Civil and Environmental Engineering
Vol. 20, Issue 2, 1024-1039, DOI: 10.2478/cee-2024-0074

FORECASTING MODELS FOR TIME AND COST PERFORMANCE PREDICTING OF INFRASTRUCTURAL PROJECTS

Saja Hadi Raheem ALDHAMAD[1*] , Rana MAYA[2], Suha Falih Mahdi ALAZAWY[3], Faiq M. ALZWAINY[4]

[1]Department of Civil Engineering, College of Engineering , Al-Iraqia University ,Baghdad, Iraq.

[2]Department of Construction Engineering and Management, Faculty of Civil Engineering, Tishreen University, Lattakia, Syria.

[3]Department of Highway and Airport Engineering, College of Engineering, University of Diyala, Ba'aqubah , Diyala, Iraq

[4]Forensic DNA Centre for Research and Training, Al-Nahrain University, Jadriya, Iraq.

[*] corresponding author: saja.h.raheem@aliraqia.edu.iq

Abstract

Developing artificial forecasting models to predict the cost index and time index is the primary objective of this research. The efficacy of residential property investment projects will be assessed through the implementation of models that incorporate Artificial Neural Networks and Multiple Linear Regression. Historical information of thirteen boundaries for twenty finished Private Property Venture Tasks were separated from the records of the Directorate of Lodging, then four models were created by utilizing Multiple Linear Regression strategy and Artificial Neural Networks method. The main model was Time Index model with test information that gave average accuracy equivalent to 78.48 % and delivered a high connection coefficient (R) equivalent to 95.3 % utilizing (MLR), the subsequent model was (CI) model with test information gave (AA) equivalents to 66.6 % and created (R) equivalents to 85.4 % utilizing (MLR), and the third model was (TI) model with test information gave (AA) equivalents to 90.31 %, and delivered a high (R) equivalents to 95.4 % utilizing (ANN), the fourth model was (CI) model with test information gave (AA) equivalents to 90.21 % and delivered a high (R) equivalents to 99.6 % utilizing (ANN). It was found that the NN model showed more exact assessment results than the MLR models. Thus, it was resolved that the NN model is generally appropriate for anticipating the time execution and cost execution of Residential Property Investment Projects.

Keywords:

Forecasting;
MLR;
ANNI;
Residential Property Investment;
Projects;
TI.

1 Introduction

One of the most important economic sectors, construction significantly increases the gross domestic product of most countries. During this stage of civil engineering, the conceptual ideas of the planner and the detailed plans of the designer are transformed into tangible physical structures.. The effectiveness of this task depends on how well all parties involved in the project work together. The aim of the construction teams is to be able to finish the project on time, under budget, with no disagreements, rework, lost time, or any other issues, and to everyone's satisfaction. As a result, construction projects need to be more productive, and every stage of management needs to be done seriously. [1].

Evaluation of the performance of construction projects offers a way to assess current and future usage trends for project facilities while also learning from the past. As a result, it is thought that gathering, interpreting, and analyzing data on project performance holds the key to more effective future planning

© Author(s) 2024. This work is distributed under the Creative Commons BY 4.0 license (https://creativecommons.org/licenses/by/4.0/).

and design [4]. whereas performance assessment enables one to assess the project's quality, determine which goals have been met and to what extent, highlight its strengths and weaknesses, pinpoint challenges that hampered its development that year, provides ideas for the following year's improvements, create a clear portfolio for the report, and increase the project's visibility. The majority of residential complex projects in Iraq are not finished within the time and budget allotted in the contracts, which causes issues with project execution. A forecasting model was created to help in this situation so that the concerned agencies may predict the performance of ongoing residential complex construction projects in terms of their Time Index (TI) and Cost Index (CI). [5].

A comprehensive literature analysis was undertaken, examining journals and theses from local and worldwide databases, to assess the various approaches and strategies employed in evaluating and measuring performance for building projects.As per [2] study on which analyzes 17 different forecast model examinations led beginning around 1971 in the cost building field and presented forecast models with different boundaries in different construction cost and risk spaces .The different models where utilized for cost forecast are: ; ML = machine learning; ANN = artificial neural network; TS = time series; RA = regression analysis; MLR = multi-linear regression; SA = statistical analysis; QUN = quantitative data; QUL = qualitative data; EVM = earned value management; NS = not specified. Therefore, Multi leaner regression (MLR), Artificial Neural networks (ANN), were adopted to achieve They demonstrated that they are models of high prediction accuracy, which was the goal of the study [4, 5].

Based on data from a computer experiment utilizing regression models, Mackovaa [3] provided a prediction model for construction time estimation for residential buildings in Slovakia. The prediction model has the maximum predictive ability R Square value of 0.808. Since there hasn't been any prior study done on forecasting models for residential complex building projects, The primary aim of this study is to develop artificial forecasting models for the purpose of evaluating the performance of these projects through the prediction of the time index and cost index. It is imperative to comprehend the factors that impact time and cost performance in order to arrive at this understanding. Consequently, the investigator in this research endeavour is to develop and evaluate time and cost index models by employing the methodologies delineated subsequently.:

1) Identify the factors that influence the time and cost indices.
2) Create and assess the suggested MLR models
3) Create and assess the suggested ANN models.
4) Examine modeling approaches

2 Identification of MLR Model Variables

Thirteen carefully chosen and precisely defined explanatory factors were employed for every building project in the Housing Directorate based on the data sheet findings used to monitor the work's progress and field investigations. The explanatory factors can be categorised into two distinct groups: objective variables and subjective variables, as seen in Tables (1) and (2) provided below. The parameters that were chosen for this study have also been identified and utilised as explanatory variables in previous research. The user has provided two references, numbered 2 and 11.

