Indian Journal of Science and Technology, *Vol 10(20), DOI: 10.17485/ijst/2017/v10i20/108567, May 2017*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Forecasting the Cost of Structure of Infrastructure Projects Utilizing Artificial Neural Network Model (Highway Projects as Case Study)

Faiq M. S. AL-Zwainy[1] and Ibraheem Abd-Allah Aidan[2]

[1]College of Engineering, Al-Nahrain University, Iraq; faiqalzwainy@eng.nahrainuniv.edu.iq
[2]Alfarabi University College, Iraq

Abstract

Objectives: The main purpose of this study is to introduce modern technique to using artificial neural network for predicting the cost of structure works for highway project at the feasible study phase. **Methods:** Multi-layer perceptron trainings utilized back-propagation algorithm was used. In this study, the feasibility of ANNs approach for modeling these cost characters was inspected. A lot of problem in relation to ANNs construction such as internal parameters and the effect of ANNs geometry on the performance of ANNs models were inspected. Information on the relative importance of the variable's affecting on the cost parameters predictions was given and mathematical equations in order to estimating the cost of structure works for highway project were determined. **Findings:** One model was developed for the prediction the structure works cost of highway project. Data and information utilized in this model was collected from Stat Commission for Roads and Bridges in republic of Iraq. ANNs model have the ability to predict the cost for structure works for highway project with very good degree of accuracy equal to 93.19% and the coefficient of correlation (R) was 90.026%, **Applications:** Neural network has shows to be a promising approach for use in the initial phase of highway projects when typically only a limited or minus data and incompleted information set is ready for cost analysis.

Keywords: Artificial Neural Network, Cost of Structure Works, Coefficient of Correlation, Highway Project, Predicating

1. Introduction

Cost management of highway projects just as important as the other elements[1], the main advantage of cost management process is to provide direction for the project costs management during the project life. This process is the approach that presents the procedures, documentation, and policies, for planning, managing, expending, and controlling the construction costs[2]. There are many constructed projects fail in time performance, others fail in cost performance and others fail in other performance indicators[3]. Due to lack of data and information in the conceptual phase of the construction projects, therefore the predicating of the construction cost is very difficult[4].

The conceptual phase of the construction project is considered as one of the first phases that deal with costing, in order to help the stakeholders to choose the suitable alternatives for the project within the planned time and quality. It is known that the accuracy of the cost estimate increases as the project life progresses to the availability of accurate data. One of the most important obstacles and difficulties experienced by estimators to estimate costs in the conceptual stage in the Republic of Iraq and from the researcher's point of view are:

- The absence of an integrated database on previously implemented construction projects.
- Do not use modern and advanced tools in estimating the costs of construction projects.
- Lack of interest in engineering alternatives in the planning, design and execution of construction projects.

**Author for correspondence*

One of the earliest endeavors to promote such a database was complete by the World Bank Transport Unit which give in form of ROad Costs Knowledge System (ROCKS) to be applied in grow homelands. It was designed to build an international knowledge system on highway work costs in order to establish an institutional memory, and get mean unit costs based on historical information that could improve the accuracy of new cost estimates and decrease the failure generated by cost overruns[5]. In grow homelands, highway companies can be had such database to help successfully evaluate new highway projects. In order to utilize these information more efficiently and to resolve these problems must be applying artificial intelligent such as Artificial Neural Networks (ANN) and Support Vector Machine (SVM).

For the duration of the last few many years many improvement in computer technology, software program programming and application production had been followed with the aid of diverse engineering disciplines. Those trends are on the whole focusing on synthetic Intelligence strategies[6]. Neural networks are defined as the science that studies mathematical methods that can be formulated based on the simulation of biological cells in living organisms. Neurons are characterized by high speed in processing data and are able to learn and deal with different types of data[7].

Some studies have addressed the matters correlating with cost estimation in the conceptual stage of construction project. Firstly, In[8] developed a approach through building a prototype model for predicating range estimates to evaluate the risk of cost overran in building project by artificial neural networks. . Secondly, In[9] used a artificial neural network model to manage cost data and develop a parametric cost predicting model for highway projects, two alternative tools to train network's weights: Genetic Algorithms (GAs) and simplex optimization. Also, in[10] introduced a regularization neural network model and specified that highway costs are extremely noisy and this noise results from many unpredictable factors such as factors related to weather conditions, market conditions, and experts judgment. In[11] examined different methods of cost estimation models in the conceptual stage of building projects such as Case – Based Reasoning (CBR), Neural Networks (NNs) and Multiple Regression Analysis (MLR). Recently, in[12] utilized the multi-layer perceptron trainings by back-propagation algorithm neural network is formulated and presented for estimation of the productivity of construction projects. In[13] introduced a new and alternative approach of using a neural network for cost estimation at the early stage. One model was built for the prediction the total cost for paving works of highway project. Data used in this model was collected from Stat Commission for Roads and Bridges in Iraq. Multilayer perceptron trainings using the back-propagation algorithm were used. It was found that ANNs have the ability to predict the total cost for paving works of highway project with a good degree of accuracy of the coefficient of correlation (R) was 88.15%, and average accuracy percentage of 84.63%.

Through the previous studies, the researcher reached a number of facts, the most important of which are:

- Artificial neural networks have more efficiency than traditional estimation techniques and have the ability to predict the future in the absence of data and information.
- The possibility of benefiting from artificial neural networks in both developing and developed countries, and in the industrial, financial and other sectors.

