

Neural network method for control valve cost estimation on the EPC project bidding

Gilang Almaghribi Sarkara Putra^{1,*}, and *Rendra Agus Triyono*^{1,*}

¹Industrial Plant Department, PT Wijaya Karya (Persero) Tbk, Jakarta, Indonesia

Abstract. Cost estimation on the bidding phase is a crucial stage that determines the success of the Engineering, Procurement and Construction (EPC) project. If the cost offered to the client is too high then it could not compete with the other bidder, but if the cost offered are too low it can reduce profit margins and result in losses for the EPC companies. This paper describe the use of Back Propagation Neural Network method to help determine cost estimation. This method is applied specifically to determine control valve cost estimation on the bidding phase so that the retrieved costs will be accurate. When there is no technical and price quotation from vendors as well as the narrowness of the bidding processing time, this method can be an alternative choice to determine the price based on previous vendor quotation. In the future, this method could be developed and applied for other instrumentation equipment such as transmitter, switch, analyzer, control system and others to achieve total cost estimation of instrumentation equipment in EPC bidding proposal.

1 Introduction

Control valve is an instrumentation component that almost exist throughout the industrial plants mainly in the oil-gas industry and petro-chemical. On the bidding phase, control valve cost estimation count as critical item because it is included in the list of major instrumentation equipment. Control valve contributed a considerable portion of the cost in a bidding proposal. Control valve cost estimation on a less appropriate bidding phase can lead to profit and loss of the project when it was already running.

Biddings which are carried out by Engineering, Procurement and Construction (EPC) companies often have a short time limit, between two to eight weeks for accumulation to the client. Their workflow to process control valve Request for Quotation (RFQ) takes about two to four weeks depending on the workload of each control valve vendor and evaluation of their engineering team. Sometimes, control valve vendor could not meet the deadline set by EPC so that there will no price quotation from vendor. Control valve cost estimation is being inaccurate if there is no price quotation from vendor. Therefore, EPC estimators need a tools to make control valve cost estimation in a short time of bidding phase.

* Corresponding author: almaghribi@wika.co.id; rendra.agus@wika.co.id

Artificial Neural Network (ANN) is an artificial intelligence system that can change its structure to solve the problem based on the external and internal information that flow through the network. ANN can be used as a tool to make the mathematical model of cost estimation, so the accuracy of cost estimation become better. Cost estimation based ANN has been developed in Large Scale Project [1] and Transformer Main Materials [2]. Today, most of developed cost estimation software is based on ANN with various method [3-4]. The expected results of ANN system is to achieve a (one) 1 correlation coefficient value which means that estimated value is not much different with actual value, avoid overfitting cost estimation and give results of control valve cost estimation accurately.

2 Methods

2.1 Network architecture

The ANN method of control valve cost estimation presented in this paper using back propagation. Back propagation learning use gradient steepest descent method to determine the weights of neuron connections. Different learning algorithm will give different output. [5]. ANN method has been applied in cost estimation for construction project [6].

Network design consists of determining the size of network, number of neurons and layers required and type of activation function. Construction design of network architecture estimated cost of control valve with the ANN consists of three layers namely input layer, output layer and hidden layer. Input layers for this network are parameters of data as much as eleven pieces which are shown in Table 1, while the output layer is cost of control valve. Fig. 1 shows ANN architecture control valve cost estimation.

Table 1. Input Parameter Design

Parameter Design	Parameter Definition	Parameter Range
X1	Body Size	1-10 (1" – 20")
X2	Connection Rating	1-7 (150# - 2500#)
X3	Body Material	1-10 (Cast Iron – Duplex SS Forged)
X4	Trim Material	1-10 (Cast Iron – Duplex SS Forged)
X5	Actuator Type	1-9 (Single Acting Spring - Electrohydraulic)
X6	Area Classification	1-10 (Ex n – Ex ia)
X7	Valve Type	1-6 (Gate Valve – Choke Valve)
X8	Leakage Class	1-6 (I – VI)
X9	Fluid Condition	1-10 (Low – High Viscosity, Flow, Pressure, Temperature, Toxic, Erosive & Corrosive)
X10	Characteristic Valve	1-7 (Quick Opening – Hyperbolic)
X11	Average Inflation	1-7 (2010, 5.12% – 2017, 3.76%)

- Number of neuron input, hidden layer and output layer
 Number of neuron input determination is performed based on parameters design which amounted to 11, while the number of hidden layer nodes is determined based on evaluation of trial network with combinations level start from 10 to 30. The best hidden neurons combination which applied can be seen in the absence of overfitting and low value of RMSE (Root Mean Square Error) of the testing and training data.
- Activation function
 Activation function for a Back Propagation network should have several important characteristics, e.g.: continuous, differentiable and monotonous not decreasing [5].

