

Guess the time of implementation of residential construction projects using neural networks ANN

Elaf Dheyaa Abdulridha AL-Zubaidi¹, Ali Hashim Yas², Hayder Fadhil Abbas³

¹ University of ALQadisiyah

² College of Engineering, University of Kerbala

³ Construction Management Engineering, Al-Farabi University College

Article Info

Received Feb 24, 2019

Keyword:

Artificial Neural Network
Implementation
Residential Construction
Time

ABSTRACT

The construction duration of residential projects, especially in building processes, significantly impact the business of a construction company. The balance between the planned cost, direct cost, and overheads directly depend on the precision of the implementation phase of the project. The application of the artificial neural network (ANN) to predict the duration of implementation of a residential construction project from the pre-design stage to completion is comprehensively discussed in this research. The study applies the back-propagation (BP) network made of nodes for error evaluation of the training states. Further, the proposed system illustrated that the artificial neural network (ANN) satisfy the three crucial criteria (cost, quality, and time) used for the evaluation of projects. The ANN provided accurate data for the training and estimation of, the duration of a residential construction project with adequate resources of implementation. The performance of the results for the ANN at 105 iteration shows that the prediction was 99.841% accurate for the overall system. The best fit occurred at the 99th epoch with an MSE of 0.10286.

Corresponding Author:

Elaf al-zubaidi
University of ALQadisiyah,
elaf.alzubaidi@qu.edu.iq

1. Introduction

The success of contract construction projects is a belt on the realization of the project within the plant time for the agreed cost. The primary prerequisite for this success is the extensive consideration of the project at all the implementation stages [1]. In every contract, special attention is accorded to the design stage to agree on the cost and length of the project. For every contractor, the duration and the cost are among the vital parameters that determine the success of a residential project. The construction company carries out a pre-contract assessment of the cost and duration of the project to provide investors with accurate and reliable information [2]. The cost and duration estimation evolved from conceptual phases to the detailed, definitive estimates face that is dependent on the quantity of information collected by the contractor. Traditionally information regarding the potential work and duration is usually available in minimal amounts, and this is based on the estimates made by the contractor before making an offer for a project. Prior to making any offer for the intention of undertaking the project, the contractors undertake a conceptual estimation of the cost and time of the project [3]. These activities are undertaken just before placing an offer but immediately after receiving a query from an investor. Similarly, a comprehensive analysis of the possible project terms, i.e., the cost and time required are presented to the investor based on the amount of work to be completed.

Studies have shown a profound discrepancy between the prospected duration and cost of a project and the real cost and duration, i.e., defines the actual project completion parameters [4]. The figures in these researches

have demonstrated that it is vital for construction companies to create internal databases to be used as benchmarks for future projects. It thus becomes vital to enter the construction data into the databases from the inception of the query to the final stage of project documentation. Additionally, the projects tender price, as well as the real cost and duration, should be stored [5]. The conceptual estimates function to evaluate the validity of the potential project, i.e., decide on the viability of the project., Considering that the formation of the preliminary parameters is created by the contractor using sparse information, making the use of modern prediction technique very significant such as the utilization of artificial neural networks [6]. The use of artificial neural networks Computing the conceptual estimates to a substantial extent would decrease the time and the cost for data processing. The application of databases in the assessment of the viability of a prospective project is significant in the provision of historical data on the real cost and completion time of a successful project [7]. The database would make the estimations more accurate for both the cost and duration of the potential project, thus improving the decisions made by the contractor in taking up the project.

This research creates, with the help of an artificial neural network, a design for the conceptual estimation of the length of a residential construction project. This proposed integrated model predicts the time to complete the project. The preliminary estimates of the parameters of the project were based on the bill of quantities, which is the most vital information present for the potential project as defined in the completed project database and tender documentation.

2. Background and Related work

The artificial neural network development was prompted by an article by McCulloch and Pitts, which is a recent development [8]. Thus, ANN cannot be fully defined based on a single scenario as it has a broader application. Briefly, the artificial neural network can be described as a computational technique used in the acquisition, representation and computational mapping from one multivariate space of data to another provided a set of data representing that mapping [9-13]. This technology imitates the operation of the biological neural networks applying appropriate mathematical ideals such as function, structure and processing information; thus, possess the capability to memorize, learn and recognize generalized rules.

