KICEM Journal of Construction Engineering and Project Management
Online ISSN 2233-9582
www.jcepm.org
http://dx.doi.org/10.6106/JCEPM.2014.4.4.009

A Neural Network Model for Building Construction Projects Cost Estimating

Nabil Ibrahim El-Sawalhi [1] and Omar Shehatto [2]

Abstract: The purpose of this paper is to develop a model for forecasting early design construction cost of building projects using Artificial Neural Network (ANN). Eighty questionnaires distributed among construction organizations were utilized to identify significant parameters for the building project costs. 169 case studies of building projects were collected from the construction industry in Gaza Strip. The case studies were used to develop ANN model. Eleven significant parameters were considered as independent input variables affected on "project cost". The neural network model reasonably succeeded in estimating building projects cost without the need for more detailed drawings. The average percentage error of tested dataset for the adapted model was largely acceptable (less than 6%). Sensitivity analysis showed that the area of typical floor and number of floors are the most influential parameters in building cost.

Keywords: Building construction projects, Artificial Neural Network, Gaza Strip, Cost estimation

I. Introduction

Cost is one of the three main challenges for the construction manager, where the success of a project is judged by meeting the criteria of cost with budget, schedule on time, and quality as specified by the owner [1]. Poor strategy or incorrect budget or schedule forecasting can easily turn an expected profit into loss [2]. Therefore, effective estimating is one of the main factors of a construction project success [3].

In recent decades, researchers and participants in construction industry have recognized the potential impact of early planning to final project outcomes. Therefore, they started to put more emphasis on early planning process, where the project definition in the early planning process is an important factor leading to project success [4].

The cost estimate becomes one of the main elements of information for decision making at preliminary stage of construction. Thus, improved cost estimation techniques will facilitate more effective control of time and costs in construction projects [5]. Actually, estimates are prepared and used for different purposes including feasibility studies, tendering phase, avoidance misuse of funds during the project, etc. The primary function of cost estimation is to produce a credible cost prediction of a construction project. However, the predicted cost depends on the requirements of a client and upon the information and data available [6].

The largest obstacles standing in front of a cost estimate, particularly in early stage, are lack of preliminary information and larger uncertainties as a result of engineering solutions.

As such, to overcome this lack of detailed information, cost estimation techniques are used to approximate the cost within an acceptable accuracy range [7].

Due to the inadequacy of traditional estimating techniques in conceptual stage and most common methods for estimating in Gaza strip are still the traditional and spreadsheet methods. Therefore, it is desirable to create a new method that helps a user with little knowledge to quickly create an accurate cost estimate.

Moreover, in conceptual stage, there is a limited of available data and lack of appropriate cost estimation methods, which need to search for a technique that capable to deal with limited information and give accurate estimate with very limited data. ANNs are utilized as systems able to generalize solutions for these problems by learning from a set of historical examples with a little of time and effort.

Many researchers apply neural network approach in various fields of engineering prediction and optimization. However, the authors reckon that the researches and studies on utilizing neural networks to estimate the cost of construction projects at various stages of the work are very limited [12] ,[13].

This research developed a model of ANN to estimate the cost of building projects at the conceptual stage depending on the historical data of projects implemented in Gaza strip between “2009-2012” to help the users in predicting cost of projects at early stage with high level of accuracy.

The aim of this research is develop a prototype model for estimating the cost of building construction projects including skeleton and architecture factors using artificial neural networks ANNs. Besides, giving a better understanding of the underlying factors of buildings cost, which can help the users to prepare preliminary cost estimate of building projects.

This research is limited to the buildings sector of

[1.] Depty Dean faculty of Engineering , The Islamic University- Gaza, E-Mail: nsawalhi@iugaza.edu.ps (*Corresponding Author)
[2] Master degree, The Islamic University- Gaza, E-Mail: shehato@hotmail.com

construction projects in Gaza strip; including the main two phases of building construction; skeleton and finishing phase. Thus, collecting data on building projects that were implemented between 2009 and 2012 were conducted. It is assumed in this study that the initial architecture drawings of the building project are available which include the information of adopted model factors.

II. PARAMETRIC COST FACTORS

The "parametric" method of estimating involves collecting relevant historical data, usually at an aggregated level of detail, and relating it to the product to be estimated through the use of mathematical techniques. Since parametric methods typically capture cost at a very high level, less detail is required for this approach than for other methodologies [8]. Parametric estimation is expressed as an analytical function of a set of variables. These usually consist in some features of the project (performances, type of materials used), which are supposed to influence mainly the final cost of the project. Commonly, these analytical functions are named ''Cost Estimation Relationships'' (CER), and are built through the application of statistical methodologies [9].

