Estimation of Construction Project Cost Based on GA-BPNN

Guangying Jin*[a] and Chunhui Yang[b]
School of Maritime Economics and Management, Dalian Maritime University, Dalian, China

Keywords: Construction Cost Estimation, Backward Propagation, Genetic Algorithm, Construction Enterprises, Project Management.

Abstract: Elevated uncertainty over recent years with the unprecedented crisis of COVID - 19 have led to greater difficulties for construction enterprises to control project costs. Meanwhile, estimating the project cost is sophisticated, especially considering about actual situation of their project management team. Thus, it is a challenge for construction enterprises to implement advanced decision support tools to estimate their project cost in the construction process. This paper proposes a model to help construction enterprises to estimate project costs. The model was developed using an approach by incorporating analytic hierarchy process (AHP) and backward propagation neural network (BPNN) optimized by the genetic algorithm (GA). Thus, it can involve subjective influencing factors and objective influencing factors. Moreover, a case study was carried out on actual projects by different construction enterprises located in Shandong province. The result shows that the implementation of the model can help construction enterprises to estimate construction costs with an average error of less than 8%. Consequently, the model can help managers estimate project costs taking into account the impact of the project management team.

1 INTRODUCTION

In 2022, Chinese national investment in real estate development decreased by 5.4% and the investment in housing decreased by 4.5% to 5,180.4 billion RMB from January to June through the China National Bureau of Statistics. As housing investment decreases, local Chinese construction companies will face greater and broader competition in the market.

Since the pandemic, uncertainties in the project management process have increased. Meanwhile, construction companies usually face problems such as increased costs and schedule delays (Abdel-Hamid, M. et al., 2021). Thus, calculating and controlling the cost of construction projects is more complex and time-consuming than before (Ahiaga-Dagbui et al, 2014).

Construction project cost, known as the sum of all costs of a construction project. Effective control of project costs is the key to ensuring corporate profits (Fazil, 2021). Therefore, construction project cost estimation remains a key part of construction project management (El-Kholy,2021).

Until now, the widely used method for cost estimation of construction projects has been modeling with Glodon software. However, the construction industry is highly data intensive and generates a large amount of data during the construction phase, the value of which is far from being fully utilized (Chakraborty, D.et al., 2020). The application of various data mining techniques has started to make people aware of the value of data in the construction industry. As a result, data mining techniques are widely used for construction project cost estimation (Elmousalami, 2020).

Hence, the main objective of this study is to propose an estimation method considering management factors and non-management factors to help construction managers estimate their construction project cost.

2 LITERATURE REVIEW

To study the application of data mining technology in the field of construction cost, we conducted a search of the relevant literature. The keywords searched were “estimate construction cost’’, ‘‘predict construction cost’’, and ‘‘construction cost management’’. As a

[a] https://orcid.org/0000-0002-4459-8425
[b] https://orcid.org/0000-0001-6441-9511

Jin, G. and Yang, C.
Estimation of Construction Project Cost Based on GA-BPNN.
DOI: 10.5220/0011917600003612
In *Proceedings of the 3rd International Symposium on Automation, Information and Computing (ISAIC 2022)*, pages 196-201
ISBN: 978-989-758-622-4; ISSN: 2975-9463

result, the word cloud in the abstract of the searched literature is shown in Figure 1.

In Figure 1, we can find that construction cost estimation has become an active field in the past decade (Mahmoodzadeh, A.et al.,2021). Meanwhile, the machine learning method is widely used in construction cost estimation （Fan, M et al.， 2021） . In more detail, we see that support vector machine techniques (SVM), case-based reasoning (CBR), decision tree, and neural network (NN) have been widely used.