The fundamental processing unit in ANN is the neuron and can be represented using the mathematical model of McCulloch-Pitts demonstrated in figure 1 below.

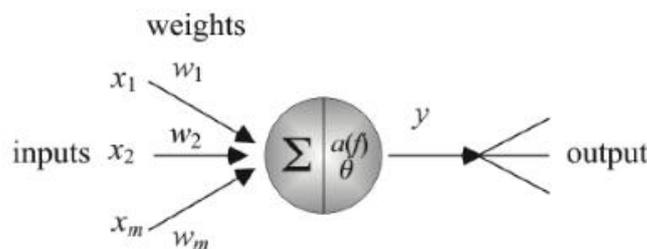


Figure 1: The McCulloch-Pitts design of the neuron

The above figure 1 illustrates the neuron-mathematical where the input data (x_1, x_2, \dots, x_m) with the corresponding weight coefficients of (w_1, w_2, \dots, w_m) resemble the dendrites [1]. Similarly, the body of the neuron represents the processing unit of the signals based on the neuron activated [9]. Once there is an activation a signal is then transmitted via the output (axon) to the succeeding in the neuron in the connection. The fundamental tasks associated with the processing components are receiving input from adjoined neurons providing activation, compute the data and transmitting the output to the succeeding neuron [1]. The neurons in an ANN can be categorized into input neurons, output neurons and hidden neurons.

2.1. Related work

Jain et al. conducted a comprehensive analysis on the application of NN in the construction engineering and management. The research demonstrated that the use of artificial intelligence in the field of construction

engineering is far from ideal. Further research, as illustrated the neural networks provide a robust system that can be used as a complementary computational system to the conventional expert systems used for forecasting costs and project durations [7]. Jain et al. introduce the aspect of neural networks as a prospective design and conceptual tool used to improve automation ideal in the engineering and construction industry. The inclusion of the underlying neural network in their research shows that there are possibilities of incorporating neural networks in expert systems for enhancing efficiency in the intelligent systems [7]. This research dwells on the inaugural use of neural networks in the construction industry, but for this research, we create a comprehensive ANN model using three vital AI aspects to compute the time for completion of a residential project. This research has shown that ANN is a crucial component of expert systems used in the construction industry.

Le-Hoai et al., acknowledge that the time performance is a crucial aspect used in benchmarking construction project. In this research, the project time can be affected by various factors that can arise in the course of the project execution phase [6]. Le-Hoai et al., believe that the identification of problems that are substantial to the execution of construction schedules and considering them in any model was useful for project management and implementation. The research was based on 70 projects that applied the multiple regression techniques to integrate vital parameters into the design to forecast the project time [6]. In this research, the regression model is compared with the ANN for cross-validation, which demonstrates that the regression model performs exemplary [6]. Despite the milestone achieved in this study, the regression model does not consider the mean square error associated with forecasting, which is solved in our proposed model.

Peško et al. proposed an artificial intelligence system for the approximation of duration and cost of a construction project [4]. The balance between the planned costs, overheads and direct cost considerably affect the prospected profits there; it is essential to prepare a concise offer for the available resources and the required time [4]. The study evaluates and compares the support vector machine (SVM) and the artificial neural network (ANN) in predicting the cost and time of a construction project. Best on the mean square error (MSE) the SVM has higher precision than ANN. Our research considers the performance of ANN in the training state, regression and error histogram [4]. These aspects demonstrated that the ANN is a robust expert system for forecasting of project time.

3. Implementation of the Proposed System

The proposed artificial neural network system was created in three phases. The three phases are vital in the formation of an accurate ANN for prediction of a project time [10]. Further, the quality and applicability of the data in the database for training were considered the most fundamental elements to determine the precision of the forecast.