One of the most significant keys in modelling buildings early stages cost estimating is identifying the factors that have real impact on the cost of building projects. Depending on this great importance of selecting these factors, several techniques were adopted carefully to identify these parameters for building projects. The factors affecting parametric cost estimate are illustrated. Since a construction project's location affects the final cost, an estimator must understand what particular location factors will be encountered and what considerations should be taken into account when formulating the estimate.

General assumptions about soil conditions may be made early in the estimating process, but they may turn out to be wrong. As the estimate progresses, geotechnical data may help improve the information and prevent costly change orders and claims. In the early estimates the assumptions regarding soil conditions and the potential effects of unknown soil conditions should be clearly documented. Soil conditions can be a significant cost risk to a project. The soil type will influence the chosen foundation type [10].

The quantity and type of a given material on a project impacts the unit cost of constructing and/or supplying that item. This is not simply a supply and demand issue, but also one of production efficiency and economy of scale. Generally speaking, the unit price for larger quantities of a given material will be less than smaller quantities. Mobilization, overhead and profit are all spread out over a larger quantity, thus reducing their affect on each unit.

Small quantities of items of work are less cost effective to construct and hence lead to higher unit prices. This practice increases a contractor's overhead and usually results in a mark-up being applied to those items [10].

Types of plastering, tilling, marble, electrical, sanitary, carpentry, metal, and aluminum would largely affect cost estimating. The nine building functional element groups of cost classification system are as follows: Slab on ground, Number of floors , stairs and elevators, External walls, Windows, External doors, Floor height, Internal doors, Area of ground and typical slabs and finally the columns quantity and length in between [11].

Floor area parameter, Number of stories, Slab type, Foundation type, Number of elevators, Type of project, Type of project and External finishing parameter were adopted by several researchers [5];[12]; [13];[2]; [4]; [14]; [15].

III. NEURAL NETWORK

In the recent years, new approaches based on the theory of computer systems that simulate the learning effect of the human brain as Artificial Neural Networks (ANNs) has grown in popularity [9]. ANNs is one of these new approaches that is able to perform tasks involving incomplete data sets, fuzzy or incomplete information and for highly complex and ill-defined problems. Moreover, it is able to deal with non-linear problems. One of the distinct characteristics of ANN is its ability to learn from experience and examples and then to adapt to changing situations. It has a natural propensity for storing experiential knowledge and making it available for use [14]. Another major benefit of using ANN is its ability to understand and simulate more complex functions than older methods such as linear regression [16]. In addition, it can approximate functions well without explaining them. This means that an output is generated based on different input signals and by training those networks, accurate estimates can be generated [7]. In spite of great accuracy of using ANN model in cost estimation, it has a considerable defect, as it depends mainly on historical data; this dependency has several disadvantages as the following:

1. Diversity of variables for effective factors is limited to what available in collected data.
2. Data should contain sufficient projects for each variable.
3. New variables which was not included in adopted model will not be handled.

Despite the large number of researchers who applied neural network approach in various fields of engineering, the studies and researches on utilizing neural networks to estimate the cost of construction projects at various stages of the work are very limited[13].

Locally, there is a lack of cost estimation researches based on ANN applied in Gaza Strip. Arafa and Alqedra (2011) developed an ANN model to estimate the cost of building construction projects at early stages. A database of 71 building projects collected from the construction industry of the Gaza Strip was used in a developed ANN model. The model had one hidden layer with seven neurons. The results obtained from the trained models indicated that neural networks are reasonably succeeded in predicting the early stage cost estimation of buildings using basic information of the projects and without the need for a more detailed design [13].

Regionally, Elsawy et al., (2011) developed a neural network model to assess the percentage of site overhead costs for building projects in Egypt, which can assist the decision makers during the tender analysis process [15]. Kim et al., (2004) applied hybrid models of ANN and Genetic Algorithm (GA) to estimate the preliminary cost of residential buildings. They first optimized the parameters of the back-propagation algorithm using genetic algorithms and then obtained a set of trained weights for the ANN model using GA. The results of the research revealed that optimizing each parameter of back-propagation networks using GA is most effective in estimating the preliminary costs of residential buildings [5]. Gunaydin and Dogan (2004) developed an ANN model to estimate the cost of a square meter of the structural system of buildings in early phases of design processes. The input layer of the trained ANN model comprised eight parameters available at the early design stage. The trained ANN model was capable of providing accurate estimates of at least 93% of buildings cost per square meter [12].

Emsley et al., (2002) trained neural network cost models using a database of data nearly 300 building projects. They used linear regression techniques as a benchmark for evaluation of the neural network models. The results showed the ability of neural networks to model the nonlinearity in the data, where the model was capable of evaluating the total cost of the construction, and the trained ANN model obtained a mean absolute percentage error of 16.6 % [17].

The above researches reviewed by the authors indicated that the application of artificial neural networks to estimate the early cost of construction projects is a promising area.