The manners in which these variables were depicted can be categorized into two distinct groupings.

Table 1: The Objective variables

Variables	Description	Units
Y	Year	Year (2003, 2004, etc.)..
UOC	Complexes units	Number
EPM	Project manager experience	No. of Years
COC	Complex Cost	Currencies
DOC	Duration of complex.	No. of days
NCD	No. of changes in design.	No. of times.

Table (2) The Subjective Variables

variables	Description	Units

P	.Position	City.
S	.Sector	1=puplic, 2=privat.e
TOC	.Type of complex	1= house, 2=fla.t
C	.Type of contract.	1=BOQ, 2=C+, 3 =turnkey.
OWL	.Organizing of the Work Location	1=good, 2=medium ,3=poor.
COS	.Conditions of Security	1=good, 2=medium ,3=poor.
COW	.Conditions of Weather	1=good, 2=medium ,3=poor

3 Multiple Linear Regressions (MLR)

Regression analysis is a statistical technique that is widely employed for the purposes of prediction and analysis in several academic disciplines. The primary focus lies in examining the association between a response variable and one or more explanatory factors. and it covers numerous approaches for modeling and analyzing multiple variables. Additionally, regression analysis enables one to comprehend how, when any one of the explanatory variables is modified, the usual value of the criterion variable varies while the other explanatory variables are kept constant. [8, 9].

4 Development of (MLR) Model

It is possible to use a backward regression to determine the significance of each explanatory variable. The variable stays in the model if it is significant. However, it is taken out and a new regression analysis is performed. [10].

In Iraq, there are 75 residential complexes spread throughout various cities, of which 20 are finished and 55 are in the planning stages. Ten (10) residential complexes under construction and twenty (20) finished residential complexes, as shown in Appendix A (Table A), were randomly selected in order to validate the time index model and the cost index model. In order to verify the time index model, ten (10) residential complexes that were still under construction were also selected at random.

4.1 Time Index Model

Multiple linear regressions were utilized in this case to ascertain the statistical link between a response (in this case, the time index) and the explanatory factors (in this case, position, units of complex, etc.). The variable removed and the effective variable were included in the first run of the SPSS v20 program. Consequently, the variable deemed most significant was identified as the effective variable, and the critical t-value, calculated at a significance level of 5%, is derived as follows: The calculated t-value at a significance level of 0.05, with degrees of freedom (n-p-1) equal to 1.943, is obtained using the given values of n (20) and p (13). To assure the accuracy of the succeeding model execution in a backward stepwise regression analysis that relies on t-values, the variable with the lowest t-value will be excluded. The principal aim of the preliminary iteration of the regression analysis is to ascertain which variables ought to be omitted from the model or which have a substantial impact on it. Based on Model (1), the variable Y will be excluded from the analysis due to its minimal value of (t), which is below the critical t value obtained from the t-table distribution. The variables OWL, DOC, COW, P, and COC were excluded from the succeeding model run due to their minimum values of (t) being lower than the corresponding t value obtained from the t-table distribution for each run. A backward regression analysis was conducted. The obtained results indicate that the ANOVA statistics for the regression analysis, as presented in Table 3, reveal a highly significant F-statistic (F=18.066). When the value of F is greater than the critical value F(p, n-p-1), the statistical F-test suggests that there is a significant difference or relationship. In the given scenario, when the sample size (n) is 20 and the degrees of freedom (p) is 13, the critical value of the F-distribution at a significance level of 0.05 and with 13.6 degrees of freedom is determined to be 3.97.

Table (3): Results of analysis of variance (ANOVA) for time index regression.

Model		Sum of Squares	df	Mean Square	F
1	Regression	0.923	13	-0.190	0.747
	Residual	-0.159	06	-0.936	
	Total	0.764	19		
2	Regression	0.923	12	-0.156	0.848
	Residual	-0.159	07	-1.004	
	Total	0.764	19		
3	Regression	0.923	11	-0.119	0.939
	Residual	-0.154	08	-1.056	
	Total	0.769	19		
4	Regression	0.922	10	-0.078	1.024
	Residual	-0.148	09	-1.102	
	Total	0.774	19		
5	Regression	0.922	09	-0.032	1.112
	Residual	-0.144	10	-1.143	
	Total	0.778	19		
6	Regression	0.921	08	0.018	1.194
	Residual	-0.134	11	-1.174	
	Total	0.787	19		
7	Regression	0.919	07	0.073	1.257
	Residual	-0.104	12	-1.180	
	Total	0.815	19		

First MLR model: - Time index (TI) (Model No. 1)

TI = -1.614 +0.363S + 2.64 TOC – 0.009 UOC + 0.154 C– 0.037 EPM + 0.068 NCD + 0.432COS (1)

4.1.1 Evaluation of the Validity of MLR Model (TI Model.1):

We used the statistical equations to check if the equation (-1) was valid for the TL model in the setting of multiple linear regression (MLR).The models' effectiveness is evaluated using the statistical measures that are labeled as [11, 12]. Mean Absolute Percentage Error (MAPE);

MAPE=(∑(│a-e│)/a*100%)/N (2)

1. Two measurements of average error are the MAPE and the percentage RMSE.Average Accuracy Percentage (AA%)

AA%= 100%-MAPE (3)

2. The Coefficient of Determination(R2)

How well the model's outputs match the desired value is indicated by the coefficient of determination. The Coefficient of Correlation(R) according to the results presented in Table 4. The first model's MAPE was determined to be 21.52% and the average accuracy percentage produced by the MLR model to be 78.48%.. Consequently, adequate agreement can be stated between the MLR model and the actual measurements.