1. Modeling phase: In this section, the ANN was modelled by defining the network architecture where a series of parameters such as the number of layers, input parameters and the quantity of neurons are defined. Additionally, the quantity of the output data, the activation functions of the neuron and the type of training function based on the orientation of the network was described.
2. Training phase: the training of the ANN was either based on the training of this network structure or the training weight. In this section, the training algorithm was selected as either supervised, unsupervised or reinforcement training.
3. Testing phase: entails the comparison of the output of the proposed ANN model. This exercise is based on the dataset for testing with the prospected values. The comparison of the figures and quantities in ANN produced a satisfactory output that can cause adjustments in the network structure to enhance the accuracy

3.1. Dataset

The primary precondition for the use of the ANN is to create a comprehensive database for the data sets used for performance evaluation, training and testing [11]. The dataset for this research was collected from the kaggle database used for a similar residential project in 2016. This data was based on a similar contractor who utilized similar situation such as machinery, human and other resources. The data did not consider subcontractors for the project and was considered for 12 months for the completed project for similar residential construction. The data was prepared for processing using ANN, which was an analysis of the vital

elements of the project for the time prediction framework [12]. The commencing point of the conceptual estimates was based on the bill of quantities used for residential projects therefore in this case the following are the fundamental components were considered;

Table 1: The pre-analysis data for the duration prediction

Component	Training Set %	Testing Set %
Structural System	51, 15, 2	49, 5, 14
Building Function	35, 26, 4	3, 4, 24, 37
Foundational System	12, 35,	23, 45
Height Units	28, 40	29, 39
Accessibility to Site	47, 21	46, 21
Area/Floor Ratio	373, 20, 100	743/6907
Construction time	12/60	12/60

The information was prepared for analysis via the normalization of the data. The normalization of the data increases the validity of the trained ANN. The normalization of the entire database, i.e. the input and output for the training and testing set for the ANN. The Z-Score is used for the normalization of the performance via the transformation in the distribution where the mean (μ) 0, and the standard deviation (σ) 1 which was computed in the following expression [13].

$$S_{ij} = \frac{X_{ij} - \mu_i}{\sigma_i}$$

Where

i is the input or the output data

S_{ij} is the normalized data value

X_{ij} is the actual data in the dataset

μ_i is the mean distribution for the training of the data

σ_i is the standard deviation of the distribution

3.2. Modelling of the ANN Framework

In this phase, the quantity of the input and the output data in be defined [14-16]. The input parameters define the special parameters of the network in terms of dimension and the quantity of output parameters that determines the quantity of the solution surfaces created in the network [17-19]. The quantity of the input was considered as defined in the database for the analysis of the individual data. The output was the end-results of the processing of the neurons in the network, as shown in figure 2.

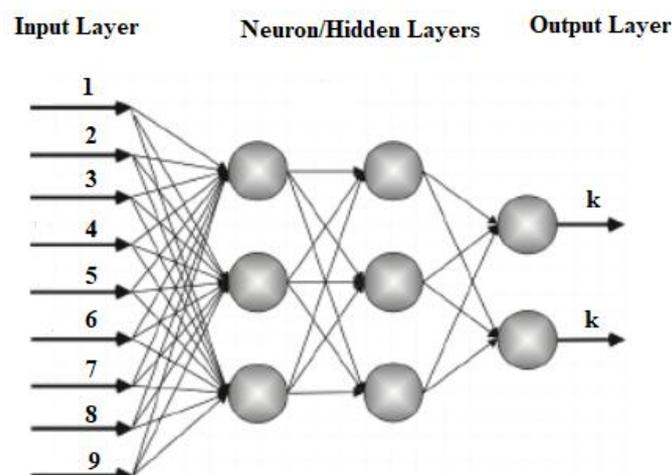


Figure 2: The model of the ANN framework, i.e., the 3L-3N1-2

The quantity of the neurons in the input layers affects the output of the neural network. The neurons are located in the hidden layer and are responsible for the processing of the data [9]. Moreover, the quantity of neurons affected the degree of freedom in the neural network, that significantly affect the accuracy of the modeled network. There were no fixed rules in the modeling other than the definition of the number of layers

and neurons. The provisions in the model will be defined as iteratively [15]. Figure 2 illustrates the ANN model with a transfer function either as the hyperbolic tangent sigmoid, which is was utilized in this model.