IV. METHODOLOGY

An extensive review of previous studies, with structured questionnaire and expert interviews were used to identify the most influential factors on building project cost in Gaza Strip. These influential factors would be the independent input parameters in the neural network model and they will form the basis of collected information from historical cases of building projects from municipalities, government ministries, engineering institutions, contractors and consultants. After analysing the data, many models would be built and trained with various structures by using Neuro-Solution 5.07 application.

4.1 Questionnaire design

A questionnaire was designed according to the identified factors that affect parametric cost estimate of projects. Thirteen cost parameter in skeleton phase (Area of typical floor, Number of floors, Area of retaining walls, Use of building, Type foundation, Number of elevators, Type of slab, Length of spans, Number of columns Number of rooms, Location of project, Number of staircases and Type of contract) and eighteen finishing phase parameters (Type of external plastering, Volume of Air-conditioning, Area of curtain walls, Type of tiling, Type of water and sanitary works, Type of electrical works, Area of gypsum board and false ceiling, Area of marble works, Fire fighting and alarm works, Quantity of electrical works, Number of windows, Quantity of water and sanitary works, Type of carpentry works, Number of internal doors, Type of Aluminium works, Type of carpentry works, Quantity of metal works and Type of painting) identified from literature were evaluated. Eighty questionnaires were distributed to various engineering institutions. Fifty-seven questionnaires with a response rate 71% have been correctly received.

The case studies used in this research was collected from different institutions concerned with construction engineering in Gaza Strip. A data sheet was prepared and used to extract all useful information from 193 bids of building projects during 2009 and 2012. However, In order to overcome any defect in collected data, some basic assumptions and criteria were defined and performed on collected projects. 24 projects of 193 projects were eliminated. Therefore, 169 projects were used to build the neural network model.

4.2 Data Analysis

To measure the accuracy of the neural network model, several methods were used. Mean Absolute Error (MAE) is one of many ways to quantify the difference between an estimated and the actual value of the projects being estimated. According to Willmott & Matsuura, (2005) the MAE is relatively simple; It involves summing the magnitudes (absolute values) of the errors to obtain the 'total error' and then dividing the total error by n, it can be defined by the following formula [18]:

$$\mathrm{MAE} = \frac{\sum_{j=0}^{P}\sum_{i=0}^{N}\left|dy_{ij} - dd_{ij}\right|}{N\,P} \qquad \text{Eq. (1)}$$

Where:

P = number of output PEs. N = number of exemplars in the data set. dyij= denormalized network output for exemplar i at PE j. ddij= denormalized desired output for exemplar i at PE j.

Mean Absolute Percentage Error (MAPE) is a quantity used to measure how close forecasts or predictions are to the eventual outcomes, according to Principe, et al., (2010) The MAPE is defined by the following formula [19]:

$$\mathrm{MAPE} = \frac{100}{N\,P}\sum_{j=0}^{P}\sum_{i=0}^{N}\frac{\left|dy_{ij} - dd_{ij}\right|}{dd_{ij}} \qquad \text{Eq. (2)}$$

Where:

P = number of output PEs. N = number of exemplars in the data set. dyi j= denormalized network output for exemplar i at PE j. ddi j= denormalized desired output for exemplar i at PE j.

According to Principe et al., (2010) the size of the mean square error (MSE) can be used to determine how well the network output fits the desired output, but it doesn't necessarily reflect whether the two sets of data move in the same direction. For instance, by simply scaling the network output, we can change the MSE without changing the directionality of the data. The correlation coefficient (r) solves this problem [19].

By definition, the correlation coefficient between a network output x and a desired output d is:

$$\boldsymbol{r} = \frac{\frac{\sum_i (x_i - \bar{x})(d_i - \bar{d})}{N}}{\sqrt{\frac{\sum_i (d_i - \bar{d})^2}{N}} \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{N}}} \qquad \text{Eq. (3)}$$

According to Hegazy & Ayed, (1998); the total MAPE methodology is defined by determining the total MAPE [20]. Training phase were represented fifty percent of the total MAPE. Likewise the test set is equal the remaining fifty percents.

$$\text{Total MAPE} = \frac{(\text{MAPE}_{\text{Tr}} \times \text{N}_{\text{Tr}} + \text{MAPE}_{\text{C.V}} \times \text{N}_{\text{C.V}})/(\text{N}_{\text{Tr}} + \text{N}_{\text{C.V}}) + \text{MAPE}_{\text{Test}}}{2}$$

…. Eq. (4)

Where: MAPETr = Mean absolute percentage error for training data set. NTr = number of exemplars in the training data set. MAPEC.V = Mean absolute percentage error for cross validation data set. NTr = number of exemplars in the cross validation training data set. MAPE Test = Mean absolute percentage error for test data set.

