptic performance in the natural neurons through calculative weights of communicative lines of the artificial neurons. Empirical nature of this method makes it applicable in some issues such as prediction process, which such attitude is observable in its structure and has a nonlinear and non-integrated behavior.

Basic concepts in artificial neural networks

A biologic neuron can produce output through collecting inputs through dendrites with a specific synaptic weight imposed on the neuron after achieving a certain level. The certain level that is same threshold level is in fact factor of neuron's activity or inactivity.

According to above, it could be mentioned that three factors should be considered in order to model a biologic neuron artificially:

- Whether the neuron is active or inactive
- The output is just depended on inputs of neuron
- Inputs should achieve to the level that output can be produced. [1]

Simple perceptron neural network

Frank Rosen Blat has presented the perceptron simply through connecting these neurons to each other. In fact, he has simulated and analyzed the model formally for the first time in digital computers. [1]

Multiplayer perceptron (MLP) neural network

In many complicated mathematical problems that can result in solving challenging equations, a MLP neural network can be easily used through defining desirable weights and functions. Different activity functions would be applied in neurons based on modes of the problem. In this kind of networks, an input layer would be applied in order to apply inputs of the problem. Also, a hidden layer and an output layer would be applied, which can finally present answers of the problem.

The nodes on the input layer are sensory neurons and nodes of output layer are responding neurons. Hidden neurons are also existed on the hidden layer. [2]

Training this kind of networks would be conducted usually through Back-propagation 1 Method. A sample of a MLP neural network has been presented in figure 1.

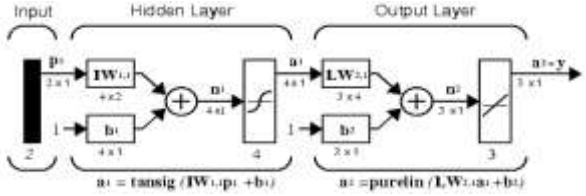

Figure 1: structure of MLP with hidden neurons and output neurons with linear function

MLP perceptron networks can be applied with any number of layers; although the acceptable issue is that a 3-layer perceptron network is able o separate every kind of space. This theory that is known as Kolmogrov Theorem presents an important issue, which is applicable in construction of neural networks [1].

A specific type of MLP networks is known as Single Layer Perceptron (SLP). This kind of network has been formed of an input and an output layer.

In order to choose the most desirable number of neurons, usually the number of hidden neurons is started with a neuron in the hidden layer and continues to maximum number of neurons (12). Marcorat-Loenberg Algorithm needs lower number of neurons than other algorithms because of high speed of convergence. The number of required neurons would be decreased along with enhancement of input data to the network and learning input and output relations in a simple structure.

Other advantages of neural networks

Neural networks with their considerable capacity for deduction of concepts from ambiguous data can be applied for exploration of patterns and identification of those methods that gaining information about them is so hard to do for human and computerized techniques. A trained neural network can be considered as an expert in the domain of information given to it for calculation and analysis. This expert can be applied in order to estimate new desired situations and answer of questions starting with "What if…"

- Adaptive learning: capability of learning the method of performing tasks based on given data for preliminary practices and experiences
- Self-organization: an ANN can provide its own organization or presentation for the information received during learning process. Real time performance: real time calculations of ANN can be done in parallel form and also it has produced specific hardware that can apply the capability using the mentioned instruments.
- Error tolerance with no pause while encoding the information: partial damage of a network can diminish its efficiency; although a number of capabilities of the network may remain even with large losses.

Discussion and results

According to high sensitivity of neural networks against type of applied data and correlation of inputs with each other and as a result their relation with relevant outputs, regardless of type of network and its application as an instrument for producing artificial process, a proper attitude should be provided of available data and quality of organizing the information.