Figure 1 displays how artificial neural network check the biases and weights using comparing target with output. Weights are not constant, but they variation over time by earning experience after several iterations[14].

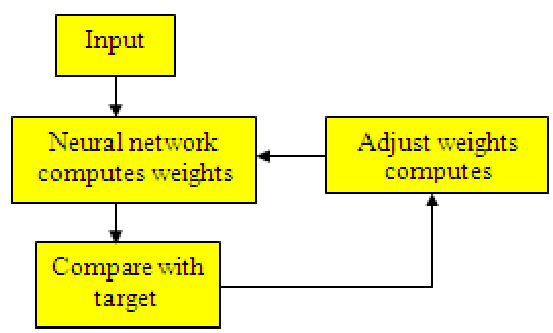

Figure 1. Correction of error using target data[14].

2. Identification of ANN Model Variables

Neural network application for highway cost estimation is an example of causal forecasting. This type of forecasting considers a number of variables that affect the variable to be predicted. This type of forecasting is more powerful than the traditional methods. The purpose of cost estimation is to predict or estimate the cost from known or assumed values of other variables related to it.

This study explained the development of neural network models of total cost structure work of highway project based on historical highway projects data. The initial problem for the research was the shortage of data obtainable that can provide reliable information about the costs. Data used to build the artificial neural network model of cost estimation were past highway data from republic of Iraq. The data collection method used in this study is the direct data gathering from highway Construction Companies and the direct interview with the concerned managers and engineers. This approach faces a large difficulty nowadays because of the security condition in republic of Iraq, and the shortage in documentation. In spite of these obstacles, the researcher succeeded in gathering well trusted data for more than 150 projects through the companies' visits and studies the concerned sheets, reports and documents for highway projects.

The model input variables for this model are the same as those used in model (TC-1)[15], expected (V3, V4, V9, V10, V11, V12, V15, V17, V20). Therefore, the model input variables consist of eleven variables (i.e. V1, V2, V5, V6, V7, V8, V13, V14, V16, V18, and V19). Only two kinds of variables that might effect on the cost estimation of highway project, as follows:

2.1 Objective Variables

This kind of variables consist of eleven variables, as the following

V1: Length of the highway in (km).

V2: Capacity – number of standard lanes.

V5: Number of interchanges.

V6: Estimate year.

V7: Length of Major Bridges, in accumulated linear meter of all bridges greater than 30m. L.

V8: Stream Crossing.

2.2 Subjective Variables

This kind of variables consist of nine variables, as the following

V13: Class- classifies the highways to border, local roads and streets, collector roads, arterial highways, expressways, and assigns them the values 1,2,3,4 and 5 respectively.

V14: Material- this classifies pavement as flexible and rigid. And assigns the values of 1 and 2 respectively to them.

V16: Technology- this defines an index which accounts for technological progress with three types old, existing and highly improved. It assigns them the values of 1, 2 and 3 respectively.

V18: Furnishing- Highway furnishing level; without (1), normal (2), high standards (3).

V19: Drainage- Highway drainage System; absent (1), surface (2), sub-surface (3).

Each one of the variables has been analyzed in order to find the best method of representation in the modeling process.

3. Development of ANN Model

In this study, the researcher used the program (NEUFRAME, Version 4), which is a simple program and easy to use and develops the neural network through the ideal with the possibility of obtaining mathematical formulas in a simple and fully controllable form. Figure 2 display the diagram of the NEUFRAME 4 program

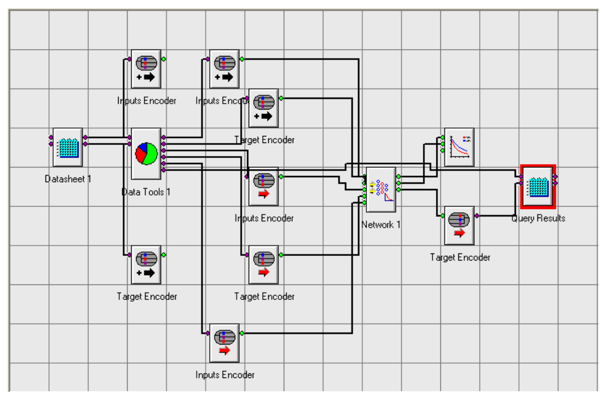

Figure 2. Graphing component of NEUFRAM 4 program.

For the purpose of improving the performance of neural network models there must be an ideal methodology for raising the efficiency of these models and by means of the following[16]:

- Development of model inputs and output.
- Data division and pre-processing.
- Development of model architecture.
- Stopping criteria, and
- Model validation.

4. Model Inputs and Outputs

There are two types of variables used in input model was objective variables and subjective variables. Where eleven variables have the most significant impact on the cost of structure work for highway projects, therefore, the performance of the ANN models affected by using these types of variables. Finally, the output model which was content only one variable's is total cost of structure work for highway project.

5. Pre-Processing and Data Division

One of the most important steps to take into account before starting the process of building the neural network is to address the real data of the inputs or outputs and to work to be homogenous data in order to get rid of statistical problems, and that is through the following:

- Data scaling,
- Normalization and
- Transformation.

In this study, the researcher used the logarithm of total cost structure work of project is taken before introducing forward in the next stages.