3.2.1. Training of the ANN

A supervision algorithm was utilized for ANN training while the backpropagation was useful error monitoring considering the broad application of the model. The modelling of the neural network was done in MATLAB R2019a. Multiple training functions were used for the ANN model. The ANN model 3L-3N1-2 was trained using the functions in table 2 below. During training the criterion limit was set at 1000 epochs, 68.8 for the performance validation, the gradient was 240, and there were six checks for the validations. The epochs define the number the data will pass via the training in the ANN.

Table 2: The functions in the training data

Training Functions	
Training A	Gradient change adaptation in the backpropagation
Training B	Levenberg-Marquardt back-propagation
Training C	Bayesian Regularization
Training C	BFGS quasi-Newton back-propagation
Training D	P-B conjugate gradient back-propagation

3.2.2. Testing of the ANN model

The testing phase was the evaluation of the performance of the proposed ANN design for guessing the time of the residential project. The primary technique utilized was based on statistical evaluation for them mean, coefficient variation and standard deviation. The mean square error was used to create the error histogram applying the equation below.

$$MSE = \frac{1}{n} \sum_{i=1}^n |I_{predicted} - I_{actual}|$$

Where, n is the quantity of the dataset for testing ANN. The MSE was conducted on the output of the ANN-based on the dataset. The error indicated the deviation of the real estimates from the real data. The error computation was calculated after training and testing of the model. The error testing led to the analysis of the stability of the neural network [10]. The MSE was used as a benchmark for the evaluation of the stability of the ANN framework for predicting the time of the residential projects. The optimization of the input data was based on a similar technique, as presented above [11]. The output of the data was maintained constant and did not vary using the ANN as the parameters for the prediction of the cost were based on a complete project.

4. Results

The simulation of the time prediction was completed using MATLAB R2019a (ntraintool) used for training ANN data, characterizing the performance and testing the error correlation of the approach. The ntraintool is demonstrated in figure 3 below, together with the related algorithm [5]. The MATLAB ntraintool took an input that was passed through a set of hidden neurons for processing then the output computed in the processing neuron before being displayed as the actual output.

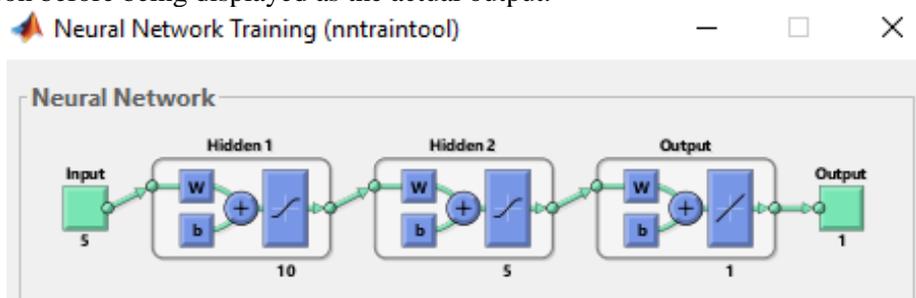


Figure 3: The MATLAB ntraintool illustrating the input node the two hidden neurons, the processing node and the output node. This tool was a clear representation of the 3L-3N1-2 model.

The algorithm used in this ntraintool to compute the collected data and generate the desired output utilized the Levenberg-Marquardt back-propagation for the training of data as shown below.

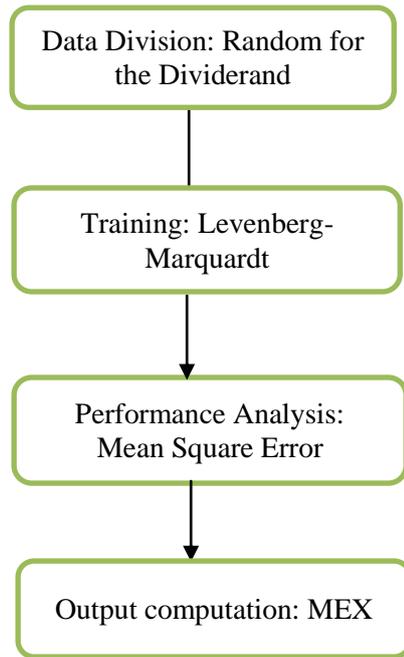


Table 3: The normalized input data used for the simulation of the prediction of the duration for the residential project