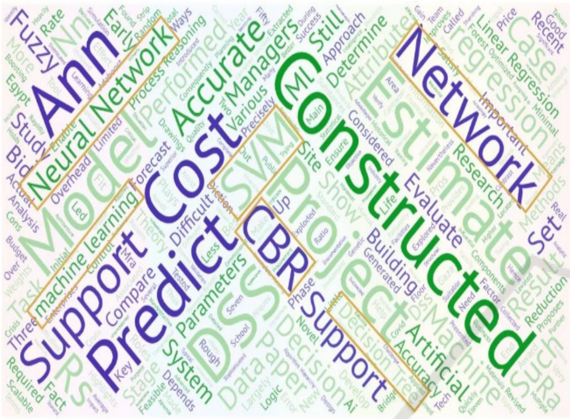


Figure 1: Word cloud of abstract of the article searched.

The methods mentioned above are all based on historical data, and hope to find certain reference value for new construction projects. As we all known, SVM performs well on small sample sets, CBR has a certain dependence on the previous field in prediction, and decision trees are prone to overfitting. Considering these characteristics, this study adopts the BPNN with strong fitting ability. In view of the shortcomings of the BPNN, we use GA to optimize it.

To the best of our knowledge, most relevant studies have focused on proposing construction cost forecasting models based on historical construction information. There are limited studies that consider the construction reality of construction firms to estimate the construction costs suitable for individual firms. However, estimating the cost of a project is important for project management (Rustamova, S., 2021). To bridge the gap, this paper proposes a model that considers the impact of management teams and estimates project costs for each firm.

3 DESIGN CONSTRUCTION COST ESTIMATION MODEL

3.1 Management and Non-Management Influencing Factors Selection and Evaluation

Both management-related factors and the characteristics of the project itself will affect the project cost, so we need to analyze these two factors.

Delphi method, also known as the expert opinion method, is a method of expressing opinions anonymously through multiple rounds of survey experts' opinions on the issues raised in the questionnaire. After repeated consultations, summaries, and modifications, the final aggregate into a consensus opinion of experts. In here, we use the Delphi method and invite experts in project cost management field to select the management factors and non-management factors that have a greater impact on the cost.

At the beginning, we invited experts to select the influencing factors related to management and obtained Table 1. In Table 1, we can find that the management factor (F1) includes three sub-factors (SF1.1---SF1.3). Among them, SF1.1 is the subjective score obtained through AHP, and SF1.2 and SF1.3 is the situation of the on-site management team obtained through investigation.

Table 1 Management-related influencing factors and reasons for choice given by experts.

Management influencing sub-factors (GA-BP model input Variable)	Reasons given by review experts in the field of construction.
Level of project management team (SF1.1)	The commitment of the site team to this project and the level of attention given to cost management will directly affect the cost.
Working years of cost manager on the project (SF1.2)	Experienced cost managers know which segments are prone to cost overruns and can take preventive measures.
The number of professionals with credentials on the project (SF1.3)	Qualified personnel with certain project management knowledge can better control and manage site materials, personnel, etc.

Moreover, a complete project management team consists of project general managers, business managers, production managers, technical managers, and so on. Each person has a different weight on the team. AHP is one of the most widely used multicriteria decision-making techniques (Luthra et al., 2016). Thus, AHP is used to analyse the importance of each project member.

In this research, experts were invited to compare the importance of two pairwise members and to rate the scale of importance of the chosen members by using the values of importance 1-9 scale (Saaty, 1990). Through the evaluation of project management experts, we get the judgment matrix. After that, we normalize the judgment matrix in columns and calculate the average of each row as the weight of members. Calculate each person's weight score based on how important each person is to cost management.

Further, we invited experts to select the non-management influencing factors. We get the non-management influencing factors in Table 2. The non-management factor (F2) includes eight sub-factors (SF2.1---SF2.8). Because the BPNN only handles digital inputs, we encode the character-type features, and the encoding results are also shown in Table 2.

3.2 GA-BPNN Design for Estimating Construction Costs

In this paper, BPNN is used to find a function to fit cost and impact factors. Thus, the input variables are the cost influencing factors and the output variable is the project cost. We show this in Figure 2. The training process of BPNN is divided into three steps:

Step (1): Forward Propagation.