Determining relationship using regression

Through standardizing data and adapting them, two variables would be formed using tabled data, which the first is considered as "cost" ratio and another one is considered as "time" and then, the data would be regulated based on the two variables. The data have been distributed as follows:

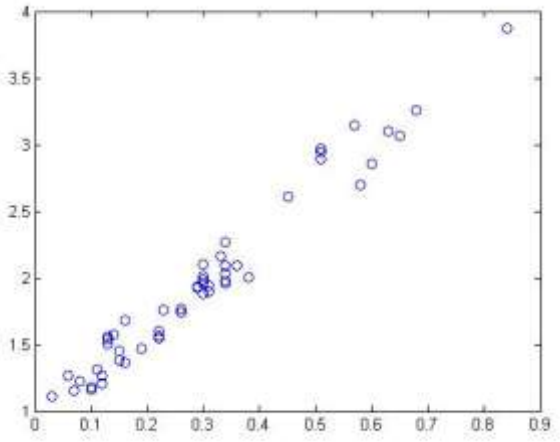

Figure 2: test data [1]

First order linear regression has is as follows:
P= 0.9235 3.4783

$$\text{cost} = 0.9235 * \text{time} + 3.4783$$

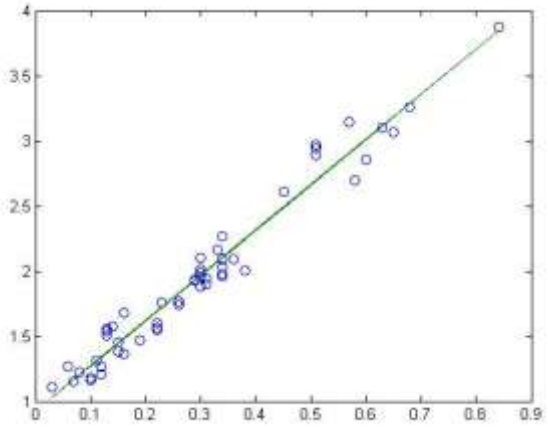

Figure 3: estimation through first order linear regression

P= 0.2200 3.3121 0.9463

$$\text{cost} = .22 * \text{time}^2 + 3.3121 * \text{time} + 0.9463$$

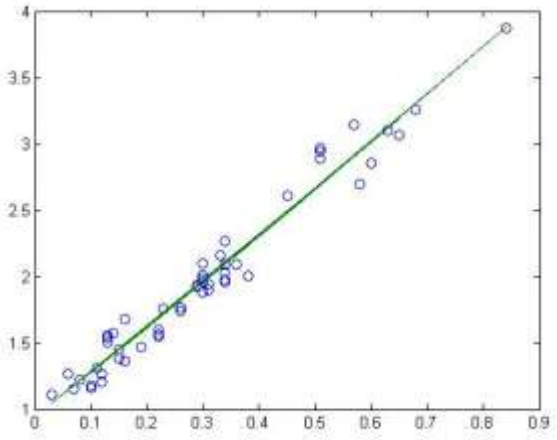

Figure 4: estimation through second order linear regression

P= -2.2082 2.9016 2.4272 1.0187

$$\text{cost} = -2.2 * \text{time}^3 + 2.9 * \text{time}^2 + 2.4 * \text{time} + 1.01$$

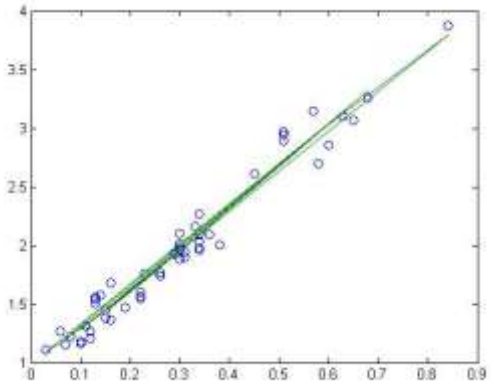

Figure 5: estimation through third order linear regression

Determining the relationship using LOWESS Algorithm

The aim by determining regression of given data is determining different polynomials with different degrees, so that they can determine the best polynomial that is fit to the data. In the examinations of the previous section, the closest polynomial to given data has been calculated in form of third degree linear regression. Through enlarging the degree of polynomial, better answer would be obtained. Without using additional knowledge and just through using main data of the problem, validation of data-based curve would be solved. For this purpose, LOWESS algorithm would be applied. The accuracy of this algorithm is depended on the parameter, which has been applied in it and is known as smoothing parameter. In this section, 99 LOWESS models would be constructed, in which smoothing parameter is varied from 0 to 1. Among the 99 models, the best one would be selected. The method can result in appearance of over-fitting problem. This is because; LOWESS algorithm estimates surplus part of data (noises) with same accuracy of estimating main part of data. In order to avoid over-fitting phenomena, cross validation technique would be applied. In this method, produced data are divided to two groups as follows: a) train data and b) test data. LOWESS algorithm applies train data for the purpose of model construction; although validation and accuracy of the presented model would be determined using test data. Applied criterion for evaluation of the model has been considered sum of squared errors.

