Research on Cost Prediction of Power Transmission and Transformation Project Based on Combination Prediction Model

Li Ma*, Qixin Wang, Shuang Liao, Ting Guo, Yi Xiong, Yue Ming and Yuxin Zou
State Grid Hubei Electric Power Company Limited Economic Research Institute, Wuhan, China

*Corresponding author: mal12@hb.sgcc.com.cn

Abstract. Based on the project division of power transmission and transformation projects, this paper analyzes a large number of practical engineering cases, and summarizes the influencing factors of substation and overhead line project cost combined with the practical engineering experience of many experts. A number of cost control indicators are extracted, and a combination prediction model of project cost based on neural network and support vector machine is constructed. The actual project data are taken as samples for empirical research. The results show that the combination prediction model is feasible and accurate for the prediction of power transmission and transformation project cost, which provides a new idea and implementation method for the prediction of power transmission and transformation project cost.

Keywords: Power transmission and transformation project, Neural network, Support vector machine, Forecast combination.

1. Introduction
In the past, the prediction of power transmission and transformation project cost mainly depended on the actual analysis and operation of technicians with many years of practical experience in this field. However, the cost of power grid project is a multivariate and highly nonlinear problem, especially when the project situation is complex and changeable, it is difficult to obtain reliable results of single project through empirical estimation to guide the project cost control. With the rise of artificial intelligence research in recent years, intelligent algorithms have been recognized again. Intelligent algorithm is an algorithm that scientists are inspired by natural laws such as human nervous system, honeycomb ant colony organization and biological genetic law, and simulate its operation process to solve corresponding problems. Intelligent algorithm can obtain unknown information from the existing information, but it has certain information value, and express it into a form that can be understood by human beings. This paper attempts to use a new combination optimization algorithm to avoid the shortcomings of the past power transmission and transformation project cost prediction by combining feature extraction and historical data learning, and effectively predict the main technical and economic indicators of its cost and guide the project cost management.
2. Research status of power transmission and transformation project cost prediction
The balance rate of power grid project investment can be traced back to the control research of power grid project investment estimation and budget estimation. By analyzing the influencing factors of power grid project investment estimation and budget estimation, establishing a reasonable estimation and budget prediction model can effectively achieve the goal of controlling the balance rate. Nowadays, in addition to the traditional list pricing method and quota pricing method, the mainstream cost forecasting method in domestic academic circles also introduces fuzzy mathematics, machine learning algorithm and other methods, and introduces the knowledge of mathematics and computer into the field of project cost forecasting, which has also achieved good academic results.

In view of the problem that there are many kinds of projects in the early stage of transmission line engineering, the deviation of cost estimation is large, and how to use a small amount of information to get more accurate project cost, many scholars combined with BP neural network algorithm to build a transmission line project cost prediction model, which has been widely used [1, 2, 3, 4]. Neural network, support vector machine and other intelligent algorithms have been successfully applied in many fields, and foreign power grid companies with high management level have used these intelligent algorithms to predict and evaluate the project cost [5, 6, 7, 8, 9].

Intelligent algorithm is a kind of "soft computing", which can get the algorithm model by simulating the habits, behaviors and body functions of animals. At present, the basic intelligent algorithms include artificial neural network, support vector machine, genetic algorithm and particle swarm optimization algorithm. Neural network, support vector machine and other intelligent algorithms have been successfully applied in many fields, which lays the foundation for the selection of prediction model in this paper.

3. Research methods of power transmission and transformation project cost prediction

3.1. Neural network
After collecting a large number of key data of historical power transmission and transformation projects, BP neural network can be used to build artificial intelligence project cost prediction model. In the training stage, we need to use a large number of historical power transmission and transformation project data. After data preprocessing, we can get the numerical data without missing value of the corresponding project as the input.

Before the model is used to predict new project data, it is necessary to train it with historical project data. By continuously fitting historical sample data and adjusting the calculation parameters of each layer of neural network, the relationship between related factors can be learned from a large number of historical project samples. The trained neural network model is equivalent to a complex function formula $y = f(x)$. By inputting the project data $X$ to be predicted, the static investment $y$ to be predicted can be calculated.