Independent variable [x] in the function [f(x)] is number of multiplication between input values and weights. There is some activation function which is often used for back propagation, among others, binary sigmoid, bipolar sigmoid, and linear functions. Binary sigmoid has a range of values from 0 to 1 and is defined by following formula:

$$f_1(x) = \frac{1}{1+\exp(-x)} \quad (1)$$

Another activation function is bipolar sigmoid with range of values of 0 to 1 and is defined by following formula:

$$f_1(x) = \frac{2}{1+\exp(-x)} - 1 \quad (2)$$

And the last activation function is linear function defined by following formula:

$$f_1 = x \quad (3)$$

Network architecture of ANN is using the activation function of binary sigmoid for hidden neurons and linear activation function for output layer.

- Learning rate [α]
Learning rate determines how efficiently and quickly the convergence of back propagation algorithm. Learning rate is a control parameter that regulates step magnitude at which the weights [W] is iterative regulated [5]. Smaller learning rate value will result slower network learning, whereas if learning rate is too large will result the oscillating system. Range of learning rate values is 0 to 1. Learning rate value usually selected based on testing data. The best learning rate value for ANN is 0.08 that obtained from the best RMSE test and there is no overfitting between range values 0.01 to 0.2.
- Momentum [μ]
One of alternatives on updating weight value of network architecture to get the smallest RMSE value is by using momentum. The momentum formula is as follows:

$$\Delta W_{jk}(t) = \alpha \delta_k z_j + \mu [W_{jk}(t) - W_{jk}(t - 1)] \quad (4)$$

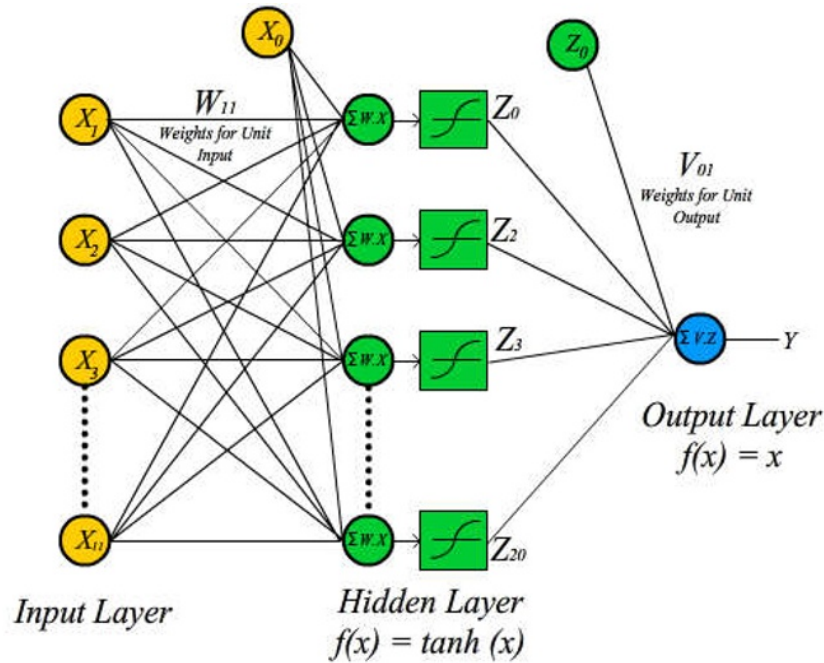


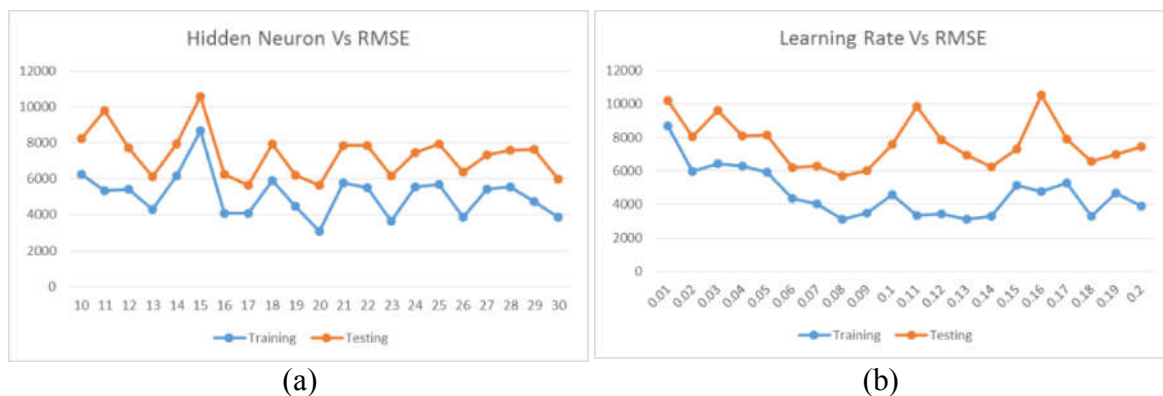
Fig. 1. Artificial Neural Network Architecture