Forward propagation proceeds in the direction from the input layer to the output layer. This process presents in Equation (1).

$$S_j = \sum_{i=1}^{m-1} w_{ij} x_i + b_j \quad (1)$$

Where: s_j is the input of activation function, m is the number of layer, w_{ij} is the weight between node i and node j, x_i is the value of node i, b_j is the threshold.

Then the node output is calculated by activation function, such as Equation (2).

$$x_j = f(s_j) \quad (2)$$

Where: x_j is the output value of the node.

Table 2 Content of non-management factors and coding of their characteristics.

Non-management sub-factors (Some other input variables of GA-BPNN)	GA-BPNN input Variable 's Category	Code
The type of foundation of building (SF2.1)	Raft foundation	1
	Independent foundation	2
	Pile foundation	3
The type of Structure of building (SF2.2)	Shear structure	1
	Frame structure	2
	Frame-shear structure	3
The number of above ground floor (SF2.3)	Low-rise residential	1
	Multi-storey residential	2
	Mid-rise residential	3
	High-rise residential	4
The number of below ground floor (SF2.4)	No basement	1
	1 Floor for basement	2
	2 Floors for basement	3
The building area of building (SF2.5)	Less than 2500 m^2	1
	2500 m^2-5000 m^2	2
	5000 m^2-7500 m^2	3
	7500 m^2-10000 m^2	4
	More than 10000 m^2	5
The type of interior wall decoration (SF2.6)	Plastered interior wall	1
	Paint interior wall	2
	Paint and plastered mix wall	3
The type of exterior wall decoration (SF2.7)	Paint exterior wall	1
	Brick exterior wall	2
	Paint and veneer brick mix wall	3
The type of ground decoration (SF2.8)	Installed underfloor heating	1
	Not installed underfloor heating	2

In here, we choose the Sigmoid function as the activation function, such as Equation (3).

$$x_j = \frac{1}{1 + e^{s_j}} \quad (3)$$

Step (2): Error Calculation.

The error between the predicted value and the real value is calculated as Equation (4).

$$E(w, b) = \frac{1}{2} \sum_{j=0}^{n-1} (o_j - y_j)^2 \quad (4)$$

Where: o_j is the output value, y_j is the real value.