Table (4): Validity of TI model.

Description	Statistical parameters
MAPE	21.520%
AA%	78.480%
R	95.60%
R^2	91.30%

4.2 Cost Index Model

To determine the statistical link between a response (in this case, the cost index) and the explanatory variables (in this case, position, units of complexity, etc.), multiple linear regressions were used. To get closer to the ideal correlation for the TI model, run-on programs were created till eight runs were completed. Based on the findings of model (1), the variable COW is deemed insignificant and will be excluded from the analysis. This decision is based on the observation that its minimum value, denoted as (t), falls below the critical t value obtained from the t-table distribution. The COS, COC, DOC, UOC, TOC, Y, and EPM models were eliminated in subsequent model iterations as their minimum values of (t) were found to be less than the t value derived from the t-table distribution for each iteration. Furthermore, the results have been produced, and the statistical measures of analysis of variance (ANOVA) pertaining to the regression model are displayed in Table 5.

Table (5) ANOVA statistical nalysis for cost index model regression.

Model		Sum of Squares	df	Mean Square	F
1	Regression	-1.357	13	-2.523	0.585
	Residual	-2.301	6	-3.000	
	Total	-3.658	19		
2	Regression	-1.357	12	-2.398	0.686
	Residual	-2.301	7	-3.000	
	Total	-3.658	19		
3	Regression	-1.367	11	-2.398	0.733
	Residual	-2.222	8	-3.000	
	Total	-3.589	19		
4	Regression	-1.377	10	-2.398	0.730
	Residual	-2.155	9	-3.000	
	Total	-3.532	19		
5	Regression	-1.398	9	-2.398	0.705
	Residual	-2.046	10	-3.000	
	Total	-3.444	19		
6	Regression	-1.409	8	-2.301	0.726
	Residual	-2	11	-3.000	
	Total	-3.409	19		
	Regression	-1.42	7	-2.301	0.785

7	Residual	-1.959	12	-3.000	
	Total	-3.379	19		
8	Regression	-1.42	6	-2.222	0.853
	Residual	-1.959	13	-3.000	
	Total	-3.379	19		

First MLR model: - Cost index (CI) (Model No. 1)

CI = 0.191+0.014 P+0.047 S+0.025 C–0.008 NCD– 0.024 OWL (4)

4.2.1 Examining the MLR Model's Validity (CI Model.1)

As evidenced by the data presented in Table 6. It was determined that the MAPE and Average Accuracy Percentage for the initial model were 33.3% and 66.7%, correspondingly. The MLR model demonstrates a satisfactory level of concurrence with the empirical measurements.

Table (6): The CI model's validity

Description	Statistical parameters
MAPE	33.3%
AA%	66.6%
R	85.4%
R^2	72.9%

5 The Progress of ANN Development

As stated previously, the objective of this study is to develop and evaluate a predictive model for the time and cost index, which is utilised to evaluate the performance of residential construction complexes. In the construction of both models, an artificial neural network (ANN) was likewise utilised. Incorporating a backpropagation algorithm into an additional multilayer feed-forward architecture facilitates data training, testing, and validation. The statistical data division method is employed to determine the optimal number of groups with respect to statistical consistency. Additionally, the influence of data division on the performance of the model is assessed. Version 4 of the commercial PC software programme Neuframe is utilised in this investigation.

6 Neuframe 4 Program

The Neuframe 4 program is a comprehensive collection of intelligent technology solutions that use neural network logic to function as flexible plug-and-play right out of the box, giving the user complete control over all network variables. The black box tester does not have access to the source code and is unaware of the system architecture. A black box test involves the tester interacting with the user interface of the system, giving inputs and assessing outputs without knowing where or how the inputs are processed. Neuframe uses a moderate amount of system resources and doesn't take up much disk space [14]. The Neuframe 4 program's general ANN structure is depicted in Figure 1 to help users understand how explanatory factors are related to response variables as outputs. [15].