When using the Neuframe software, the first step in building a neural network model is to divide the data into three sets: validation sets, testing set and training set. The method used to divide data was trial and error method. In Table 1 it can be seen that the best data subsets division is (70-15-15) % according to highest correlation coefficient of validation set (89.30%) and lowest testing (7.14%) with training error (6.94%).

Table 2 exploring the different choices for split groups (i.e., random striped, and blocked), The best achievement was acquired when the striped division was used, according to training error (7.14%), lowest testing (6.94%) with highest coefficient of correlation to validation set (89.30%).

6. Scaling of Data

When the data is divided into three sets, namely the training set, the test set and validation set, it is necessary consider treatments both input and output by converting real values to standard values according to the transfer function in the hidden and output layers (i.e., 0.0 to 1.0 for sigmoid transfer function and –1.0 to 1.0 for tanh transfer function). Note that the most common method[16], as follows:

$$\mathbf{x_n} = \frac{\mathbf{x} - \mathbf{x_{min}}}{\mathbf{x_{max}} - \mathbf{x_{min}}} \qquad (1)$$

7. Model Architecture, Optimization and Stopping Criteria

The design of artificial neural network architecture is considered one of the most complex and difficult tasks in building intelligent neural networks. In order to obtain the best performance of the artificial neural network, there must be continuous training of the neural network

Table 1. Effect of data division on performance of ANNs

Data Division			training error%	testing error%	coefficient correlation(r)%
Training%	Testing%	Querying%			
62	22	16	8.90	9.00	77.35
60	24	16	7.23	8.40	70.88
60	18	22	7.85	7.57	76.59
64	20	16	7.40	7.48	68.64
56	30	14	7.36	7.44	77.64
68	14	18	7.56	7.35	79.64
64	18	18	7.55	7.21	81.26
66	17	17	7.49	7.10	86.30
68	12	20	6.75	6.98	81.42
65	15	20	7.18	6.98	81.52
70	15	15	7.14	6.94	89.30
66	14	20	6.87	6.94	81.41

Table 2. Effects of method division on ANNs performance

Data Division%			choices of division	training error%	testing error%	coefficient correlation(r)%
Training	Testing	Querying				
70	15	15	Striped	7.14	6.94	89.30
70	15	15	Blocked	8.88	9.88	69.89
70	15	15	Random	8.06	8.38	70.40

Table 3. Effects No. of nodes on ANNs performance

Model No.	Parameters Effect	No. of Nodes	training error%	testing error%	coefficient correlation(r)%
TCSW-1	**choices of division** (Striped) **Learning Rate** (0.2) **Momentum Term** (0.8) **Transfer function in hidden layer** (Sigmoid) **Transfer function in output layer** (Sigmoid)	1	7.14	6.94	89.30
TCSW -2		2	7.41	6.80	88.36
TCSW -3		3	7.31	6.95	88.59
TCSW -4		4	7.57	6.92	85.60
TCSW -5		5	7.38	6.80	85.94
TCSW -6		6	7.20	6.83	86.53
TCSW -7		7	7.45	6.80	85.67
TCSW -8		8	7.34	6.97	87.24
TCSW -9		9	7.44	6.99	86.44
TCSW -10		10	7.54	6.82	85.98
TCSW -11		11	7.44	6.99	85.79
TCSW -12		12	7.42	6.99	85.65
TCSW -13		13	7.49	6.99	85.47
TCSW -14		14	7.51	6.97	86.66
TCSW -15		15	7.59	6.98	86.48
TCSW -16		16	7.57	6.93	86.73
TCSW -17		17	7.55	6.97	85.65
TCSW -18		18	7.56	6.93	85.69
TCSW -19		19	7.57	6.95	85.75
TCSW -20		20	7.57	6.95	86.00
TCSW -21		21	7.57	6.95	86.00
TCSW -22		22	7.57	6.95	86.00
TCSW -23		23	7.57	6.95	86.00

in order to obtain the lowest testing error followed by training error and high correlation coefficient of validation set, by achieving the optimal characteristics of the neural network, such as learning rates, number of nodes, transfer functions, and momentum terms. In order to obtain the best number of hidden nodes in the hidden layer, the default parameters were used for the program which is the learning rate equal to 0.2 and momentum term equal to 0.8. The researcher then conducted a number of experiments with a regular increase of the number of nodes in the hidden layer, as shown in Table 1 and Figure 1,It can be seen that the Model 2 with two hidden nodes have the lowest prediction error. So, it was selected in this model.

Two hidden node was chosen in model TCSW-2 as shown in Figure 3, where the neural network have two hidden node with lowest prediction error for the testing test (6.80%), training test (7.41%) and validation test (88.36%), therefore the neural network is considered best or most favorable.