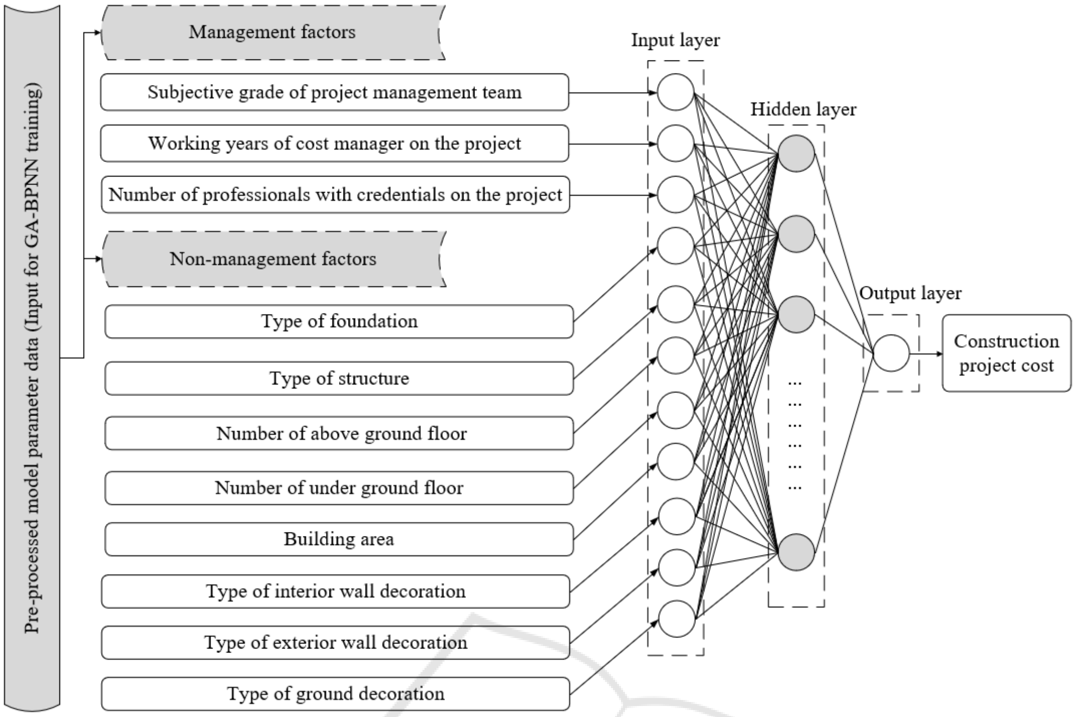


Figure 2: The structure of cost estimation BPNN model and the input and output variables.

Step (3): Error Back Propagation.
The backward propagation proceeds in the direction from the output layer to the input layer. In here, using the gradient descent algorithm as the learning algorithm of the network. The weight update formula is Equation (5).

$$\Delta w(i,j) = -\eta \frac{\partial E(w,b)}{\partial w(i,j)} \tag{5}$$

Where: η is learning rate.

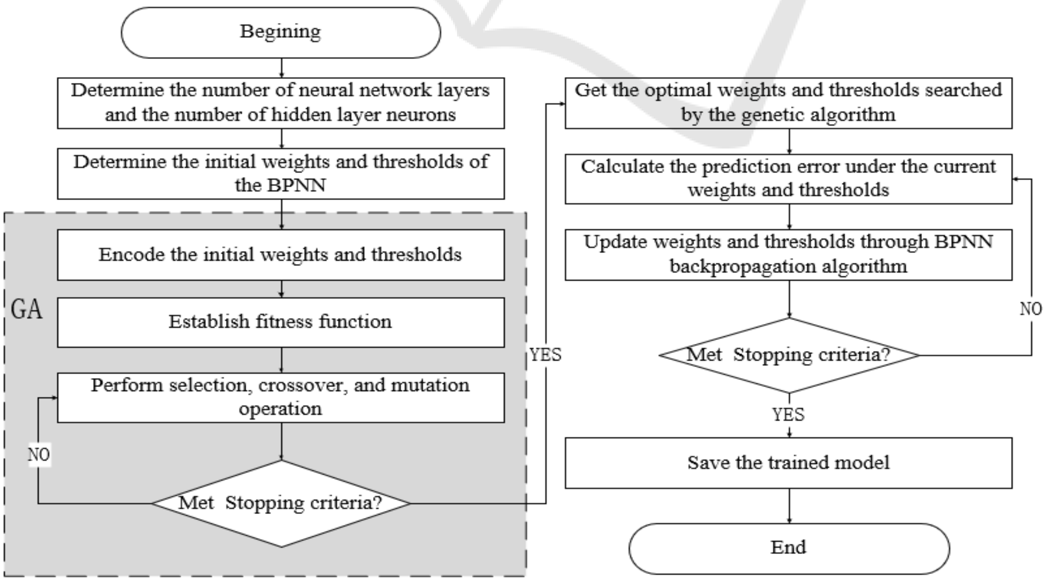


Figure 3: The structure of cost estimation BPNN model and the input and output variables.