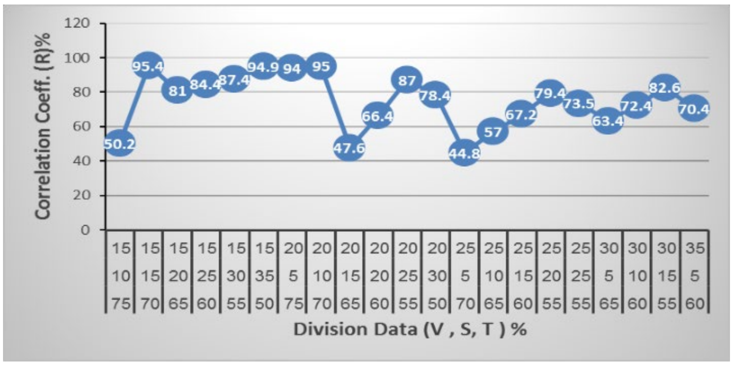
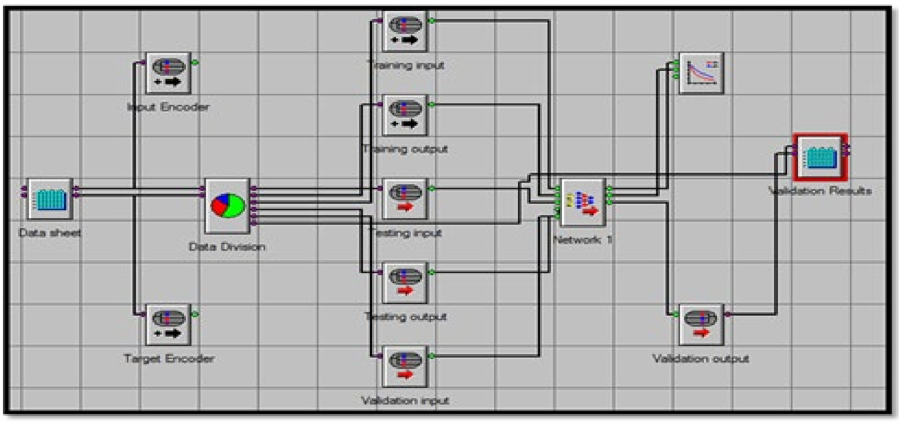


Figure (1): Graphing component of NEUFRAM 4 program



7 Time index Model (TI model.2)

The identical inputs were utilised in the final model of the time index, which demonstrates increased significance in the MLR model. These are the components:

The sector, S

TOC = Complex Type   Units of complexes, or UOC

C = Contract Type

EPM = Project manager experience

NCD = Number of Design Changes

Conditions of Security, or COS

The data from completed residential property investment projects that are used as program inputs is shown in Table (7).

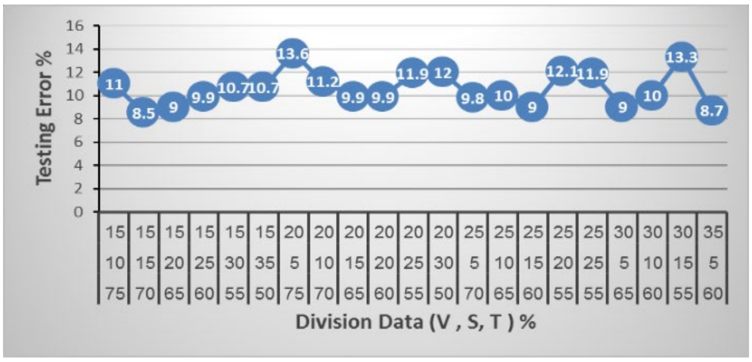
Table (7): The ANN's variable input in the TI model

NO. of project	S	TOC	UOC	C	EPM	NCD	COS	Actual TI
P1	1	2	600	2	22	21	3	-0.4112
P2	2	2	592	1	25	20	3	-0.0067
P3	2	2	504	3	10	19	2	0.2965
P4	1	2	504	3	33	22	3	0.1531
P5	1	1	49	1	20	20	2	0.3958
P6	1	2	228	2	32	22	1	0.4903

P7	2	2	504	1	10	18	1	0.0550
P8	1	2	558	3	25	26	1	-0.0568
P9	1	2	504	3	33	21	1	-0.1983
P10	2	2	504	3	33	20	1	-0.4504
P11	2	2	528	1	15	23	2	0.2634
P12	2	2	586	1	23	19	2	-0.2400
P13	2	2	504	3	15	19	2	0.2302
P14	1	1	167	1	15	19	1	0.0955
P15	2	2	504	3	30	22	2	0.2587
P16	2	2	504	3	25	25	2	0.2787
P17	2	2	480	1	33	27	2	0.1290
P18	2	2	504	3	30	21	2	0.1879
P19	2	2	512	3	20	28	1	0.2086
P20	1	2	432	2	25	24	1	0.1303
Max	2	2	600	3	33	28	3	0.4903
Min	1	1	49	1	10	18	1	-0.4112
SD	-0.2988	-0.5119	2.1599	-0.0300	0.8876	0.4588	-0.1449	-0.1605

7.1 Pre-Processing and Data Division

Pre-processing data is essential to employing neural networks properly. During the training phase, it selects the data that is made accessible to develop the model. The next stage is to divide the available data into three subsets: training, testing, and validation sets in order to develop ANN models. For generalisation, which seeks to provide better results for examples that have not yet been observed, the cross-validation set was employed. The training set, which is used to determine the weights, is the learning set. Nonetheless, the network's performance and generalisation ability are assessed using the test set. As illustrated in Figure (2), the optimal data partitioning for the (TI) model is as follows: 70% for the training set, 15% for the validation (querying) set, and 15% for the testing set. according to the testing set's maximum correlation coefficient (R) of 95.4% and lowest error of 8.5%.