Component	No.	Measurements	Real Data		Normalized Data	
			Max	Min	Max	Min
Structural System	1	m ³	8364.70	00.00	4.73	-0.83
Building Function	2	m	12470.0	28.00	4.97	-0.68
Foundational System	3	t	18167.50	94.50	6.66	-0.87
Height Units	4	m	56.00	28.00	2.96	-0.74
Accessibility to Site	5	km	12.00	5.00	1.70	-1.33
Area/Floor Ratio	6	m ²	24560.00	00.00	2.27	-0.74
Drainage Works	7	level	6	0	2.99	-0.56
Decoration	8	level	4	0	1.25	-0.36
Output Data						
Construction time	1	months	12	8	2.39	-1.86

The dataset was divided into two where the real part was utilized for training of the ANN, and the second part was used for testing. Both the input and output data used for testing considered the minimum and maximum range for real and normalized data [3]. The output for the normalization illustrated that the construction time with similar parameters as the real data would be completed within 8-12 months, which is within the original range.

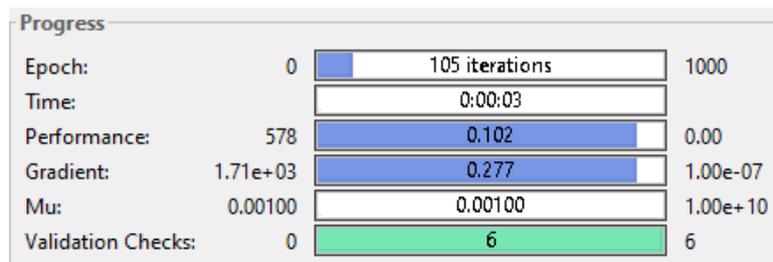


Figure 4: The parameters for artificial network training, validation and testing for the prediction of the residential project duration. 105 epochs were used, i.e., the data was passed 105 times via the training data. The validation of the data was completed after six checks.

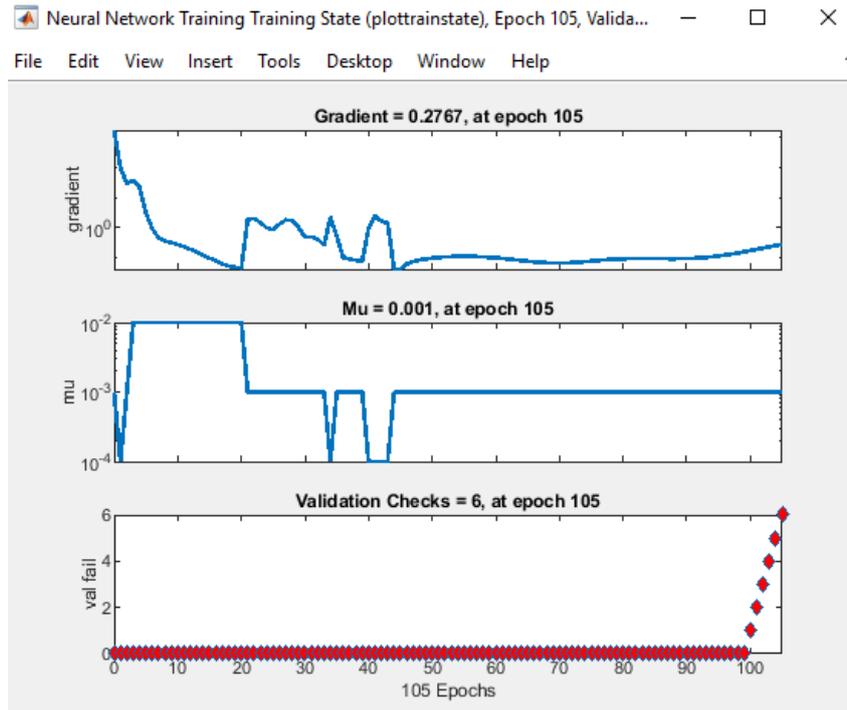


Figure 5: The plot for the training state parameters for the 105 epochs.