Table 4. Effects momentum term on ANNs performance (Model TCSW-2)

Parameters Effect	Momentum Term	training error%	testing error%	coefficient correlation(r)%
Model No. (TCSW-2) **choices of division** (Striped) **Learning Rate** (0.2) **No. of Nodes** (2) **Transfer function in hidden layer** (Sigmoid) **Transfer function in output layer** (Sigmoid)	0.01	7.71	6.86	84.93
	0.05	7.66	6.85	85.11
	0.1	7.64	6.85	85.22
	0.2	7.58	6.84	85.76
	0.3	7.47	6.83	86.37
	0.4	7.46	6.83	84.85
	0.5	7.40	6.82	85.95
	0.6	7.35	6.82	85.47
	0.7	7.29	6.82	86.29
	0.8	7.41	6.80	88.36
	0.9	7.50	6.82	88.04
	0.95	7.56	6.82	90.33

Table 5. Effects learning rate on ANNs performance (Model TCSW-2)

Parameters Effect	Learning Rate	training error%	testing error%	coefficient correlation(r)%
Model No. (TCSW-2) **choices of division** (Striped) **Momentum Term** (0.8) **No. of Nodes** (2) **Transfer function in hidden layer** (Sigmoid) **Transfer function in output layer** (Sigmoid)	0.02	8.28	7.08	82.94
	0.05	7.58	6.83	85.65
	0.1	7.58	6.83	85.97
	0.15	7.20	6.80	88.58
	0.2	7.41	6.80	88.36
	0.3	7.41	6.87	87.78
	0.4	7.43	6.88	88.76
	0.5	7.40	6.90	89.22
	0.6	7.39	6.93	89.76
	0.7	7.40	6.88	89.57
	0.8	7.20	6.87	90.12
	0.9	7.65	6.86	90.20

Table 4 and Figure 4 shows the effect of the momentum term on the performance of the model TCSW-2, where learning rate equal to 0.20, and the best favorable value for momentum term equal to 0.8, with lowest prediction error for the testing test (6.80%), training test (7.41%) and validation test (88.36%)

Table 5 and Figure 5 explained the effect of the learning rate on the ANN model (Model TCSW-2) performance, when momentum term equal to 0.8, the result indicated the optimum value for learning rate equal to 0.15, which have smallest prediction error

Table 6 shown the effects of employ varied transfer functions (i.e., tanh and sigmoid), results indicated for ANN model was comparatively insensitive to the kind of the transfer function. optimal performance for ANN model when have lowest prediction error coupled with highest correlation coefficient (r), where the tanh transfer function was used for hidden and output layers.

To ensure the data that was carried out by ***Neuframe*** software for testing, training, and validation sets represents the oneself statistical population, a statistical parameters estimation were carried out, including the average, range, maximum, minimum, and standard deviation, as shown in Table 7. The results indicated the testing, training, and validation sets are generally statistically strong.

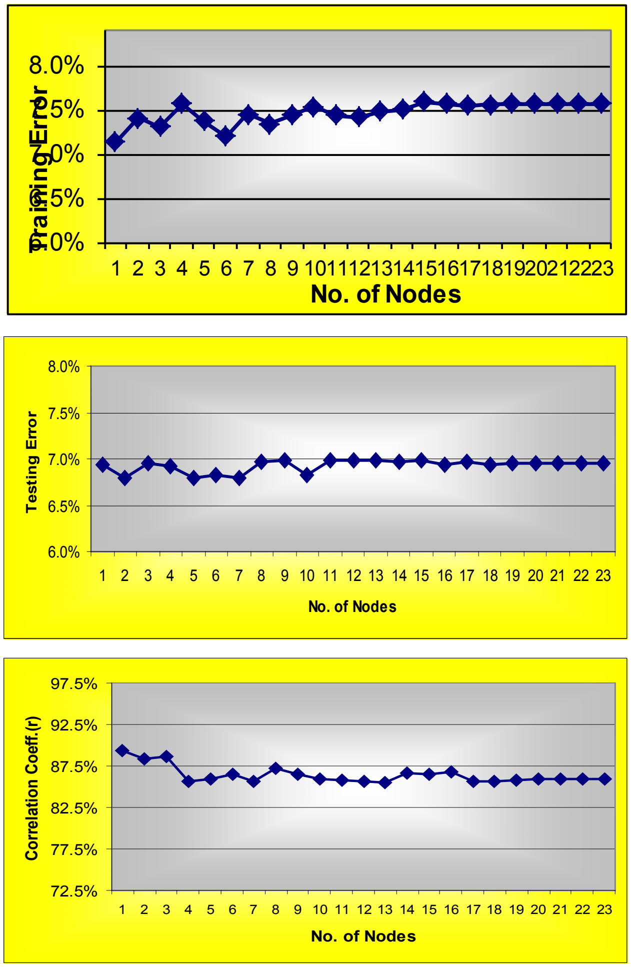

Figure 3. Performance of ANNs model with different hidden nodes (Model TCSW-2).