V.RESULTS AND DISCUSSION

Most of respondents for the thirteen skeleton phase cost factors declared that the area of typical floor (90%) and number of floors (90%) are the most influential factors on building cost, while area of retaining wall (75%), type of building (73%), type of foundation (72%), number of elevators (71%) and type of slab (71%) have a moderate influence. For remaining parameters as, length of span between columns (69%), number of columns (65%), number of rooms (59%), location of project (57%), number of stair cases (57%), and type of contract (51%) have a lower influence on the project cost. On the other hand, for the influence of eighteen finishing parameters, the external finishing type has the highest rate (78%), while volume of air-conditioning, area of curtain walls, type of tiling, type of sanitary, and type of electrical works have a rate between 72%-76%. Accordingly, five experts in construction field were selected to reach a consensus about specifying the key cost parameters. The results with those five experts were significantly close to the questionnaire results, and only three Delphi rounds were conducted due to largely degree of consensus. Table (I) shows the most influential parameters which had the highest rate adopted in several previous researches.

TABLE I
INFLUENTIAL PARAMETER OF BUILDING PROJECT COST

No.	Skeleton and finishing phase input parameter	Range
1	Area of typical floors	Less than 1200 m2
2	Number of storey	(1-8) storey
3	Use of building	Residential - Schools - Public
4	Type of foundation	Isolated – Strap – Piles - Mat
5	Type of slab	Solid - Ribbed – Slab with drop beams
6	Number of elevators	(0-1-2)
7	Type of external finishing	Plaster – Natural stone-Others
8	Presence of HVAC and false ceiling	Central conditioning - Split units
9	Type of tilling	Ceramic – Terrazzo - Porcelain
10	Type of electricity works	Basic - Luxury
11	Type of mechanical works	Basic - Luxury

VI. MODEL DEVELOPMENT

The developed model in this research based on Neuro Solution 5.07 for Excel program. It was selected for its ease of use, speed of training, flexibility of building and executing the ANN model. In addition, the modeller has the flexibility to specify his own neural network type, learning rate, momentum, activation functions, number of hidden layers/neurons, and graphical interpretation of the results.

Five steps for implementing the neural network model were followed as shown in Figure I and illustrated below:

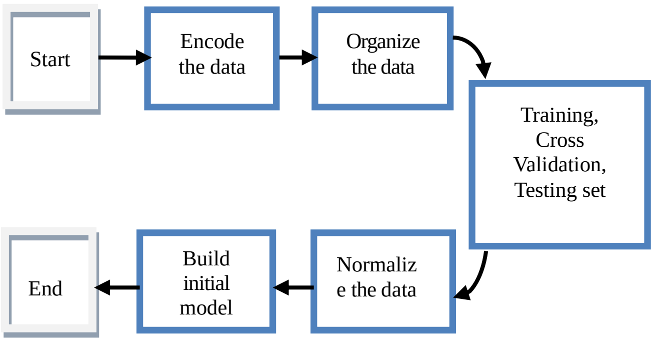


FIGURE I
MODEL IMPEMENTATION STEPS FLOWCHART

Artificial networks only deal with numeric input data. Therefore, the raw data must often be converted from the external environment to numeric form [21]. This may be challenging because there are many ways to do it and unfortunately, some are better than others are for neural network learning [19]. In this research the data is textual and numeric, so it is encoded to numeric form.

Initially, the first step in implementing the neural network model in Neuro-Solution application is to organize the Neuro-solution excel spreadsheet. Then, specifying the input factors that have been already encoded, which consist of 11 factors; type of project, area of typical floor, number of floors, type of foundation, type of slab, number of elevators, type of external finishing, type of air-conditioning, type of tilling, type of electricity, and type of sanitary. Finally, specifying the desired parameter (output) which is (total cost of the project).

As a rule of thumb, determining the number of hidden layer/neurons is one of the main drawbacks of ANNs, because there is no specific rule and it requires many trial and error processes, while considerable time must be spent [5]. Hegazy & Moselhi, (1995) stated that one hidden layer with a number of hidden neurons as 0.5 m, 0.75m, m, or 2m+1, where m is the number of input neurons, is suitable for most applications [22].

The available data were divided into three sets namely; training set, cross-validation set and test set. Training and cross validation sets are used in learning the model through utilizing training set in modifying the network weights to minimize the network error, and monitoring this error by cross validation set during the training process. However, test set does not enter in the training process and it hasn't any effect on the training process, where it is used for measuring the generalization ability of the network, and evaluated network performance [13].

In the present study, the total available data is 169 exemplars that were divided logical randomly, according to literatures, into three sets with the following ratio:

- Training set (includes 116 exemplars ≈ 69%).
- Cross validation set (includes 27 exemplars ≈ 16%).
- Test set (includes 26 exemplars ≈ 15%).