The training process includes network structure design, super parameter setting, weight initialization and updating model parameters by gradient descent algorithm.

1) Network structure design
Neural network includes input layer, hidden layer and output layer. A circle in each layer represents a processing neural unit, several neural processing units form a layer, and several layers form a network, namely "neural network", as shown in the figure below:
(2) Super parameter setting and weight initialization

Before training, neural network needs to set super parameters such as iteration times and learning rate, which need to be adjusted according to different training sets and actual training effect.

The weights of each layer of neural network will be adjusted continuously in the process of training, but the initial values need to be given. Generally, uniform distribution or random initialization is used.

(3) Network training

In the neural network training, the prediction value of the current model is calculated by forward propagation. According to the final loss function, the partial derivative of each parameter is calculated by gradient descent algorithm, and the parameters are updated to realize back propagation, so as to make our model more accurate. The loss function represents the deviation between the predicted value and the real value. Through the idea of gradient descent algorithm to update the weight of the network model, the loss function constantly approaches the global optimal value, that is, the minimum value of the loss function. The weight of gradient descent update is shown in the figure.

The abscissa represents the weight of the model and the ordinate represents the loss function. As shown in the figure, the weight of the model is updated towards the minimum value of the loss function.
3.2. Support vector machine

The power transmission project cost prediction model based on support vector machine can be equivalent to the function mapping based on pattern matching. Independent variable, \( x_i, i = 1, 2, N \), of the vector \( X = \{x_1, x_2, \ldots, x_n\} \), which represents the value of the factors affecting the cost of the power transmission and transformation project, and \( X_m = \{x_{m1}, x_{m2}, \ldots, x_{mn}\} (m = 1, 2, \ldots, n) \) represents the index value of the factors affecting the cost of the power transmission and transformation project in the m year. The dependent variable of the model is the cost index value of the power transmission and transformation project, which is set as \( y_i \), representing the predicted value of the i year. For the relation between \( y_i \) and \( X = \{x_1, x_2, \ldots, x_n\} \), suppose there is a nonlinear mapping \( F \) such that. Taking \( X_m \) as the input independent variable and \( y_i \) as the output, the regression equation is solved by support vector machine:

\[
f(x) = \sum_{i=1}^{n}(\alpha_i^* - \alpha_i)K(x_i, x) + b^*
\]  

The main steps are as follows:

1. The sample data of transmission and transformation cost impact index value is selected to determine the sample matrix and process the sample data;
2. At the same time, an error limit is set, and the optimized parameters are obtained through cross validation \((C, \sigma)\). The specific method is to put different \((C, \sigma)\) into the prediction model to verify the model, and compare the predicted value of power transmission project cost obtained by the calculation model with the actual value to calculate the mean square error \((\text{MSE})\);
3. \((C, \sigma)\) corresponding to the minimum mean square error between the predicted value of the model and the actual cost value is taken as the optimized \((C, \sigma)\) value;
4. The optimized \((C, \sigma)\) is substituted into the model to detect the sample data. For the kernel function, by substituting the \((C, \sigma)\) parameters obtained by support vector machine and kernel function into equation (1), the prediction model of power transmission project cost can be obtained as follows:

\[
f(x) = \sum_{i=1}^{n}(\alpha_i^* - \alpha_i) \left(-\frac{\|x_i-x\|^2}{2\sigma^2}\right) + b
\]

4. Empirical Study on cost prediction of power transmission and transformation project

4.1. Data sources and processing

The typical project selected in this study is the real 110kV power transmission and transformation project data of a province from 2016 to 2019. According to statistics, 580 samples were used in training. Since the modeling is in the stage of verification and testing, this stage only focuses on the data preprocessing of new substation projects and overhead lines which have a large proportion of the cost of power transmission and transformation projects. There are 218 substation projects and 362 overhead line projects in the training samples.

According to the experience of experts, the factors of substation engineering are: Survey and design cost, building area of the whole station, unit price of main transformer equipment, land acquisition area, site leveling cost, unit price of high voltage side circuit breaker, control cable quantity, retaining wall and slope protection cost, number of outgoing lines at medium voltage side, building area of main control building, power cable quantity, number of high voltage side circuit breaker, number of main transformer in current period, Number of outgoing lines at high voltage side, cost of foundation treatment, unit price of circuit breaker at low voltage side, number of circuit breakers at medium voltage side, number of circuit breakers at low voltage side, unit price of circuit breaker at medium voltage side, number of outgoing lines at low voltage side.