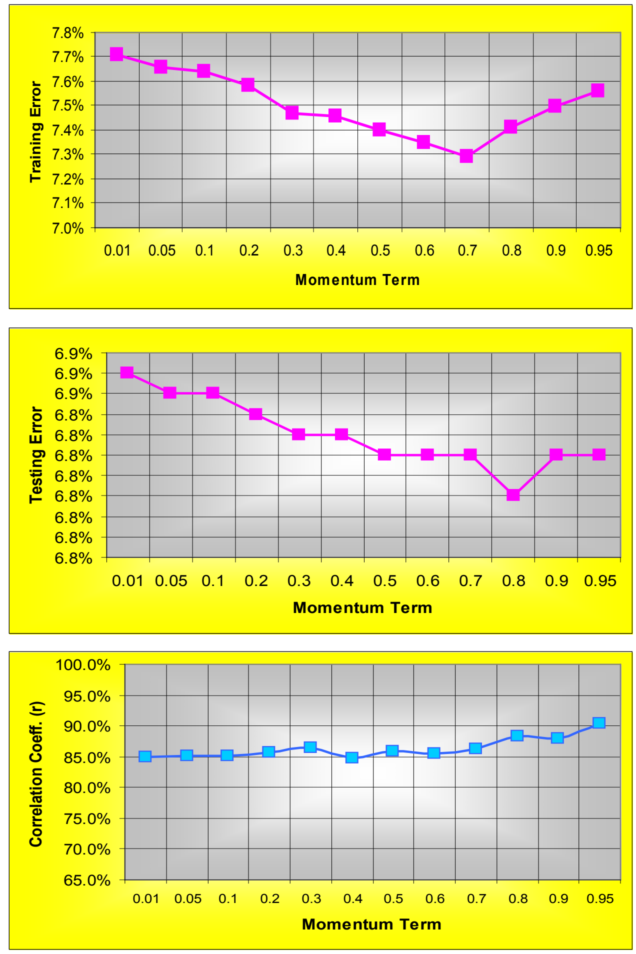

Figure 4. Effect of various momentum term on ANNs performance (learning rate = 0.2).

Table 8 shows the results of the t-test, where these results indicate that testing, training, and validation sets are representative of stylish statistical population.

Table 6. Effects of transfer function on ANNs performance

Parameters Effect	Transfer Function		training error%	testing error%	coefficient correlation(r)%
	Hidden Layer	Output Layer			
Model No. (TCSW-2) **choices of division** (Striped) **No. of Nodes** (2) **Momentum Term** (0.8) **Learning Rate** (0.15)	sigmoid	sigmoid	7.20	6.80	88.58
	sigmoid	tanh	7.78	6.78	89.66
	tanh	sigmoid	7.24	6.73	88.83
	tanh	tanh	7.58	6.48	90.03

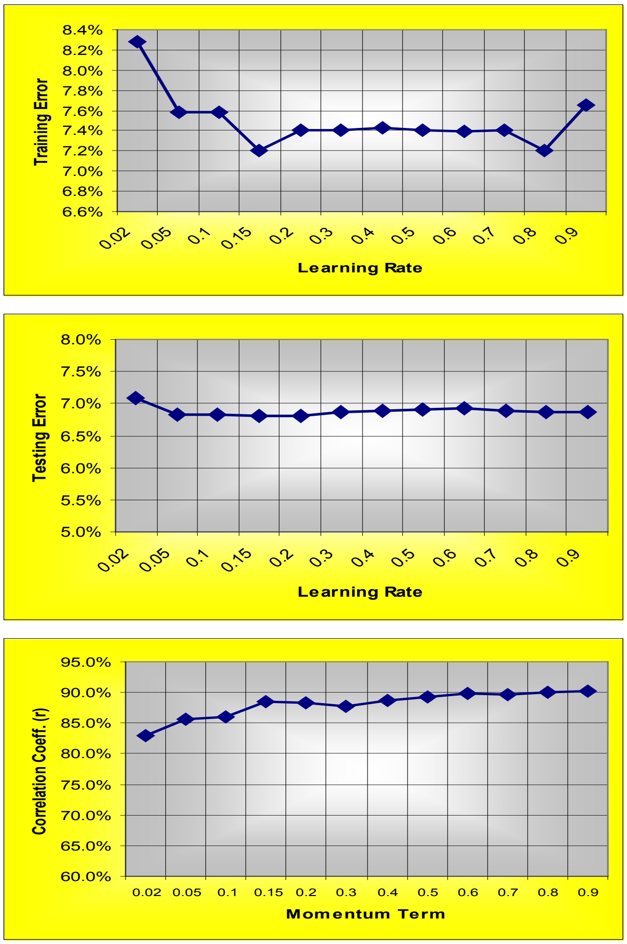

Figure 5. Effect of various learning rate on ANNs performance (momentum term = 0.8).

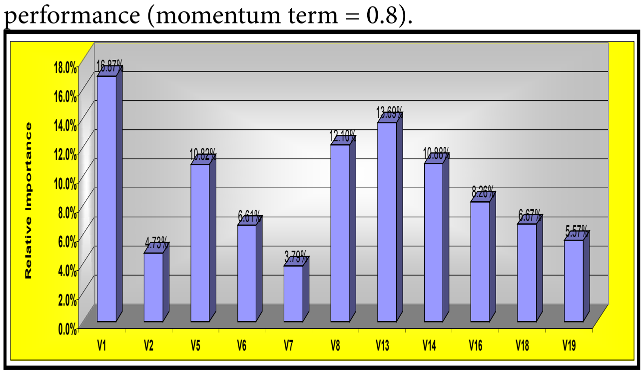

Figure 6. Relative importance of the input variables for model (TCSW-2).

8. Relative Importance of the ANN Model

Sensitivity analysis can be used to determine which of the input variables have the most significant impact on the output predictions for ANN model (model 2 (TCSW-2)). The researcher utilized an easy and innovative approach[17] to explained the relative importance of the input variables by testing connection weights of trained neural network as shown in Table 9 .

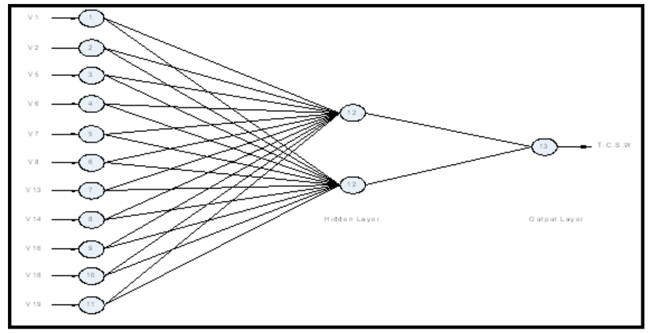

Figure 7. Structure of the ANNs optimal model (TCSW-2).