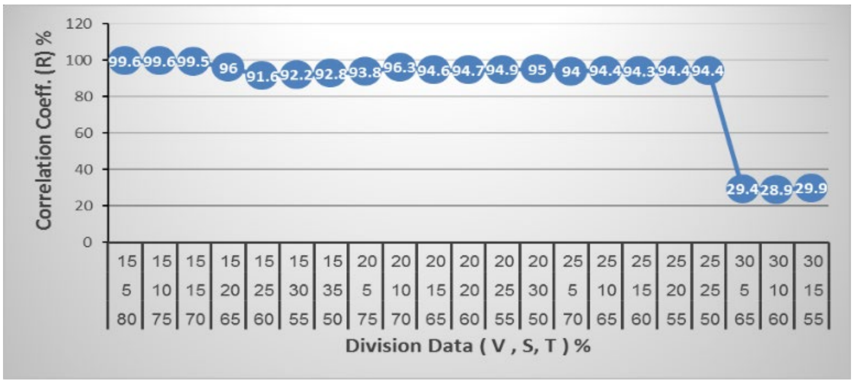


Figure (2): Performance of (TI) model related to division alternatives

Table (8) displays the findings from analysing the impact of using the striped, blocked, and random division choices. It is clear that the division strategy had no appreciable effect on the ANN model's performance. When the blocked division was used, performance improved.

Table (8): Impact of division technique on ANN performance (TI) model

ata division %			Choices of division	Testing error%	Correlation coeff. (R)%
Training	Testing	validation			
70	15	15	striped	13.4	81
70	15	15	blocked	8.5	95.4
70	15	15	random	14.9	93.3

7.2 Scaling of Data

In order to make sure that each variable got the same amount of attention during training, the input and output variables were pre-processed via scaling after the data set was divided into three subsets. Both the hidden and output layers use the sigmoid transfer function (0.0 to 1.0) and the tanh transfer function (-1.0 to 1.0), which need to be scaled to match their respective bounds. A simple linear mapping from the extremes of the variables to the real extremes of the neural network is the most common method for applying scaling (55). In this approach, we find the scaled value Xn by taking the minimum value Xmin and the maximum value Xmax for each variable X. Sections. [16, 17]

involve:X_n=(X-Xmin)/(Xmax-Xmin) (5)

7.3 ANN Model Final Equations

Neuframe determined the minimum number of link weights required for the optimal artificial neural networks (ANNs) model, specifically the TI model, based on the structure illustrated in Figure (3). This finding suggests that the network can be simplified to a very simple formula.

Table 9: lists the connection weights and the threshold levels (bias).

Hidden layer nodes	Wji (weight from node i in the input layer to node j in the hidden layer)							Hidden layer threshold θ
	I=1	I=2	I=3	I=4	I=5	I=6	I=7	
j=8	0.4261	0.1676	-0.3858	0.0404	-0.0383	-0.3131	-0.5123	0.4151
Output layer nodes								Output layer threshold θj
	i=8							
j=9	0.4229							-0.6620

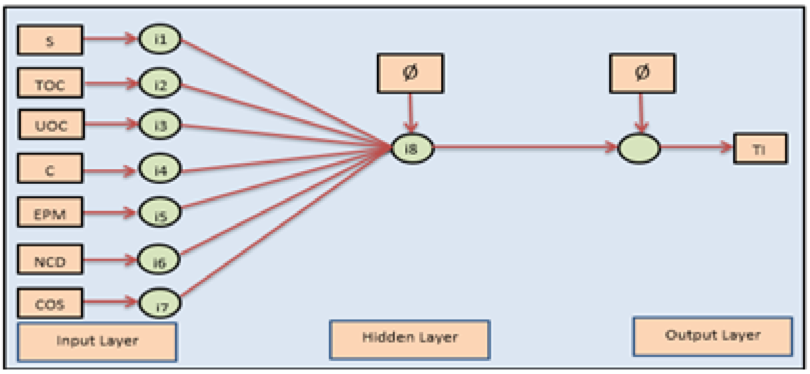


Figure (3): The Optimal Network Neuronal (TI) Model Structure

Table 8 displays the threshold levels when using the connection weight. In addition, Equation -5 must be used to scale all input variables (such as S, TOC, UOC, etc.) between 0.0 and 1.0. The following expression can be used to forecast the time index of an equation

$$\mathrm{TI} = \frac{2.748}{1+e^{-(0.422952*\tanh X-0.66205)}} + 0.354486 \tag{6}$$

Where:

$$X = 0.4216S + 0.1676TOC - 0.0007UOC + 0.0202C - 0.0016EPM - 0.0391NCD - 0.2561COS + 0.812759$$

7.4 Validation of the ANN Model (TI model.2)

According to the findings presented in Table 9, the analysis of the second ANN model reveals that its MAPE is 9.69% and its Average Accuracy Percentage is 90.31%. Thus, it is possible to conclude that the ANN model and the actual measurements are in satisfactory accord.

Table 10: Validity of ANN model

Description	Statistical parameters
MAPE	09.69%
AA%	90.31%
R	095.4%
R2	91.7%

8 Cost index Model (CI model.2)

The same inputs that demonstrate greater significance in the MLR model were kept in the final cost index model. The inputs are as follows: Position is represented by P, Sector by S, Contract Type by C, Number of Design Changes (NCD) by NCD, and Work Location by OWL. The completed residential property investment project data that is used as program input is shown in Table 10.