Before starting the training phase, it is usually necessary to scale the data, or normalize it to the network's paradigm. Kshirsagar & Rathod (2012) and Gunaydın & Dogan (2004) stated that data is generally normalized for the purpose of confidentiality and for effective training of the model being developed, where the input data must be normalized between an upper and lower bound [19]. The normalization of training data is recognized to improve the performance of trained networks. Therefore, the input/output data was scaled, zero is the lower bound and the upper bound is one to suit neural networks processing by using Neuro-solution program.

Once all data were prepared, then the subsequent step is represented in creating the initial network by selecting the network type, number of hidden layer/nodes, transfer function, learning rule, and number of epochs and runs. An initial neural network of Multilayer Perceptron (MLP), that consists of one input, hidden, and output layer, was built and a supervised learning control was checked to specify the maximum number of epochs and the termination limits.

6.1 Neural network training

The objective of training a neural network is to get a network that performs best on unseen data through training many networks on a training set and comparing the errors of the networks on the validation set [23]. Therefore, several network parameters such as number of hidden layers, number of hidden nodes, transfer functions and learning rules were trained multiple times to produce the best weights for the model.

As a preliminary step to filter the preferable neural network type, a test process was applied for most of available networks in the application. Two types Multilayer Perceptron (MLP) and General feed Forward (GFF) networks were chosen, due to their good initial results, to be focused in following training process.

The training process started with selecting the neural network type either MLP or GFF network. For each one, five types of learning rules were used, and with every learning rule six types of transfer functions were applied, and then 3 separate hidden layers were utilized with increment of hidden nodes from 1 node up to 40 nodes in each layer.

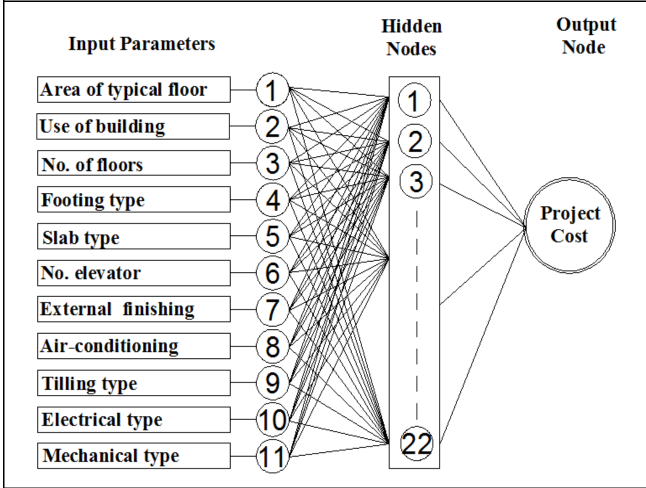
More than one and a half thousand trials contain 40 variable hidden nodes for each was executed to obtain the best model of neural network. Ten runs, in each one 3000 epochs, were applied, where a run is a complete presentation of 3000 epochs, each epoch is a one complete presentation of all of the data [19]. However, in each run, new weights were applied in the first epoch and then the weights were adjusted to minimize the percentage of error in other epochs. To avoid overtraining for the network during the training process, an option of using cross-validation was selected, which computes the error in a cross validation set at the same time that the network is being trained with the training set. The model was started with one hidden layer and one hidden node in order to begin the model with simple architecture, and then the number of hidden nodes was growing up by one node up to 40 hidden nodes.

6.2 Neural network testing

The purpose of testing phase of ANN model is to ensure that the developed model was successfully trained and generalization is adequately achieved. Therefore, testing the network is essentially the same as training [15]. The testing set is critical to confirm that the network has not simply memorized a given set of data but has learned the general patterns involved within an application [21]. The testing data is totally a different set of data that the network is unaware of; after finishing the training process testing data is used for validation and generalization of the trained network. If the network is able to generalize rather precisely the output for this testing data, then it means that the neural network is able to predict the output correctly for new data and hence the network is validated [24].

Through a system of trial and error guided by earlier recommendation, the best model that provided more accurate cost estimate without being overly complex was structured of Multilayer Preceptron (MLP) includes one input layer with 11 input neurons and one hidden layer with (22 hidden neurons) and finally one output layer with one output neuron (Total cost). However, the main downside to using the Multilayer Preceptron network structure is that it required the use of more nodes and more training epochs to achieve the desired results. Figure (II) summarizes the architecture of the model as number of hidden layer/nodes, type of network and transfer function.