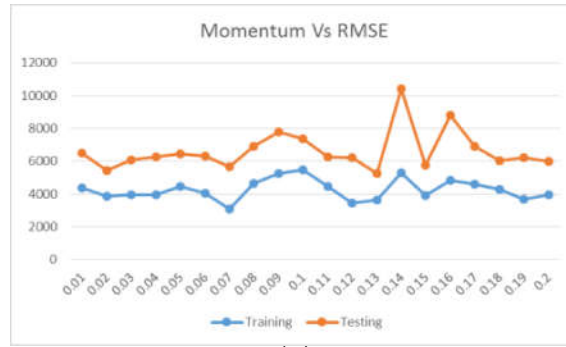
2.2 Program algorithm

The design of feed forward network then is used for back propagation learning. This learning is done by calculating error from each layer then propagated backward to correct weights and biases in network. Program algorithm used in the training of error back propagation is gradient descent. The first step is to initialize hidden weights and output [v] weights with a random value. Then enter input value to each layer of the neuron unit by following formula:

$$Y_j = f(v_{0j} + \sum_{i=1}^n x_i v_{ij}) \quad (5)$$

Calculation above is used in hidden neuron and output value, from these results then do the algorithm of back propagation of its output value. First step is calculate the error value from target minus output. Then calculate the output weights correction by equation (4) above where $[\delta_k]$ is output error, then these weights correction we calculate as the equation (5) to obtain the hidden output error. Hidden output error value is calculated again with equation (4) so that obtained hidden neurons correction. Then update the weights by summing old weights with new weights to the hidden neurons weights and output weights. The last step is to calculate the RMSE if it's appropriate, if it doesn't repeat the first step to get the expected value of the RMSE.





(c)

Fig. 2. a. Graphic of Relationship between Hidden Neuron and RMSE
 b. Graphic of Relationship between Learning Rate and RMSE
 c. Graphic of Relationship between Momentum and RMSE

3 Result discussion

3.1 Network parameter test

Network parameters test is performed by searching the best learning of combination execution between hidden layer, learning rate and momentum. Number of hidden layer neurons determine the ability of network to capture data patterns and produce nonlinear mapping between input and output. ANN software uses one hidden layer. Number of hidden layer neurons is determined through testing which is done on the network by varying number of hidden layer neurons to know network stability. Magnitude of RMSE network with numerous combination of hidden neurons is shown in Fig. 2.a, where the smallest RMSE value of hidden neurons is 20. The best learning rate value based on the results of RMSE experiments are between values range of 0.01 to 0.2. Based on Fig. 2.b and Fig. 2.c, it is shown that the best learning rate that is 0.08 and the best momentum is 0.07.

3.2 Cost estimation performance

Control valve cost estimation learning is performed with back propagation learning with number of nodes hidden neurons parameter are 20, learning rate 0.08 and momentum 0.07. Below Fig. 3 described RMSE value from learning process.

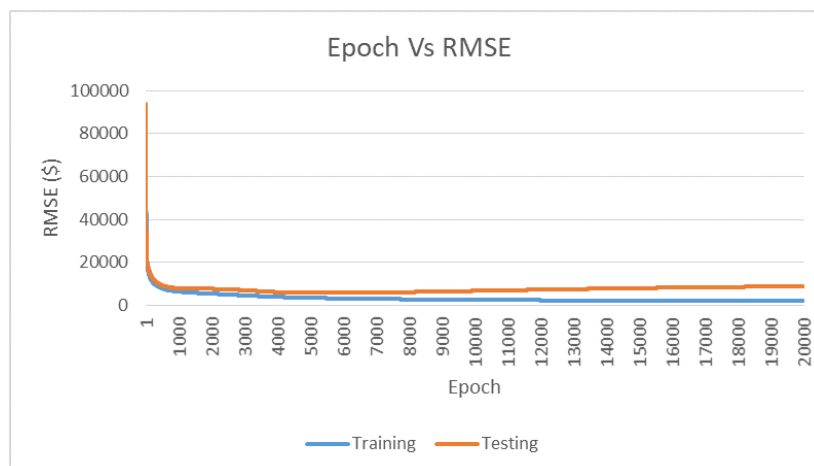


Fig. 3. Graphic of Relationship between RMSE and Epoch for Testing and Training Data

Fig. 3 shows the changing of training and testing RMSE value to epoch is decreasing and there is no overfitting, where training and testing value has a descent tendency. From Fig. 3 above also can be seen that the ANN system can generalize the control valve cost estimation system, so it can be used for testing data outside of the learning system.