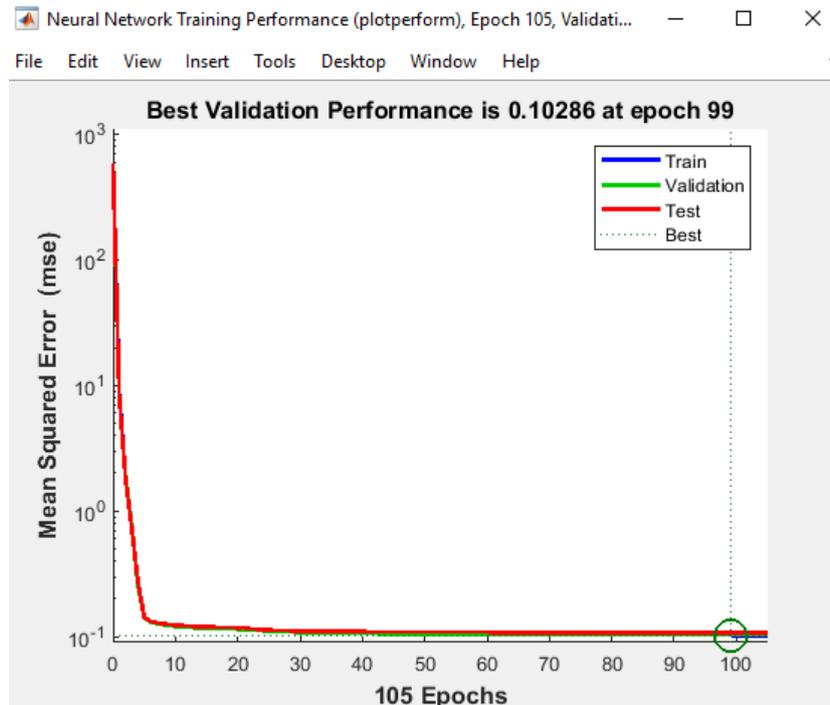


Figure 6: The results for the validation of 105 epochs with the best fit occurring at epoch 99, which is 0.10286.

Figure 6 above demonstrate the performance of the real construction data for the prediction of the construction time using the MSE and regression [14]. The training states were developed used the Levenberg-Marquardt (LM) using the back-propagation algorithm comparing the output and the target of the analysis using the ANN. The performance illustrates the best fit value for the selected epoch value [7]. The Mean Square Error (MSE) shows the deviation of the predicted value from the real data. The best fit value according to the results

in figure 6 shows that the most desirable value or the results of the iterations was 0.10288 for the 99th epoch. The LM was applied in the hidden layer where the backpropagation was used for fitting the results.

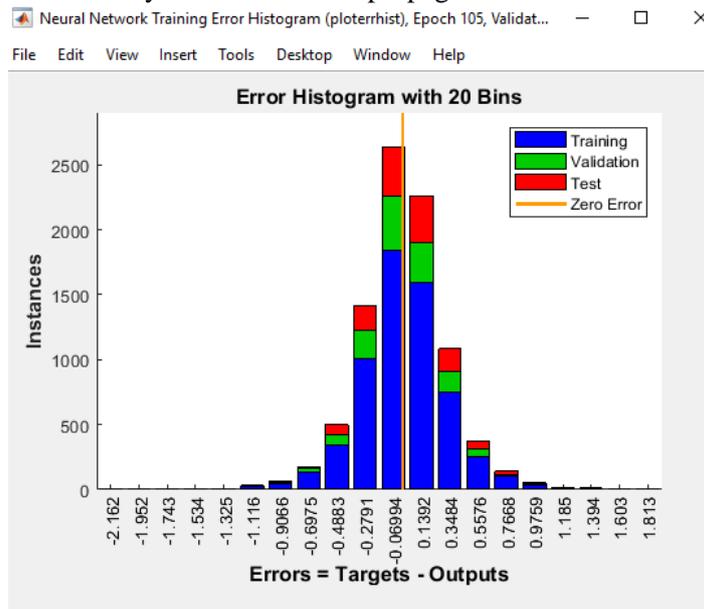


Figure 7: The error histogram for the training of the data using the Levenberg-Marquardt (LM) using the back-propagation.