Figure 6 shows the results sensitivity analysis, where the variables (V1,V13 and V8) has the most significant effect on the predicted cost of structure works for highway projects with a relative importance (16.87%, 13.69 and 12.10%) respectively. Followed by V14 and V5 with a relative importance of 10.88% and 10.82 respectively. Results also indicate that (V16, V18 and V6) have a temperate impact on prediction with a relative importance equals to 8.26%, 6.67% and 6.61% respectively, while the variables (V2 and V7) have the smallest effect on prediction with relative importance of 4.73% & 3.79% respectively.

9. Equation of Artificial Neural Network Model

Neuframe v.4 can be used to obtain a number of connection weights for the optimal artificial neural network model (Model TCSW-2), these weights which enables the network to be converted into easy and simple equation. Figure 7 demonstrate the structure of the ANN model, while Table 10 show the weights and threshold levels (bias) of neural network for developer's model. The equation of forecasting the cost of structure works for highway projects can be obtained by using the threshold levels and connection weights shown in Table 10, as follows:

$$TCSW = \frac{1}{1 + e^{(-0.29288 + 0.10052 \tanh x_1 + 0.93419 \tanh x_2)}} \quad (2)$$

Where:

Table 7. Input and output statistics for the ANNs (Model TCSW-2)

Statistical parameters	Output	Input Variables										
	log (cost)	V1	V2	V5	V6	V7	V8	V13	V14	V16	V18	V19
	Training set											
MAX	5.63347	602	10	30	1995	600	180	5	2	3	3	3
MIN	1.69897	0.4	2	0	1980	0	0	1	1	1	1	1
RANGE	3.9345	601.6	8	30	15	600	180	4	1	2	2	2
MEAN	3.539	49.625	3.8113	1.4057	1988.78	26.802	7.6792	2.2453	1.0755	1.9151	1.7264	1.387
S.D	0.69017	71.724	1.6683	4.6266	2.74297	95.793	27.434	0.9443	0.2654	0.4804	0.61	0.562
	Testing set											
MAX	5.10721	176	8	28	1995	570	110	5	2	3	3	3
MIN	2.68842	7	2	0	1979	0	0	1	1	1	1	1
RANGE	2.41879	169	6	28	16	570	110	4	1	2	2	2
MEAN	3.64476	60.191	3.7273	2.5455	1988.32	46.818	12.182	2.4545	1.0909	1.9545	1.7273	1.409
S.D	0.61829	56.388	1.8818	6.9606	3.3006	148.4	30.455	1.1434	0.2942	0.4857	0.6311	0.59
	Validation set											
MAX	5.4624	246.8	8	27	1995	612	154	5	2	3	3	2
MIN	2.25285	3	2	0	1981	0	0	1	1	1	1	1
RANGE	3.20954	243.8	6	27	14	612	154	4	1	2	2	1
MEAN	3.62283	74.967	3.5238	2.5714	1988.76	60.762	12.667	2.381	1.0952	1.9524	1.7143	1.333
S.D	0.75996	68.982	1.7782	7.5005	2.62497	175.51	38.6	1.0713	0.3008	0.4976	0.6437	0.483

Table 8. t-test for the ANNs input and output variables (Model TCSW-2)

Variables	Output	Input Variables										
	log (cost)	V1	V2	V5	V6	V7	V8	V13	V14	V16	V18	V19
Data set	Testing											
t-value	0.708580408	0.170022	-0.81539	-0.28596	1.311447	-0.25277	0.152604	-0.48151	-0.50073	-1.87151	-1.28188	-1.01411
Lower critical value	-1.98027223	-1.98027	-1.98027	-1.98027	-1.98027	-1.98027	-1.98027	-1.98027	-1.98027	-1.98027	-1.98027	-1.98027
Upper critical value	1.980272226	1.980272	1.980272	1.980272	1.980272	1.980272	1.980272	1.980272	1.980272	1.980272	1.980272	1.980272
t-test	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
Data set	Validation											
t-value	-0.03199057	0.560925	-0.8125	-1.45584	0.542826	-1.58979	-1.48727	-0.99788	-1.81205	-0.54429	-0.92275	-0.27783
Lower critical value	-1.97897058	-1.97897	-1.97897	-1.97897	-1.97897	-1.97897	-1.97897	-1.97897	-1.97897	-1.97897	-1.97897	-1.97897
Upper critical value	1.978970576	1.978971	1.978971	1.978971	1.978971	1.978971	1.978971	1.978971	1.978971	1.978971	1.978971	1.978971
t-test	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept

$X_1 = \{\theta 12 + (w_{12\text{-}1} * V1) + (w_{12\text{-}2} * V2) + (w_{12\text{-}3} * V5) + (w_{12\text{-}4} * V7) + (w_{12\text{-}5} * V7) + (w_{12\text{-}6} * V8) + (w_{12\text{-}7} * V13) + (w_{12\text{-}8} * V14) + (w_{12\text{-}9} * V16) + (w_{12\text{-}10} * V18) + (w_{12\text{-}11} * V19)\}$ (3)