Architecture of the model



Model Type	Transfer Function	Update Methods	Gradient Search
Multilayer Perceptron	Tanh	Batch	Momentum
No. of hidden layer	No. of PEs in the input layer	No. of PEs in the 1st hidden layer	No. of PEs in the output layer
1	11	22	1

FIGURE II
ARCHITECTURE OF THE ADOPTED MODEL

The most common statistical performance measures were applied on the adopted model to ensure the validity of this model in estimating the cost of new projects as the following:

The Mean Absolute error (MAE) for the adopted model equals (33,757 $), it is largely acceptable for projects worth hundreds of thousands dollars. However, it is not a significant indicator for the model performance because it proceeds in one direction, where the mentioned error may be very simple if the total cost of a project is large, and in turn; it may be a large margin of error in case the total cost of a project is small.

The mean absolute percentage error of the model is calculated from the test set, which equals 6%, this result can be expressed in another form by accuracy performance (AP) according to Willmott and Matsuura (2005) which is defined as (100−MAPE) %[18].

Regression analysis was used to ascertain the relationship between the estimated cost and the actual cost. The results of linear regressing are illustrated graphically in Figure (3). The correlation coefficient (R) is 0.995, indicating that; there is a good linear correlation between the actual value and the estimated neural network cost at testing phase.

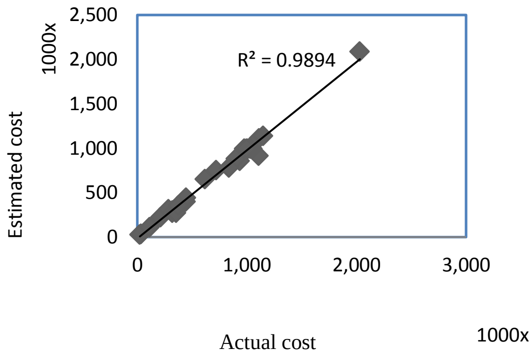


FIGURE III
LINEAR REGRESSION OF ACTUAL AND ESTIMATED COSTS

By reviewing many researches that used ANN in cost estimation, it is shown that no specific percent of allowable error for model estimate is available. However, the acceptable accuracy performance for ANN model is equal 10% according to [25] and [26]. In this study and according to Eq 4, the Total MAPE = 10%, where this error includes all datasets as training, cross validation, and test datasets.

The results of performance measures are summarized in Table (II), where the accuracy performance of adopted model is 94%. In which the average error is 6%.

TABLE II
RESULTS OF PERFORMANCE MEASUREMENT

	MAE	MAPE	AP	R	Total MAPE
MLP Model	33,757	6%	94 %	0.995	10 %

Figure IV describes the actual cost comparing with estimated costs for cross validation (C.V) dataset. It is noted that there is a slight difference between two cost lines.

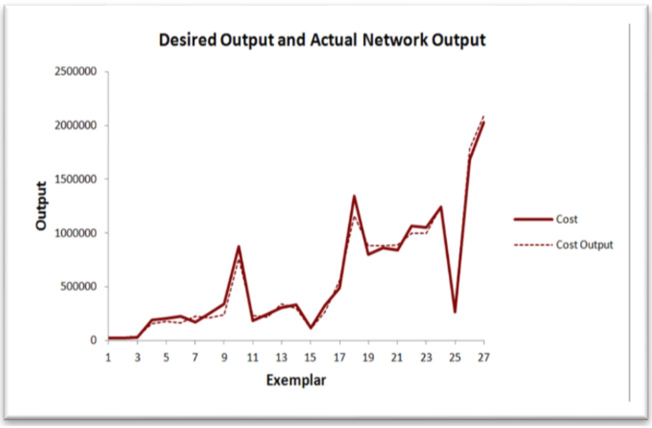


FIGURE IV
DESIRED OUTPUT & ACTUAL NETWORK FOR C.V SET

For test dataset, a perfect agreement between the actual and estimated cost is shown in Figure V which means the estimated values equal the actual ones.

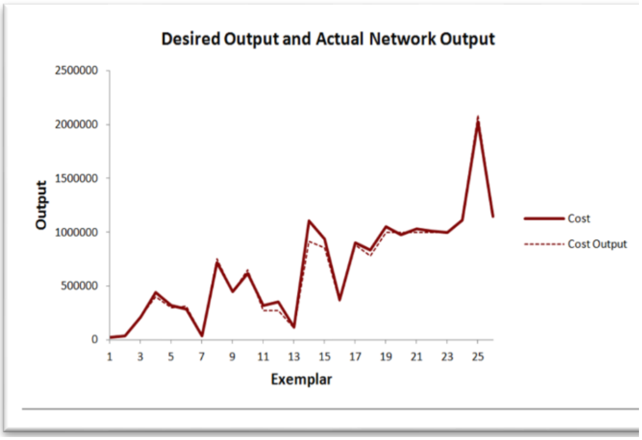


FIGURE V
DESIRED OUTPUT & ACTUAL NETWORK OUTPUT
FOR TEST SET

As presented in results, the average estimate accuracy of the adopted model is (6%). By comparing this estimate accuracy, where no need for drawings or details, with literature studies as Enshassi, et al., (2007) who stated that the level of accuracy for a project with no design work it may range from +40% to –20%. After preliminary design work, it may range from +25% to –10%. On completion of detailed design work it may range from +10% to –5%., it is clearly shown the high potential of using neural network models in cost estimate [27].