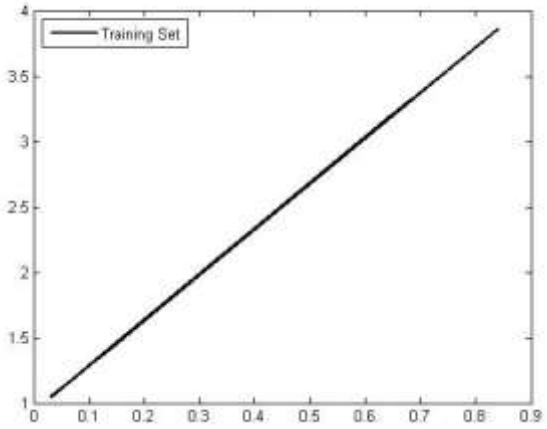

Figure 6: trained line

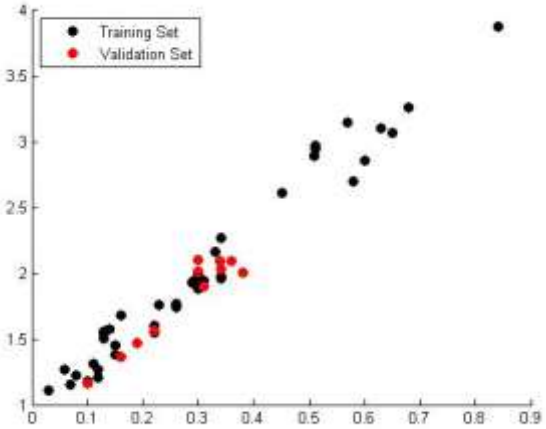

Figure 9: test and train data

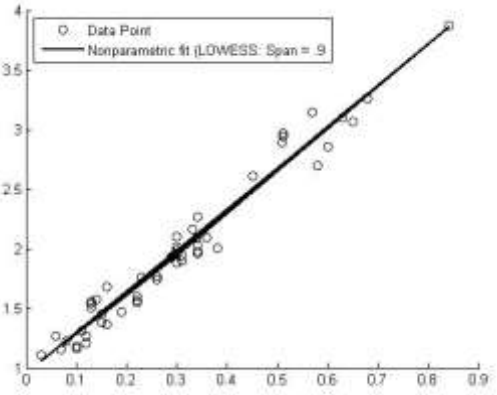

Figure 7: data scattering and trained line

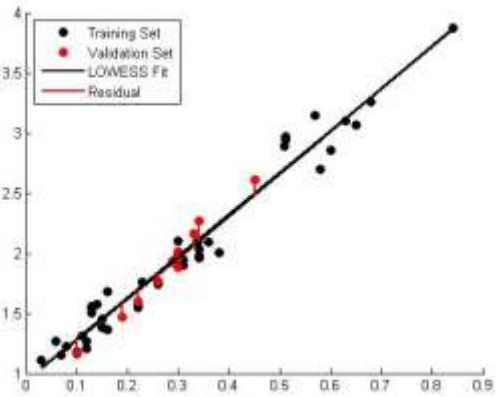

Figure 10: error determination in validation data

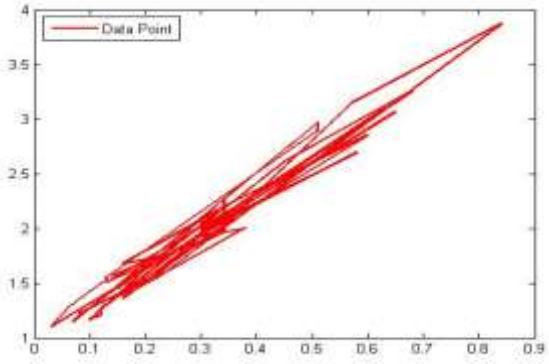

Figure 8: data scattering

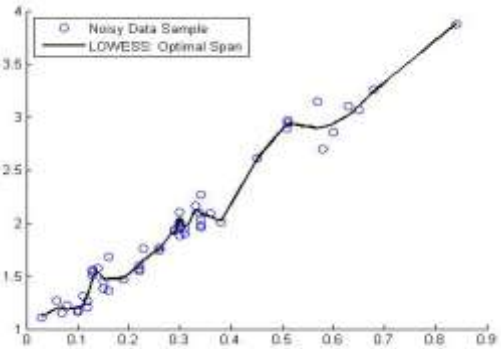

Figure 11: estimation through using LOWESS

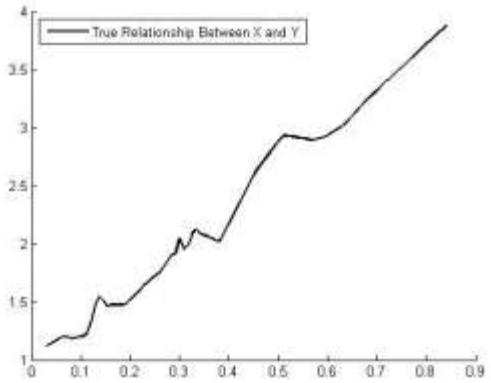

Figure 12: estimation through using LOWESS algorithm

Determining the relationship using neural network

The first step for collection of useful train data is testing and validating neural networks (systems). A perceptional collection of test results for compressive strength of concrete has been presented in FRP.

Applied parameters as input nodes in ANN modeling are as follows:

- Cost of the project
- Completion time of the project
- Applied sources for the project
- Work results and outcomes

In order to make data scalable from 0.1 to 0.9, minimum and maximum values for using in linear relations among values have been applied. Table 1 would present statistical characteristics of collected data.