Network validation is done by making a regression plot by calculating the correlation coefficient (R) value. Regression plot shows the relationship between network output value and target of observation data. Blue line shows the best match value of regression line between target and output. Correlation coefficient (R) value is a value which states linear relationship between target and output that has a value of 0 to 1. If the value approaches 1, then the linear relationship between target and output will be bigger. Fig. 4.a and Fig. 4.b shows the regression plot for training and testing data.

Fig. 4.a and Fig. 4.b below shows the results of training data correlation coefficient by 0.99998 while correlation coefficient of testing data is 0.94352. This indicate that the result of estimated value and actual value is same for training data, whereas for testing data ,the result of estimated value is approaching true value.

Testing data of 28 control valve shows 91.81% average in accuracy. ANN software has a standard deviation of 5.65% and the average error to 8.19%. This result is good enough to be applied in the control valve cost estimation on the bidding phase.

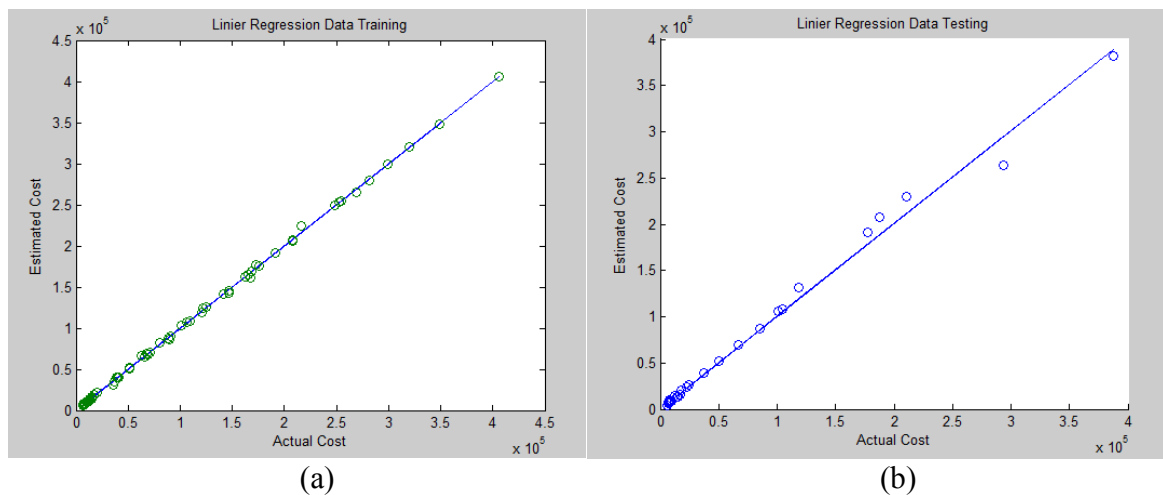


Fig. 4. a. Graphic of Relationship between Hidden Neuron and RMSE
b. Graphic of Relationship between Learning Rate and RMSE

4 Conclusion

Artificial neural network method with back propagation learning algorithm can be implemented to determine the advisability of control valve cost estimation in bidding phase. Back propagation algorithm has been trained well and provide the accurate output for testing data, so offered cost estimation that is given to the client has a small error. The characteristic of back propagation which is generalization, make this system can be used for early estimation on control valve cost determination. Correlation coefficient result of data testing is 0.94352, it means that the estimation value resemble the actual value. The expected output target is accurate in testing, as it occurs adjustment of weight value and better bias on learning process.

References

1. Hui Zhao, Weijie Huang, Shengping Wang, Tao Liu. *Intelligent Cost Estimation Method for Large scale Projects Based on PCA Optimized RBF Neural Network*. 8th International Symposium on Computational Intelligence and Design. IEEE. (2015)
2. H.D. Mehta, R.M. Patel, P.H. Trivedi. *Cost Estimation of Transformer Main Materials using Artificial Neural Networks*. Nirma University International Conference on Engineering. IEEE. (2012)
3. D.Manikavelan, R. Ponnusamy. *Software Cost Estimation by Analogy Using Feed Forward Neural Network*. Information Communication and Embedded Systems (ICICES). IEEE. (2014)
4. Lixue Xia, Rong Luo, Bin Zhao, Yu Wang, Huazhong Yang. *An Accurate and Low cost PM 2.5 Estimation Method Based on Artificial Neural Network*. Design Automation Conference, 20th Asia and South Pacific. IEEE. (2015)
5. L. Fausett. *Fundamentals of Neural Networks : Architecture, Algorithms and Applications*. Prentice Hall, Inc : New Jersey. (1993)
6. Zeyan Du, Binyong Li. *Construction Project Cost Estimation Based on Improved BP Neural Network*. International Conference on Smart Grid and Electrical Automation. IEEE. (2017)