The influencing factors included in the model study of overhead line engineering include: OPGW price, tower material quantity, conductor quantity, number of high voltage side circuit breakers, concrete quantity, strain ratio, tower material price, tower base quantity, total line length, foundation steel
quantity, proportion of high strength steel tower material, survey and design fee, Double circuit length, bidding fee, project preliminary work fee, project legal person management fee, construction loan interest, total tower base, single circuit length, total line length.

4.2. Design of model weight
In order to make the prediction result more accurate, the combination model will be used to predict the static investment of substation project. Among them, the combination model includes neural network and support vector machine. The reciprocal of the square sum of the error between the predicted value of each model prediction set and the actual value of the training set is taken as the weight.

Taking a group of samples as an example, the average deviation between the predicted value of neural network and support vector machine model and the actual value is as follows:

| Model type          | Static investment real value (ten thousand yuan) | Static investment projections (ten thousand yuan) | Mean deviation rate |
|---------------------|-----------------------------------------------|--------------------------------------------------|---------------------|
| The neural network  | \( A_i \)                                      | \( B_1i \)                                       | \( F_1=(B_1i-A_i)/ Ai \) |
| Support vector machine | \( A_i \)                                   | \( B_2i \)                                       | \( F_2=(B_2i-A_i)/ Ai \) |

The calculation method of the combined weight of various model results is as follows:

| Model type          | Sum of squares of deviation | Sum squared of deviation | The weight (normalization of deviation square and inverse) |
|---------------------|-----------------------------|--------------------------|----------------------------------------------------------|
| The neural network  | \( (B_1i - A_i)^2 \)        | \( D_1=1/ (B_1i - A_i)^2 \) | \( W_1=D_1/(D_1+D_2) \) |
| Support vector machine | \( (B_2i - A_i)^2 \)       | \( D_2=1/ (B_2i - A_i)^2 \) | \( W_2=D_2/(D_1+D_2) \) |

According to the weight obtained by the above calculation method, the predicted value and predicted interval of each model are combined and weighted to get the final deviation rate as follows:

\[ F=F_1*W_1+F_2*W_2 \]

Through the deviation rate calculated in the portfolio model, combined with the actual static investment, the predicted static investment of the project is obtained.

4.3. Combination prediction model calculation.
(1) Substation engineering

Taking 5 substation projects verified by the prediction model as examples, the combined model calculation is carried out by using the predicted results of the neural network and support vector regression prediction model. The results are shown in the table below:
Table 3. Calculation table of static portfolio prediction model of substation

| The serial number | Single Project Name        | Voltage grade (kV) | Final accounts of the static investment (ten thousand yuan) | The neural network Deviation rate F1 (%) | Support vector regression Deviation rate F2 (%) | The weight W1 | The weight W2 | Combination deviation rate (%) |
|-------------------|---------------------------|--------------------|-------------------------------------------------------------|----------------------------------------|-----------------------------------------------|---------------|---------------|-------------------------------|
| 1                 | A Substation engineering  | 110                | 4966.72                                                     | 0.63                                   | 97.18%                                         | 3.70          | 2.82%         | 0.72                          |
| 2                 | B Substation engineering  | 110                | 3322.72                                                     | 17.43                                  | 5.39%                                         | 4.16%         | 94.61%        | 4.88                          |
| 3                 | C Substation engineering  | 110                | 3939.30                                                     | 1.37                                   | 49.60%                                        | 1.36          | 50.40%        | 1.36                          |
| 4                 | D Substation engineering  | 110                | 2975.61                                                     | 1.40                                   | 5.38%                                         | 0.33          | 94.62%        | 0.39                          |
| 5                 | E Substation engineering  | 110                | 4279.83                                                     | 2.62                                   | 2.95%                                         | 0.46          | 97.05%        | 0.52                          |

It can be seen from the above table that the error rate calculated by various artificial intelligence prediction models, such as neural network and support vector machine, is calculated by the combined prediction model and given a certain weight value to get the combined error rate. From the results, the combination error rate is within the range of 5%, and the result is good.