6.3 Sensitivity analysis

Once the optimum Neural Network model has been selected, sensitivity analysis was run on the best model to evaluate the influence of each input parameters on output variable. This provides a feedback to which input parameters are the most significant. It allows the user to see how input affects output over a range of values if all other inputs remain constant.

The NeuroSolution program provides a useful tool to identify sensitive input variables: ‘‘Sensitivity about the Mean’’. The sensitivity analysis was run after fixing the best weights. The start was by varying the first input between the mean ± one standard deviation, while all other inputs are fixed at their respective means. The network output was computed for 50 steps above and below the mean. This process was then repeated for each input, and the corresponding change in the output for each input is recorded then the standard deviation is calculated according to the following formula:

$$\sigma = \sqrt{\frac{(x-\bar{x})^2}{(n-1)}}$$

Eq. (5)

Where x: is the output value.
$\bar{x}$: is the mean of the output values.
n: is the number of the outputs in the sample.

It is very easy to compute in the trained network and effectively measures how much a change in a given input affects the output across the training data set. Inputs with large sensitivities have more importance in the mapping and therefore are the ones we should keep. The inputs with small sensitivities can be discarded [19](Principe et al., 2000). Finally, a report summarizing the variation of output with respect to the variation of each input was generated and presented in Figure VI as shown.

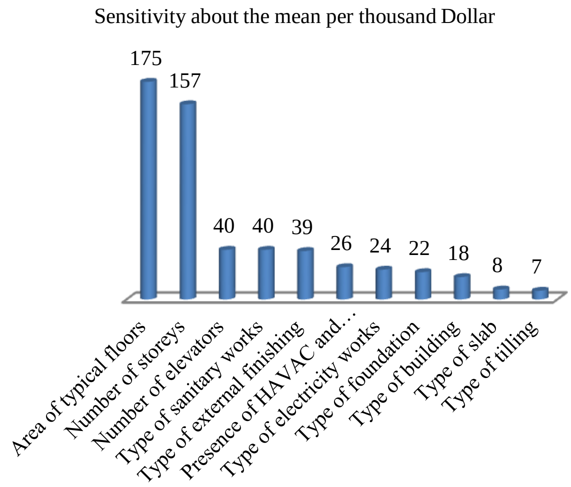


FIGURE VI
SENSITIVITY ABOUT THE MEAN

The increase of Standard Deviation refers to the strong influence of this parameter on the overall cost of the project. Figure (6) shows that the area of typical floor has the highest rate of influence on the total cost of projects. The value 175 for the area of typical floor input parameter is the value of the standard deviation for 116 output values. These output values are recorded after training the model with fixing the best weights on the mean value for each raw except the area of typical floor value which is varied between (the mean – standard deviation) to (the mean + standard deviation). Number of stories has also a very significant influence. Other parameters have a considerable gab of influence on total cost.

VII. CONCLUSION

This study aimed at developing ANN model for early cost estimate of building projects in Gaza Strip. This help parties involved in construction projects (owner, contractors, and others) in obtaining the total cost. Questionnaire survey and exploratory search of previous studies were used to identify the most effective parameters on building projects cost at the early stages of project. As well, historical data of building projects, which were executed between 2009 and 2012 in Gaza Strip, were used to build up the model.

ANN was used to predict future parametric cost of project. The best model that provided more accurate results was Multilayer Perceptron network model (MLP). The model was structured from one input layer included 11 input neurons (represents 11 parameters), one hidden layer contained 22 hidden neurons, one output neuron represent the predicted cost, Tanh transfer function, and Moentum learning rate which belongs to Back-propagation algorithm.

The accuracy performance of the adopted model recorded 94% where the model performed well and no significant difference was discerned between the estimated output and the actual budget value. The model is capable of providing very good estimates prediction of project cost at early stage of project.

REFERENCES

[1] A. Rezaian, "Time-cost-quality-risk of construction and development projects or investment". *Middle-East Journal of Scientific Research*, vol. 1, no. 2, pp. 218-223, 2011.

[2] M. Y. Cheng, H.-C. Tsai, E. Sudjono, "Conceptual cost estimates using evolutionary fuzzy hybrid neural network for projects in construction industry", *Expert Systems with Applications*, vol. 37, pp. 4224–4231, 2010.