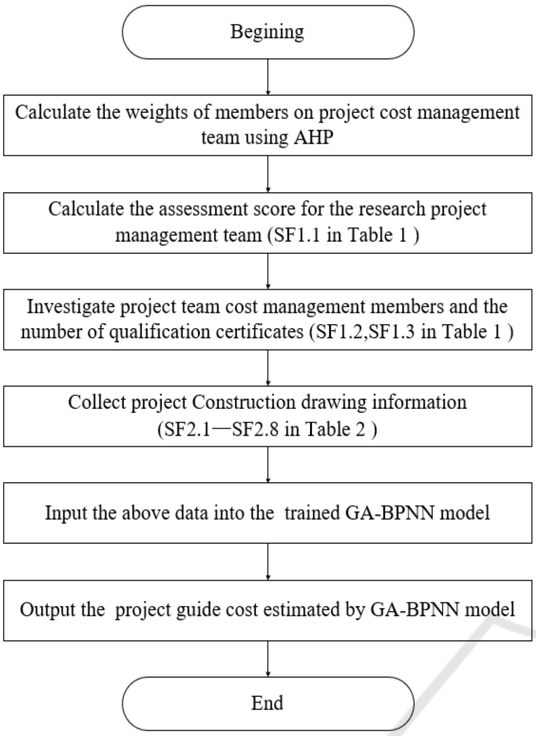


Figure 4: The work process of GA-BPNN model in estimating project cost.

When taking derivative layer by layer, follow the chain derivative Equation (6).

$$\frac{\partial E(w,b)}{\partial w(i,j)} = \frac{\partial E(w,b)}{\partial out}\frac{\partial out}{\partial net}\frac{\partial net}{\partial w(i,j)} \quad (6)$$

Where: ∂out is the output value of the previous level.

Finally, the updated weight w' is as Equation (7).

$$w' = w + \Delta w(i,j) \quad (7)$$

The network often cannot guarantee that the randomly set weights and thresholds can obtain the best results. GA is used to optimize the initial weights and thresholds of BPNN. As shown in Figure 3.

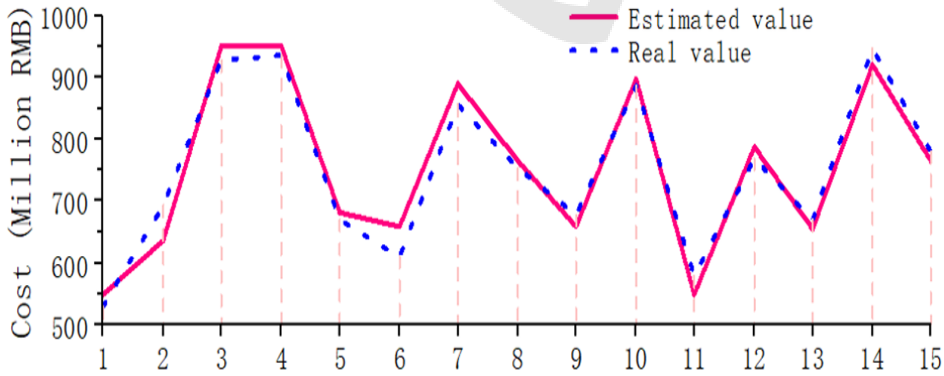
3.3 The Work Process of Construction Cost Estimation

In this study, we would like to explore the feasibility of using models to predict project costs while also considering managerial factors. The process design of the project cost estimation is shown in Figure 4.

4 CASE STUDY

As mentioned earlier, to train, validate and test the GA-BPNN model. we have utilized a database that we assembled based on civil construction projects located in Shandong province (China) as the research data.

We train the GA-BPNN model with the dataset. The fit of the model is high during the training process, and the R^2 of multiple training is above 0.92, which indicates that the constructed network can better best fit the parameters. Thus, we save the cost estimation model. To test the estimation effect of the model, we used another 15 projects to compare the estimation results. We calculate the average error rate to be within 8%. And the results are shown in Figure 5.