$X_2 = \{\theta 13 + (w_{13\text{-}1} * V1) + (w_{13\text{-}2} * V2) + (w_{13\text{-}3} * V5) + (w_{13\text{-}4} * V7) + (w_{13\text{-}5} * V7) + (w_{13\text{-}6} * V8) + (w_{13\text{-}7} * V13) + (w_{13\text{-}8} * V14) + (w_{13\text{-}9} * V16) + (w_{13\text{-}10} * V18) + (w_{13\text{-}11} * V19)\}$ (4)

In order to using the Equations (3) and (4), must be converted all input variables (i.e., V1, V2, V5, V6, V7, V8, V13, V14, V16, V18, and V19) into scaling values between 0.0 and 1.0 using Equation (1) and also using data in Table 7 and for only the ANN model training. But the Equation (2) must be converted from scaling value into actual value to predicted value for cost of structure works for highway projects, as follows:

$$TCSW = \frac{3.9345}{1 + e^{(-0.29288 + 0.10052 \tanh x_1 + 0.93419 \tanh x_2)}} + 1.6989 \qquad (5)$$

And

$X_1 = \{17.088 + (0.00071*V1) + (0.02195*V2) + (-0.009*V5) + (-0.008*V6) + (0*V7) + (-0.002*V8) + (-0.143*V13) + (-0.511*V14) + (0.1479*V16) + (-0.0865*V18) + (0.144*V19)\}$ (6)

$X_2 = \{46.8954 + (-0.0013*V1) + (-0.0186*V2) + (-0.017*V5) + (-0.023*V6) + (0*V7) + (-0.003*V8) + (-0.0013*V13) + (-0.216*V14) + (-0.1375*V16) + (0.149*V18) + (0.0393*V19)\}$ (7)

Table 9. Connection weights between hidden node and input layer and output layer

Hidden Nodes	Weights					
	V1	**V2**	**V5**	**V6**	**V7**	**V8**
Hidden 1	0.42750	0.17563	-0.26178	-0.12562	-0.00157	-0.35379
Hidden 2	-0.76698	-0.14931	-0.50644	-0.35024	-0.28470	-0.49468
Hidden Nodes	Weights					
	V13	**V14**	**V16**	**V18**	**V19**	**output**
Hidden 1	-0.57164	-0.51102	0.293994	-0.17278	0.288731	-0.10052
Hidden 2	-0.35638	-0.21573	-0.27494	0.298744	0.078543	-0.93419

Table 10. Weights and threshold levels for the ANNs optimal model (Model TCSW-2)

Hidden layer nodes	**w_{ji} (weight from node i in the input layer to node j in the hidden layer)**					**Hidden layer threshold θ_j**
	i=1	**i=2**	**i=3**	**i=4**	**i=5**	
j=12	0.42750	0.17563	-0.26178	-0.12562	-0.00157	0.103748
	i=6	i=7	i=8	i=9	i=10	
	-0.35379	-0.57164	-0.51102	0.293994	-0.17278	
	i=11					
	0.288731					
j=13	i=1	i=2	i=3	i=4	i=5	0.40337
	-0.76698	-0.14931	-0.50644	-0.35024	-0.28470	
	i=6	i=7	i=8	i=9	i=10	
	-0.49468	-0.35638	-0.21573	-0.27494	0.298744	
	i=11					
	0.078543					
Output layer nodes	w_{ji} (weight from node i in the hidden layer to node j in the output layer)					Output layer threshold θ_j
	i=12	i=13	-	-	-	
j=14	-0.10052	-0.93419	-	-	-	0.29288

10. Validation of Artificial Neural Network Model

In this study, the researcher used a set of statistical measures to prove the validity of the developed artificial neural network model[18,19,20,21].

1. Mean Percentage Error (MPE);

$$MPE = \{\sum_{i=1}^{n}\left(\frac{A-E}{A}\right)/n\}*100\% \tag{8}$$

Where:
A= total cost
E= predicted cost
n= number of projects (30 for validity set).

2. Root Mean Squared Error (RMSE);

$$RMSE = \sqrt{\frac{\sum_{i-1}^{n}(E-A)^2}{n}} \tag{9}$$

3. Mean Absolute Percentage Error (MAPE),

$$MAPE = \left(\sum_{i=1}^{n}\frac{|A-E|}{A}*100\%\right)\Big/ n \tag{10}$$

4. Average Accuracy Percentage (AA %)

$$A\ \% = 100\% - MAPE \tag{11}$$

5. Correlation Coefficient (R)
6. Determination Coefficient (R^2).

Table 11 shows the results of the statistical measures used to prove the validity of the neural network. MAPE % equal to 6.81% and AA% equal to 93.19%, these results suggest that the neural network model (TCSW-2) shows very good compatibility with actual costs for highway projects.

Table 11. Results of the comparative study

Description	ANN for Model TCSW -2
MPE	-2.99%
RMSE	0.30772
MAPE	6.81%
AA %	93.19%
R	90.026%
R^2	81.05%

Table 12. Error categorization (%)[22].

MAPE		
Good	Fair	Poor
Less than 25	25-50	More than 50

Schexnaydr and Mayo[22] proposed the error of estimation at the planning stage was approximately from -25% to +25%. In this research, MAPE% can be depended as an error categorization, according to Table 12, MAPE of model TCSW-2 was very good.