The histogram in figure 7 demonstrates the error for the 20 bins for the training data, the validation and the testing of the performance of the normalized data [2]. The data is used to display the zero error for the validation of the performance. The best fit was recorded at epoch 99 for the hidden layer. The error value was close to the zero error, illustrating that most of the predicted value was close to the real data. The error histogram verified the performance of the training, validation and the testing phase.

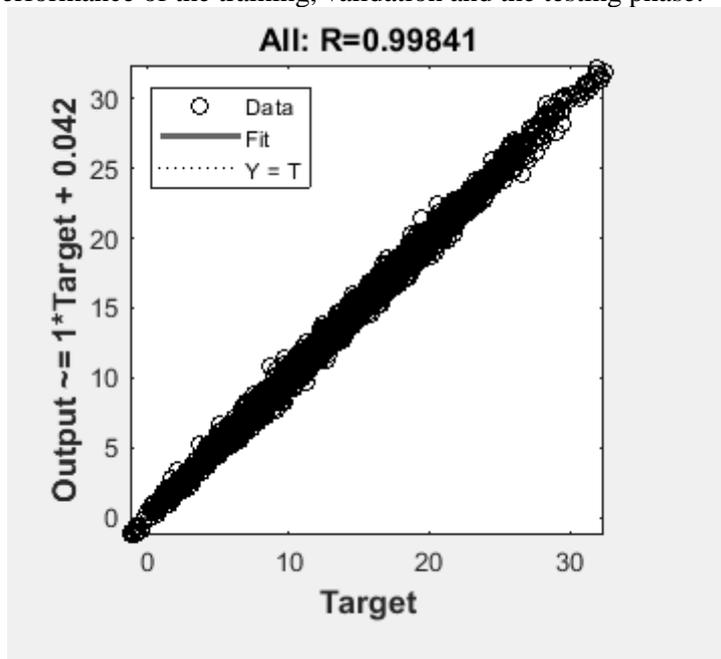


Figure 8: The results for the regression between the output data and the targets for the LM approach for the hidden layer.

The regression analysis (R) show that the results were 99.841% accurate with the most accurate prediction occurring at epoch 99. The accuracy of the results demonstrates that the collected data was significant in predicting the time for the development of the residential project [1]. The regression analysis in this research was used for comparison of the performance of the ANN technique showing that this approach is robust in

creating conceptual estimates for the predesign and development of offer for the construction project contracts.

5. Conclusion and Future Work

The ANNN has been demonstrated as a robust model compared to the conventional arithmetic and statistical model for the conceptual design of complex residential design projects. The proposed mode proved that ANN could handle the complexity in terms of quantitative and qualitative evaluation of design projects. The proposed model was used to predict the time for the conceptual design and estimation of the duration of implementation of a residential project. The dataset from a similar real project was used to evaluate the performance of the proposed system. Based on the results of the research, the MSE can be minimized through the use of multiple databases. Future research is recommended to evaluate the applicability of multivariate datasets and databases in enhancing the accuracy of ANN in the conceptual estimation of the duration of implementation of a project