[3] Y. Al-Shanti, "A cost estimate system for Gaza Strip construction contractors", The Islamic University of Gaza Strip, Palestine, 2003.

[4] Y. Wang, C. Yu, H. Chan, "Predicting construction cost and schedule success using artificial neural networks ensemble and support vector machines classification models", *International Journal of Project Management*, vol. 30, pp. 470–478, 2012.

[5] G.-H. Kim, S.-H. An, K.-I.Kang, "Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning", *Building and Environment*, vol. 39, pp. 1235–1242, 2004.

[6] S. Elhag, H. Boussabaine, A. Ballal, "Critical determinants of construction tendering costs: quantity surveyors_ standpoint", *International Journal of Project Management*, vol. 23, no.7, pp. 538–545, 2005.

[7] B. Verlinden, R. Duflou, D. Collin, D. Cattrysse, "Cost estimation for sheet metal parts using multiple regression and artificial neural networks: A case study*", Journal of Construction Engineering and Management (ASCE)*, vol. 127, no. 2, pp. 93–100, 2007.

[8] J. Long, "Parameteric cost estimating in the new millennium", 2006.

[9] S. Cavalieri, P. Maccarrone, R. Pinto, "Parametric vs. neural network models for the estimation of production costs: a case study in the automotive industry", *International Journal of Production Economics*, vol. 91, pp. 165–177, 2004.

[10] CEM, "Cost estimating manual for WSDOT projects", Washington State Department of Transport, 2008.

[11] H. Yaman, E. Tas, A building cost estimation model based on functional elements. Istanbul Technical University, Faculty of Architecture, *Istanbul Turkey, ITU A|Z* ,vol 4, no. 1, pp. 73-87, 2007.

[12] M. Gunaydın, Z. Dogan, "A neural network approach for early cost estimation of structural systems of buildings", *International Journal of Project Management*, vol. 22, pp. 595–602, 2004.

[13] M. Arafa, M. Alqedra, "Early stage cost estimation of buildings construction projects using ANN", *Journal of Artifical Intelligence*, vol. 4, no. 1, pp. 63-75, 2011.

[14] S., Doğan, "Using machine learning techniques for early cost estimation of structural systems of buildings", İzmir Institute of Technology, İzmir, 2005.

[15] I. ElSawy, H. Hosny, M. Abdel Razek, "A Neural network model for construction projects site overhead cost estimating in Egypt". *International Journal of Computer Science Issues*, vol. 8, no. 1, 2011.

[16] G. Weckman, H. Paschold, J. Dowler, H. Whiting, W. Young, "Using neural networks with limited data to estimate manufacturing cost", *Journal of Industrial and Systems Engineering*, vol. 3, no. 4, pp. 257-274, 2010.

[17] M.W. Emsley, D.J., Lowe, A.R. Duff, A.Harding, A. Hickson, "Development of Neural Networks to predict Total Construction Costs", *Journal of Construction Management and Economics*, Taylor and Francis Group Ltd., no. 20, pp. 456-472, 2002.

[18] C. Willmott, K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance", *Climate Research*, vol. 30, pp. 79–82, 2005.

[19] J. Principe, C. Lefebvre, G. Lynn, C. Fancourt, D. Wooten,"NeuroSolution Help" , NeuroDimension, Inc, 2010.

[20] T. Hegazy, A. Ayed, "Neural network model for parametric cost estimation of highway projects", *Journal of Construction Engineering and Management*, pp. 210-218, 1998.

[21] P. Kshirsagar, N. Rathod, "Artificial neural network" in MPGI National Multi Conference 2012 proceedings, International Journal of Computer Applications, Maharashtra, 2012.

[22] T. Hegazy, O. Moselhi, "Elements of cost estimation: "A survey in Canada and the United States", *Cost Engineering*, vol. 37, no.5, pp. 27-31, 1995.

[23] Z. Dindar, "Artificial neural networks applied to option pricing", University of the Witwatersrand, Johannesburg, 2004.

[24] P. Edara, "Mode choice modeling using artificial neural networks", Virginia Polytechnic Institute and State university, Virginia, 2003.

[25] N. Bakhary, K. Yahya, N. C. Nam, "Univariate artificial neural network in forecasting demand of low cost house in Petaling Jaya". Jurnal Teknologi, Vol. 40, pp. 67-75, 2004.

[26] P. Samphaongoen, "A Visual approach to construction cost estimating. Marquette University", Milwaukee, Wisconsin, 2010.

[27] A. Enshassi, S. Mohamed, I. Madi,. "Cost estimation practice in the Gaza Strip: A case study", *The Islamic University Journal, vol*. 15, no. 2, pp. 153-176, 2007.

[28] N. Sinclair, P. Artin, S. Mulford, "Construction cost data workbook", Conference on the International Comparison Program ,World Bank, Washington, D.C., 2002