Regression values (R values) would be measured between outputs and targets in the network. R value of 1 refers to a close relationship and R value of 0 refers to a random relationship.

Figure 15 has illustrated regression values of networks with different numbers of hidden nodes.

Figure 16 has presented maximum squared error (MSE) against the number of hidden neural layers after initial acceptance of desired networks. The best networks include NN-11-1 and NN6-8-1. In order to achieve ideal models, NN6-11-1 has been selected, since the model can present desirable and good results for R values and includes minimum MSE rate among all networks. Results of NN6-11-1 have been briefly presented in figures 17.

Figure 17 has presented squared errors in the networks in small and large scales. In other words, the figure indicates that the network is in learning mode. A neural network with 10 neurons on the first layer has been trained using given data. A neural network is as follows:

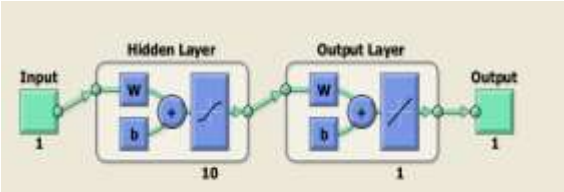

Figure 13: applied neural network

Using trained input and output data, the combination of test and train data and validation data would be presented as follows:

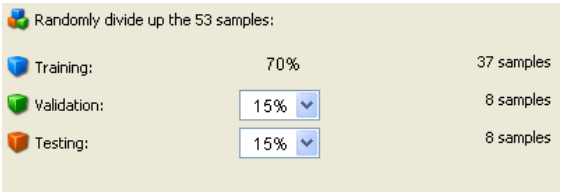

Figure 14: distribution of test and train data

After training operations, error and train diagrams have been obtained as follows:

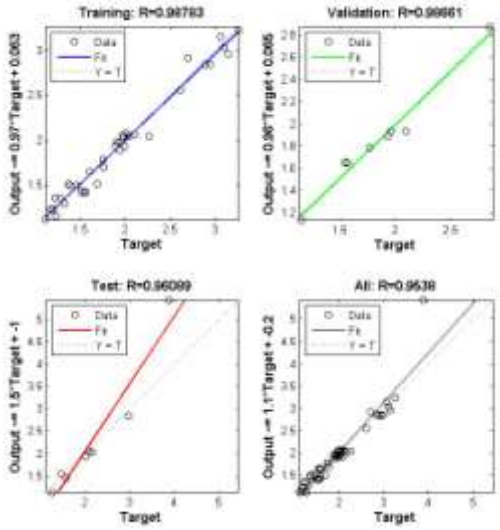

Figure 15: relationship between inputs and outputs

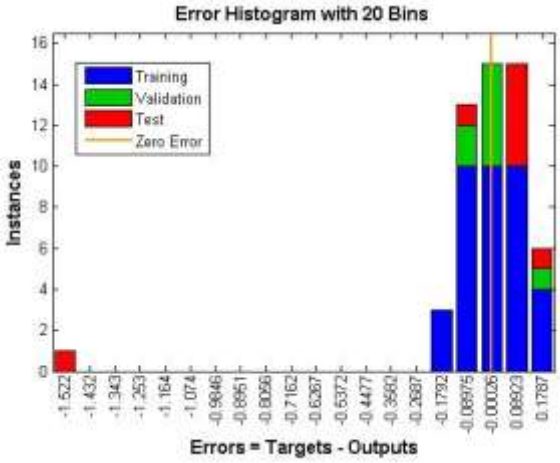

Figure 16: error curve

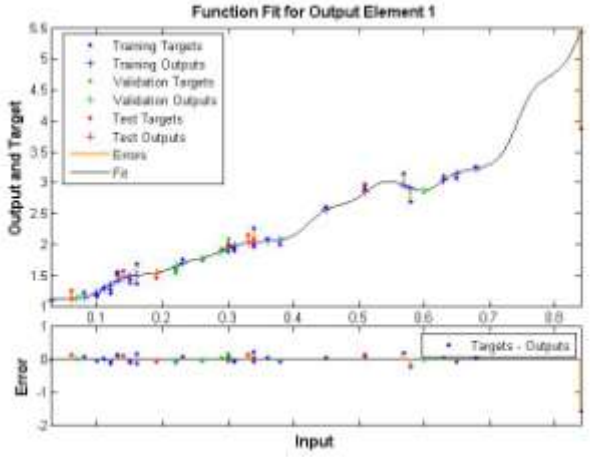

Figure 17: error estimation and fitted diagram with the neural network

Table 1: obtained data from the tests

S. No.	D (mm)	L (mm)	$\frac{L}{D}$	Type of FRP	t (mm)	f_{frp} (N/mm^2)	E_{frp} (N/mm^2)	f_{co} (N/mm^2)	$\frac{f_l}{f_{co}}$	f_{cc} (N/mm^2)		$\frac{f_{cc,cal}}{f_{cc,exp}}$
										Experiment	From Eq. (5)	
1	152	305	2.0	CFRP	0.31	755	73300	38.6	0.08	47.2	48.9	1.04
2	152	305	2.0	CFRP	0.61	1047	70600	38.6	0.22	60.6	67.0	1.11
3	152	305	2.0	CFRP	0.61	1047	70600	38.6	0.22	61.9	67.0	1.08
4	152	305	2.0	CFRP	0.92	1105	77500	38.6	0.34	80.9	82.6	1.02
5	152	305	2.0	CFRP	0.92	1105	77500	38.6	0.34	76.4	82.6	1.08
6	152	305	2.0	CFRP	0.92	1105	77500	38.6	0.34	75.8	82.6	1.09
7	152	305	2.0	CFRP	0.92	822	54000	38.6	0.26	68.3	72.2	1.06
8	152	305	2.0	CFRP	0.92	822	54000	38.6	0.26	67.3	72.2	1.07
9	152	305	2.0	CFRP	1.22	388	27700	38.6	0.16	52.6	59.3	1.13
10	100	200	2.0	CFRP	0.42	1285	576600	30.2	0.36	63.3	66.6	1.05
11	100	200	2.0	CFRP	0.14	1579	628600	30.2	0.15	41.7	45.3	1.09
12	152	610	4.0	GFRP	1.0	383	21600	26.2	0.19	38.4	42.8	1.11
13	152	610	4.0	GFRP	1.0	383	21600	26.2	0.38	52.5	59.5	1.13
14	152	610	4.0	CFRP	1.0	580	38100	26.2	0.29	50.6	51.7	1.02
15	150	300	2.0	CFRP	0.12	2600	200000	34.9	0.12	44.3	48.9	1.10
16	150	300	2.0	CFRP	0.12	2600	200000	34.9	0.12	42.2	48.9	1.16
17	150	300	2.0	CFRP	0.24	1100	420000	34.9	0.10	41.3	46.5	1.13
18	150	300	2.0	CFRP	0.24	1100	420000	34.9	0.10	40.7	46.5	1.14
19	153	305	2.0	GFRP	1.45	524	37233	29.6	0.34	67.1	63.3	0.94
20	153	305	2.0	GFRP	1.45	524	37233	29.6	0.34	60.2	63.3	1.05
21	153	305	2.0	GFRP	2.21	579	40336	29.6	0.57	93.0	86.1	0.93
22	153	305	2.0	GFRP	2.97	641	40749	29.6	0.84	114.7	112.9	0.98
23	153	305	2.0	GFRP	1.45	524	37233	32.0	0.31	60.8	65.2	1.07
24	51	102	2.0	CFRP	0.09	3500	235000	41.0	0.30	86.0	82.2	0.96
25	51	102	2.0	CFRP	0.18	3500	235000	41.0	0.60	117.0	123.4	1.05
26	152	305	2.0	CFRP	0.3	380	25000	43.7	0.03	48.4	48.1	0.99
27	150	300	2.0	GFRP	0.3	583	52000	36.3	0.06	46.0	43.6	0.95
28	150	300	2.0	GFRP	0.6	583	52000	36.3	0.13	55.8	52.1	0.93