In Figure 8, shows the comparison of predicted and observed total cost structure work for validation data, therefore , the correlation between the actual cost value and the predicted costs can be considered a strong, where, the determination coefficient equal to 81.05|%. This indicates that ANN model have a very good correlation between the actual cost and predicted cost.

11. Conclusion

The main conclusions can be summarized as follows:

- Models of artificial neural networks are an important and useful technique that can be used in the early stages of the life of highway projects for the purpose of forecasting costs, especially when data are scattered, inaccurate and incomplete
- The artificial neural network architecture developed in this study consists of three layers, namely the input layer, It includes eleven variables and the hidden layer, It contains two hidden nodes and the output layer, It consists of one variable, which is the cost of structure works for highway projects. The results included the following:
- Mean Percentage Error (MPE) equal to -2.99%.
- Mean Absolute Percentage Error (MAPE) equal to 6.81%.
- Average Accuracy Percentage (AA %) equal to 93.19%.
- Determination Coefficient (R2) equal to 81.05%.
- Correlation Coefficient (R) equal to 90.026%.
- With this models, the effects of cost-related parameters on the total cost of highway construction projects can be investigated through its sensitivity analysis procedure
- ANNs models could be translated into a simple and practical formula from which (TCSW-2) may be calculated.

12. References

1. Faiq MA, Reem A. Investigation cost deviation of highway project. Research Journal of Applied Sciences, Engineering and Technology. 2016; 13(11):843–55. Crossref
2. Faiq MSA, Reem A, Tareq K. Reviewing of the simulation models in cost management of the construction projects. Civil Engineering Journal. 2016; 2(11):607–22.
3. Faiq MSA, Mohammed SKhA. Investigation and evaluation the performance of the construction residential complexes projects. International Journal of Applied Engineering Research. 2016; 11(16):8863–77.
4. Al-Zwainy FMS, Hadhal NT. Investigation and evaluation of the cost estimation methods of iraqi communication projects. IJEMR. 2015; 5(6):41–8.
5. ROCKS. Road Costs Knowledge System. The World Bank, Transport Unit – TUDTR. Available from: http://www.worldbank.org/transport/roads/rd_tools/rocks_main.htm
6. Al-Zwainy FMS, Hadhal NT. Application artificial forecasting techniques in cost management (review). Journal of Engineering. 2016; 22(8):1–15.
7. Al-Zwainy FMS. Development of the mathematical model for predicating the construction productivity in iraq using the artificial neural perceptron network. Journal of Engineering and Development. 2014; 18(2):1–21.
8. Pearce AR, Gregory RA, Williams L. Range estimating for risk management using artificial neural networks. Journal of Parametric Estimating. 1996; 14(1):88–93.
9. Hegazy T, Amr A. Neural network model for parametric cost estimation of highway projects. Journal of Construction Engineering and Management. ASCE. 1998; 124(3):210–8. Crossref
10. Adeli H, Wu M. Regularization neural network for construction cost estimation. Journal of Construction Engineering and Management. ASCE. 1998; 124(1):18–24. Crossref
11. Gwang-Hee K, Sang-Hoon A, Kyung-In K. Comparison of construction cost estimating models based on regression analysis, neural networks and case-based reasoning. Journal Building and Environment. 2004; 39(1):1235–42.
12. Al-Zwainy FM, Hatem A, Huda F. Development of the construction productivity estimation model using artificial neural network for finishing works for floors with marble. ARPN Journal of Engineering and Applied Sciences. 2012; 7(6):714–22.
13. Al-Suhaili RH, Saco ZM, Al-Zwainy FM. Estimating the cost of paving the highway construction project by using artificial neural network. The 2nd Scientific Conference of College Engineering; Iraq. 2009. p. 846–67.
14. Demuth H, Beale M. Hagan M. Neural network toolbox user's guide. Natick, Massachusetts; 2006.
15. Al-Zwainy FM. The use of artificial neural network for estimating total cost of highway construction projects [Doctoral diss]. College of Engineering, Baghdad University, Iraq; 2009.
16. Al-Zwainy FMS, Al-Suhaily RH, Saco ZM. Project management and artificial neural networks: Fundamental and application. LAP LAMBERT Academic Publishing; 2015.
17. Shahin MA, Jaska MB, Maier HR. Predicting settlement of shallow foundations using neural networks. Journal of Geotechnical and Geoenviromental Engineering, ASCE. 2002; 128(9):785–93. Crossref
18. Al-Suhaili RH, Saco ZM, Al-Zwainy FM. Using the artificial neural networks for prdicting the total cost of highway projects in Iraq. The 2nd Regional Conf for Eng Sci College of Eng; Iraq. 2010. P. 846–67.
19. Al-Zwainy F, Hadhal NT. Building a mathematical model for predicting the cost of the communication towers projects using multifactor linear regression technique. International Journal of Construction Engineering and Management. 2015; 05(1):25–9.
20. Al-Zwainy FM, Rasheed HA, Ibraheem HF. Development of the construction productivity estimation model using artificial neural network for finishing works for floors with marble. ARPN Journal of Engineering and Applied Sciences. 2012; 7(6):714–22.
21. Al-Zwainy FMS, Abdulmajeed MH, Aljumaily H. Using multivariable linear regression technique for modeling productivity construction in Iraq. OJCE. 2013; 3(3):127–35. Crossref
22. Schexnayder CJ, Mayo E. Construction management fundamentals. Boston, MA: McGraw-Hill Higher Education; 2003. PMCid:PMC1693145.