Table (10): ANN's CI model's input variables

NO. of project	P	S	C	NCD	OWL	[Actual CI]
P1	01	01	01	19	01	-0.003597
P2	02	01	02	24	01	0.001233
P3	03	01	02	21	02	0.01292
P4	01	01	01	20	01	0.002094
P5	05	02	03	19	03	0.030683
P6	06	01	03	26	02	0.044986
P7	14	02	01	18	02	0.171414
P8	07	02	3	20	01	0.045429
P9	07	02	001	23	01	-0.005312
P10	08	02	01	19	02	0.036981
P11	05	02	03	19	02	0.074014
P12	09	02	01	20	02	0.03396
P13	03	01	03	21	02	0.088947
P14	11	02	03	25	01	0.144702
P15	12	02	01	27	03	-0.001972
P16	13	02	03	21	02	0.08051
P17	4	01	03	22	03	0.005868
P18	10	02	03	22	02	0.101858
P19	13	02	03	28	02	0.100595
P20	1	01	02	22	01	0.024954
Max	14	02	03	28	03	0.1714
Min	01	01	01	18	01	0.0053
SD	4.3392	0.5026	0.9333	2.876	0.6958	0.0513

8.1 Pre-Processing and Data Division

The ideal distribution of data for the CI model was determined to be 75% for the training set, 10% for the validation (querying) set, and 15% for the testing set, as seen in Figure (4). This was determined by looking at the testing set's lowest error rate of 7.5% and highest correlation coefficient (R) of 99.6%.The results of an investigation into the impact of using various division options (i.e., striped, blocked, and random) are displayed in Table 10. The performance of the ANN model was unaffected by the division method, as can be seen. The improved performance was obtained by using the blocked division.

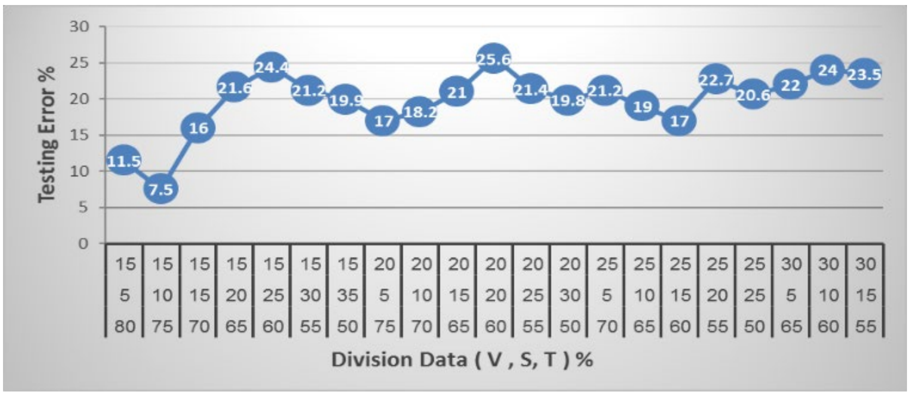


Figure (4): The CI model's performance as it relates to division choices

Table (10): Effect method of division on ANN performance (CI) model

Data division %			Choices of division	Testing error%	Correlation coeff. (R)%
Training	Testing	validation			
`75	10	15	Striped	14	66.60
75	10	15	Blocked	7.5	99.60
75	10	15	Random	10.5	50.10

8.2 Data Scaling

Equation (4) is used to scale the input and output variables after the available data have been split into their respective subsets. This removes the dimension and guarantees that each variable is trained with equal weight.

8.3 ANN Model Final Equations

As a result of the limited quantity of connection weights acquired by Neuframe for the ideal ANNs model (CI model) depicted in Figure 5, the network can be formulated in a comparatively concise manner (as summarised in Table 11).

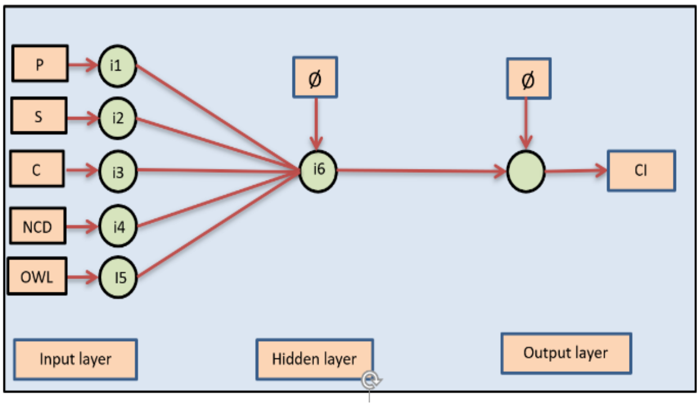


Figure (5): The best (CI) structure for ANNs. Model

Table (11): The ANN optimum CI model's weight and threshold levels

Hidden layer nodes	Wji (weight from node i in the input layer to node j in the hidden layer)					idden layer threshold θ
	I=1	I=2	I=3	I=4	I=5	
j=6	2.0216	-0.0249	0.9269	-0.9471	-0.5398	-0.7974
Output layer nodes						Output layer threshold θj
	i=6					
j=7	5.6106					-3.4106

Table 11 displays the threshold levels based on the connection weight. Equation (5) must also be used to scale all input variables (P, S, C, etc.) between 0.0 and 1.0. The equation cost index prediction is as follows:

$$\mathbf{CI} = \frac{0.17018}{1+e^{-(5.6106*\tanh X-3.4106)}} + (\mathbf{0.0012}) \quad (7)$$

Where:

X = 0.1555P - 0.0249S + 0.4634C - 0.1052NCD - 0.2699OWL + 0.772782

8.4 The ANN Model's Validation (CI model.2)

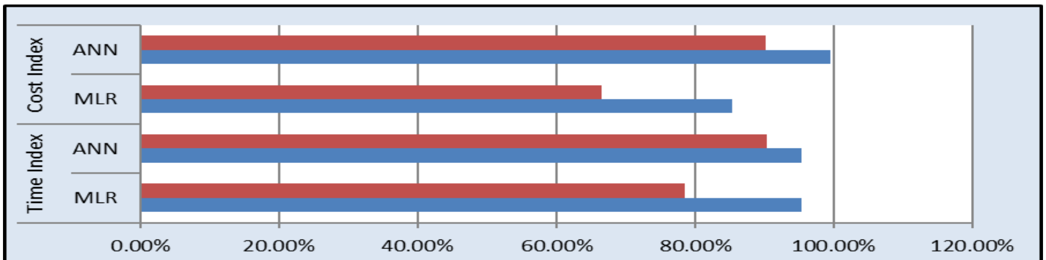
The MAPE and Average Accuracy Percentage that the ANN model for the second model produced were determined using the data in Table (12), and they were 9.69% and 90.31%, respectively. As a result, it is feasible to draw the conclusion that there is suitable agreement between the actual data and the ANN model.The average accuracy percentage (AAP) and maximum accuracy percentage (MAPE) that the second model's ANN model generated were determined using the data displayed in Table (12).

Table (12): ANN model validity

Description	Statistical Parameters
MAPE	9.88
AA%	90.12
R	99.6
R2	99.2

9 Evaluation of ML.R and ANN as Two Methods

A comparison is made between MLR and ANN predictions for the Time Index Model (TI) and the Cost Index Model (CI) in Figure 6. These comparisons imply that the ANN approach outperforms the MLR approach in terms of prediction utility. This could be because neural networks for distributed and parallel computing types could uncover relationships with a considerable reduction in data, whereas the massive quantity of data needed to develop a regression model could account for this. A second explanation is the potential nonlinear impact of the predictors on a response variable.



	Time Index		Cost Index	
	MLR	ANN	MLR	ANN
AA %	78.48%	90.31%	66.60%	90.12%
R %	95.30%	95.40%	85.40%	99.60%

Figure (6) Comparison between ANN and MLR models

10 Conclusions

1. In order to estimate the time and cost performance of building residential property investment projects throughout the execution stage, the current study develops forecasting models using MLR and ANN.
2. The time index model (TI) contained seven variables that were still in place: security conditions, complex units, type sector, number of design modifications, complex type, project management experience, and contract type.

3. There were still five factors in the Cost Index model (CI): type sector, type of contract, number of design modifications, work location organization, and position. The best models were constructed by the ANN technique; these comprised the variables found in MLR models, one hidden layer with one neuron that had a sigmoid transfer function, and a sigmoid for the output neuron's Cost Index (CI) and Time Index (TI).
4. Results from using the same training data for the MLR and ANN models show that they are quite good at predicting whether or not residential complex development projects in Iraq would be successful.
5. When comparing the equations produced by the MLR method and the ANN method, it becomes clear that the ANN method is superior for predictive purposes. Both methods can be used with a high degree of certainty when modelling the results of residential real estate investment projects.
6. One possible explanation is that, for distributed and parallel computing types, neural networks are better at detecting relationships with less data than regression models. The influence of predictors on a response variable is another possible cause. This effect is not necessarily linear.

References

[1] AKOGBE, R. - FENG, X. - ZHOU, J.: Project performance evaluation based on time-cost-design capacity and plant utilization, Quest Journals Journal of Architecture and Civil Engineering, vol. 2, no. 2, pp. 1-9, 2004.

[2] ALSUGAIR, A. M. - ALSANABANI, N. M. - AL-GAHTANI, K. S.: Forecasting the final contract cost on the basis of the owner’s cost estimation using an artificial neural network, Buildings, vol. 13, no. 3, p. 786, 2023, doi: 10.3390/buildings13030786.

[3] MACKOVA, D. - KOZLOVSKA, M. - BASKOVA, R. - SPISAKOVA, M. - KRAJNIKOVA, K.: Construction-duration prediction model for residential buildings in Slovak Republic based on computer simulation, International Journal of Applied Engineering Research, vol. 12, no. 13, pp. 3590-3599, 2017.

[4] KEVIN, - CHARLES, : Analysis of building performance evaluation and value management as tools in building facilities management, Civil and Environmental Research, vol. 3, no. 1, 2013.

[5] AL-ZWAINY, F. M. S. - AL-MARSOMI, M. S. K.: Developing a novel approach for evaluation performance of the engineering departments managers using 360 technique, Eurasian Journal of Science & Engineering, vol. 2, no. 2, pp. 56-67, 2017.

[6] AL-ZWAINY, F. M. S. - HADHAL, N. T.: Building a mathematical model for predicting the cost of communication towers projects using multifactor linear regression technique, International Journal of Construction Engineering and Management, vol. 5, no. 1, pp. 25-29, 2016.

[7] JABER, F. K. - JASIM, N. A. - AL-ZWAINY, F. M.: Forecasting techniques in construction industry: Earned value indicators and performance models, Scientific Review Engineering and Environmental Sciences, vol. 29, no. 2, pp. 234-243, 2020.