29	150	300	2.0	GFRP	0.6	583	52000	36.3	0.13	56.4	52.1	0.92
30	150	300	2.0	GFRP	2.4	583	52000	36.3	0.51	104.9	98.3	0.94
31	150	300	2.0	GFRP	2.4	583	52000	36.3	0.51	106.9	98.3	0.92

Conclusion

At the present study, based on characteristics of different samples of projects and based on time, cost, and several other indicators, suitable formulation between time and cost has been presented using classic (regression) and intelligent (artificial neural networks) methods. In fact, learned information and data in previous projects would be applied in order to complete new projects. According to the presented methods, neural network has indicated better accuracy and validity than other methods in terms of estimating the functions. Moreover, regression method has also indicated the lowest degree of accuracy in estimation of relationship between time and cost of completion of the project.

References

[1] H. Naderpour"Empirical Approach Using ANN for Determining Confined Compressive Strength in FRP-St rengthened Concrete Members"15 wcee.

[2] H. Demuth, M. Beale, "Neural network toolbox for use with MATLAB, user's guide version 4, neural network toolbox user's guide," copyright 1992 - 2002, Math Works, Inc.

[3] C.F. Kemerer, "An Empirical Validation Of Software Cost Estimation Models,» Communications of the ACM 30 416–429, 1987.

[4] D.V. Ferens, R.B. Gumer, «An Evaluation For Three Function Point Models Of Estimation Of Software Effort,» IEEE National Aerospace and Electronics Conference, 1992.

[5] S.S.Vicinanza, T.Mukhopadhyay, J.J. Prietula, "Software-Effort Estimation: An Exploratory Study Of Expert Performance,» Information Systems Research, December , 243–262, 1991.

[6] T.Menzies,et al,» Validation Methods For Calibrating Software Effort Models, « 27th International conference on Software Engineering, Missouri,USA,15-21,2005.

[7] J. Hakkarainen, et al "Neural Networks In Specification Level Software Size Estimation," Proceeding of the Twenty-Sixth Hawaii International Conference on Volume iv, 5-8 Jan. 1993 Page(s):626 - 634 vol.4.

[8] C.Mair,et al, "An Investigation Of Machine Learning Based Prediction Systems," Journal of systems and software, volume 53,Issue 1, 2000.