References

- [1] A. Bakr, A. A. Ragab and N. A. B. Yehia, "Estimates of Site Overheads for Residential Projects In Egypt Using Neural Network," in *Sustainable Construction and Project Management-Sustainable Infrastructure and Transportation for Future Cities*, Aswan, Egypt, 2018.
- [2] A. A. A. Alshahethi and K. L. Radhika, "Estimating the Final Cost of Construction Project Using Neural Networks: A Case of Yemen Construction Projects," *International Journal for Research in Applied Science & Engineering Technology*, vol. 6, no. 11, pp. 2141-2151, 2018.
- [3] N. I. El-Sawalhi and O. Shehatto, "A Neural Network Model for Building Construction Projects Cost Estimating," *KICEM Journal of Construction Engineering and Project Management*, vol. 4, no. 4, pp. 9-14, 2014.
- [4] I. Peško et al, "Estimation of Costs and Durations of Construction of Urban Roads Using ANN and SVM," *Complexity*, vol. 2017, pp. 1-13, 2017.
- [5] V. V. Thendral Tharmalingam, "An Efficient Convolutional Neural Network Based Classifier to Predict Tamil Writer," *Period. Eng. Nat. Sci.*, vol. 6, no. 1, pp. 285–295, 2018.
- [6] V. B. Chandanshive and A. R. Kambekar, "Prediction of Early Stage Construction Cost of Building Projects Using Artificial Neural Network," *International Journal of Scientific & Engineering Research*, vol. 5, no. 7, pp. 453-463, 2014.
- [7] L. Le-Hoai, Y. D. Lee and A. T. Nguyen, "Estimating time performance for building construction projects in Vietnam," *KSCE Journal of Civil Engineering*, vol. 17, no. 1, pp. 1-8, 2013.
- [8] M. Jain and K. K. Pathak, "Applications of Artificial Neural Network in Construction Engineering and Management-A Review," *International Journal of Engineering Technology, Management and Applied Sciences*, vol. 2, no. 3, pp. 142-134, 2014.
- [9] M. Mishra, P. K. Singh, and A. Brahmachari, "A Robust Pest Identification System using Morphological Analysis in Neural Networks," *Period. Eng. Nat. Sci.*, vol. 7, no. 1, pp. 483–495, 2019.
- [10] H. Yahia, H. Hosny and M. E. Abdel Razik, "Time Contingency Assessment in Construction Projects in Egypt using Artificial Neural Networks Model," *IJCSI International Journal of Computer Science Issues*, vol. 8, no. 4, pp. 523-531, 2011.
- [11] B. S. Babu, A. Suneetha, G. C. Babu, Y. J. N. Kumar, G. Karuna, and G. W. Optimization, "Medical Disease Prediction using Grey Wolf optimization and Auto Encoder based Recurrent Neural Network," *Period. Eng. Nat. Sci.*, vol. 6, no. 1, pp. 229–240, 2018.
- [12] R. M. Choudhry, S. Rafiq, A. Ahmed and T. Ahmed, "Estimating Probabilistic Cost and Time Contingency for Residential Building Projects," in *Proceedings of the 3rd International Conference on Civil, Structural and Transportation Engineering*, Niagara Falls, Canada, 2018.
- [13] M. Y. Hegab and G. R. Smith, "Delay time analysis in microtunneling projects (Technical notes).," *Journal of Construction Engineering and Management (ASCE)*, vol. 133, no. 2, pp. 191-195, 2007.

- [14] B. Durakovic and H. Basic, "Continuous Quality Improvement in Textile Processing by Statistical Process Control Tools: A Case Study of Medium-Sized Company," *Periodicals of Engineering and Natural Sciences*, vol. 1, no. 1, pp. 39-46, 2013.
- [15] M. A. Azman, Z. Abdul-Samad and S. Ismail, "The accuracy of preliminary cost estimates in Public Works Department (PWD) of Peninsular Malaysia," *International Journal of Project Management*, vol. 31, no. 7, pp. 994-1005, 2013.
- [16] A. A. G. Allah, A. H. Ibrahim and O. A. Hagrass, "Predicting the construction duration of building projects using artificial neural networks," *International Journal of Applied Management Science*, vol. 7, no. 2, pp. 123-141, 2015.
- [17] B. S. Waziri, K. Bala and S. A. Bustani, "artificial Neural Networks in Construction Engineering and Management," *International Journal of Architecture, Engineering and Construction*, vol. 6, no. 1, pp. 50-60, 2017.
- [18] S. Petruseva, V. Zujo and V. Zileska-Pancovska, "Neural Network Prediction Model for Construction Project Duration," *International Journal of Engineering Research & Technology*, vol. 2, no. 11, pp. 1646-1654, 2013.
- [19] M. Juszczuk, A. Leśniak and K. Zima, "ANN Based Approach for Estimation of Construction Costs of Sports Fields," *Complexity*, pp. 1-11, 2018.