Test project	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Estimated value(Million RMB)	550.8	637.3	950.4	949.5	680.1	659.0	890.1	765.8	658.6	897.4	547.7	789.7	654.8	922.3	764.4
Real value(Million RMB)	531.6	692.6	926.8	934.4	671.8	609.9	854.5	754.3	678.4	886.3	578.8	767.4	669.6	945.6	778.5
Error rate(%)	3.62	7.98	2.54	1.61	1.23	8.05	4.17	1.52	2.93	1.25	5.37	2.90	2.22	2.46	1.81

Figure 5: Comparison of estimated and real values of GA-BPNN model and calculation of error rate.

Meanwhile, we compared the prediction results of other machine learning models. Each model is run 20 times and averaged. To assess the degree of applicability of different models to this problem, we counted the MAE, RMSE, and R^2 for each prediction.

Table 3: Comparison of prediction performance of different machine learning models.

NO.	Model	MAE	RMSE	R^2
1	Random Forest	177.42	250.78	0.73
2	SVM	94.68	152.67	0.91
3	BPNN	102.31	160.72	0.88
4	GA-BPNN	54.35	89.29	0.94

In Table 3, we can find that the GA-BPNN performs best on our data set, followed by SVM, and the random forest algorithm performs worst. These results show the applicability of the GA-BPNN model in estimating the cost of construction project.

5 CONCLUSION

This paper presents comprehensive descriptions of the proposed GA-BPNN model and its application in project cost estimation for construction enterprises. Meanwhile, we considered the influence of engineering project management factors and used the Delphi method to effectively select the factors that have a large impact on the project cost. And the weights of different members were calculated by AHP for project team level.

After simulation, the error is within the allowable range, which has certain guiding significance for managers to estimate the project cost according to their own project conditions. Thus, it maybe be feasible to consider the actual characteristics of the enterprise management team in the process of project cost estimation. Meanwhile, GA-BP model might be reliable in solving the problem of project cost estimation.

ACKNOWLEDGEMENTS

The study was supported by the Talent Research Start-up Founding of Dalian Maritime University, authorization code: 02502329.

REFERENCES

Abdel-Hamid, M., & Abdelhaleem, H. M. (2021). Project cost control using five dimensions building information modelling. International Journal of Construction Management, 1-10.

Ahiaga-Dagbui, D. D., & Smith, S. D. (2014). Rethinking construction cost overruns: cognition, learning and estimation. *Journal of financial management of property and construction*.

Chakraborty, D., Elhegazy, H., Elzarka, H., & Gutierrez, L. (2020). A novel construction cost prediction model using hybrid natural and light gradient boosting. Advanced Engineering Informatics, 46, 101201.

El-Kholy, A. M. (2021). Exploring the best ANN model based on four paradigms to predict delay and cost overrun percentages of highway projects. International Journal of Construction Management, 21(7), 694-712.

Elmousalami, H. H. (2020). Data on field canals improvement projects for cost prediction using artificial intelligence. Data in brief, 31, 105688.

Fan, M., & Sharma, A. (2021). Design and implementation of construction cost prediction model based on SVM and LSSVM in industries 4.0. International Journal of Intelligent Computing and Cybernetics.

Fazil, M. W., Lee, C. K., & Tamyez, P. F. M. (2021). Cost estimation performance in the Construction Projects: A systematic review and Future Directions. International Journal of Industrial Management, 11, 217-234.

Luthra, S., Mangla, S. K., Xu, L., & Diabat, A. (2016). Using AHP to evaluate barriers in adopting sustainable consumption and production initiatives in a supply chain. International Journal of Production Economics, 181, 342-349.

Mahmoodzadeh, A., Mohammadi, M., Daraei, A., Farid Hama Ali, H., Ismail Abdullah, A., & Kameran Al-Salihi, N. (2021). Forecasting tunnel geology, construction time and costs using machine learning methods. Neural Computing and Applications, 33(1), 321-348.

Rustamova, S. (2021). Application of PMI PMBOK standard in the development of the Construction cost management projects plan. International journal of trends in business administration, 11(1).

Saaty, T. L. (1990). The analytic hierarchy process in conflict management. International Journal of Conflict Management.