[8] WANG, Y. R. - GIBSON, G. E.: A study of pre-project planning and project success using ANNs and regression models, Automation in Construction, vol. 19, no. 3, pp. 341-346, 2010.

[9] AL-ZWAINY, F. M. S. - SALIH, S. A. - ALDIKHEELI, M. R.: Prediction of residual strength of sustainable self-consolidating concrete exposed to elevated temperature using artificial intelligence technique, International Journal of Applied Science and Engineering, vol. 18, no. 2, pp. 1-15, 2021.

[10] AL-ZWAINY, F. M. S. - ALA'AMAR, H. - IBRAHEEM, H. F.: Predicting productivity in construction industry utilizing multiple linear regression technique and artificial neural network technique: A review for research and applications, International Journal of Research in Advanced Engineering and Technology, vol. 6, no. 3, pp. 7-13, 2020.

[11] MAYA, R. - HASSAN, B. - HASSAN, A.: Develop an artificial neural network (ANN) model to predict construction projects performance in Syria, Journal of King Saud University-Engineering Sciences, 2021.

[12] MEMON, A. H. - RAHMAN, I. A. - AZIS, A. A. A.: Time and cost performance in construction projects in southern and central regions of Peninsular Malaysia, International Journal of Advances in Applied Sciences, vol. 1, no. 1, pp. 45-52, 2012.

[13] HAMMOODY, O. - AL-SOMAYDAII, J. - AL-ZWAINY, F. - HAYDER, G.: Forecasting and determining of cost performance index of tunnels projects using artificial neural networks, International Journal for Computational Civil and Structural Engineering, vol. 18, no. 1, pp. 51-60, 2022.

[14] JASIM, N. A. - MARUF, S. M. - ALJUMAILY, H. S. M. - AL-ZWAINY, F.: Predicting index to complete schedule performance indicator in highway projects using artificial neural network model, Archives of Civil Engineering, vol. 66, no. 3, pp. 541-554, 2020.

[15] AIDAN, I. A. - AL-JEZNawi, D. - AL-ZWAINY, F. M. S.: Predicting earned value indexes in residential complexes' construction projects using artificial neural network model, International Journal of Intelligent Engineering and Systems, vol. 13, no. 4, pp. 248-259, 2020.

[16] ZAMIM, S. K. - FARAJ, N. S. - AIDAN, I. A. - AL-ZWAINY, F. M. S. - ABDULQADER, M. A. - MOHAMMED, I. A.: Prediction of dust storms in construction projects using intelligent artificial neural network technology, Periodicals of Engineering and Natural Sciences, vol. 7, no. 4, pp. 1659-1666, 2019.

[17] JABER, F. K. - AL-ZWAINY, F. M. S. - HACHEM, S. W.: Optimizing of predictive performance for construction projects utilizing support vector machine technique, Cogent Engineering, vol. 6, no. 1, pp. 1685860, 2019.

Appendix A

Table (A): The Explanatory and Response Variables

Project NO.	Explanatory													Response	
	P	Y	S	TOC	UOC	C	EPM	COC	DOC	NCD	COS	COW	COW	TI	CI
P_1	1	2003	1	2	228	2	11	96900	270	22	1	1	1	0.490	0.024954
P_2	2	2003	1	2	432	2	19	30000	900	24	1	1	1	0.130	0.001233
P_3	3	2003	1	2	600	2	20	50000	1080	21	2	3	2	-0.411	0.01292
P_4	4	2004	1	2	504	3	22	42775	1334	22	3	3	2	0.153	0.005868
P_5	1	2007	1	1	167	1	11	19740	610	19	1	1	1	0.095	-0.0036

P_{6}	1	2007	1	1	49	1	11	51587	367	20	1	2	2	0.395	0.002094
P_{7}	5	2007	2	2	504	3	23	39729	670	19	3	2	2	0.296	0.030683
P_{8}	6	2007	1	2	558	3	16	46815	914	26	2	1	3	-0.056	0.044986
P_{9}	3	2007	1	2	504	3	13	47500	914	21	2	1	1	-0.198	0.088947
P_{10}	7	2007	2	2	504	3	25	42000	914	20	1	1	2	-0.450	0.045429
P_{11}	7	2008	2	2	528	1	23	58927	609	23	1	2	2	0.263	-0.00531
P_{12}	8	2008	2	2	586	1	21	59896	914	19	2	2	3	-0.240	0.036981
P_{13}	5	2004	2	2	504	3	20	39114	1219	19	2	2	1	0.230	0.074014
P_{14}	9	2008	2	2	592	1	20	70142	913	20	2	3	1	-0.006	0.03396
P_{15}	10	2005	2	2	504	3	25	44660	1082	22	2	2	2	0.258	0.101858
P_{16}	11	2007	2	2	504	3	17	46606	1159	25	1	2	3	0.278	0.144702
P_{17}	12	2008	2	2	480	1	11	64906	911	27	3	2	3	0.129	-0.00197
P_{18}	13	2005	2	2	504	3	15	40815	1066	21	2	2	2	0.187	0.08051
P_{19}	13	2007	2	2	512	3	25	52448	913	28	2	1	1	0.208	0.100595
P_{20}	14	2011	2	2	504	1	22	39600	570	18	2	1	2	0.055	0.171414