(2) Overhead line engineering

Taking 4 overhead line projects verified by the prediction model as examples, the combined model calculation is carried out by using the predicted results of the neural network and support vector regression prediction model. The results are shown in the table 4.

It can be seen from the above table that the error rate calculated by various artificial intelligence prediction models, such as neural network and support vector machine, is calculated by the combined prediction model and given a certain weight value to get the error rate of the combined prediction model. According to the results, there are 2 errors within 5% and 2 within 10%-16% of the combination prediction, indicating a good result.

Table 4. Calculation table of line engineering static portfolio prediction model

| The serial number | Single Project Name        | Voltage grade (kV) | Line length (single fold) km | Final accounts of the static investment (ten thousand yuan) | The neural network Deviation rate F1 (%) | Support vector regression Deviation rate F2 (%) | The weight W1 | The weight W2 | Combination deviation rate (%) |
|-------------------|---------------------------|--------------------|------------------------------|-------------------------------------------------------------|----------------------------------------|-----------------------------------------------|---------------|---------------|-------------------------------|
| 1                 | A Substation engineering  | 220                | 20                           | 2759.40                                                     | 1.98                                   | 47.70%                                         | 1.89          | 52.30%        | 1.93                          |
| 2                 | B Substation engineering  | 220                | 64                           | 6149.45                                                     | 5.84                                   | 2.31%                                         | 0.90          | 97.69%        | 1.01                          |
| 3                 | C Substation engineering  | 220                | 8                            | 1314.71                                                     | 10.62                                  | 67.15%                                        | 15.18         | 32.85%        | 12.12                         |
| 4                 | D Substation engineering  | 220                | 48                           | 4758.83                                                     | 13.98                                  | 60.53%                                        | 17.31         | 39.47%        | 15.29                         |
5. Conclusions
Based on the existing problems of power transmission and transformation project cost management and control, based on the analysis of a large number of practical engineering case cost levels, experts and combining with practical engineering experience, to produce several substations and overhead line construction cost control points, combined with intelligent algorithms, such as neural network and support vector machine (SVM) based on neural network and support vector machine (SVM) is constructed of combination forecast model of engineering cost, and with the actual engineering data as samples for empirical research, the empirical results show that this combination forecast model accuracy is high, has the very good practicability and feasibility, It provides a new and feasible solution to the increasingly serious problem of power transmission and transformation project cost prediction.

Acknowledgments
This work was financially supported by Science and Technology Project of State Grid Hubei Electric Power Company Limited Economic Research Institute (521538200007)

References
[1] Ling Yunpeng, Yan Pengfei, Han changzhan, et al. Prediction model of transmission line project cost based on BP neural network [J]. China electric power, 2012, 45 (10)
[2] Xu Hongdong. Research on cost prediction of power transmission and transformation project based on artificial neural network [D]
[3] Sun Anli, Xiang Chun, Wu enthalpy Xi. Research on transmission line project cost prediction model based on BP neural network [J]. Modern electronic technology, 2018
[4] Wang mianbin, Geng Pengyun, an Lei. Establishment and implementation of power engineering cost prediction model based on ANN [J]. Modern electronic technology, 2017 (24): 174-176
[5] Shayeghi H, Ghasemi A, Moradzadeh M, et al. Day-ahead electricity price forecasting using WPT, GMI and modified LSSVM-based S-OLABC algorithm [J]. Soft Computing, 2017, 21(2): 525-541.
[6] Razak I A B W A, Abidin I B Z, Siah Y K, et al. Support vector machine for day ahead electricity price forecasting [J]. 2015.
[7] Kaur M A, Sawhney B K , Gupta O P . To Optimize the Cost of Electricity Generated by Hydro Power Plant Using Artificial Neural Networks [J]. Ersra Publications.
[8] Alanis A Y . Electricity Prices Forecasting using Artificial Neural Networks [J]. IEEE Latin America Transactions, 2018, 16(1): 105-111.
[9] Tang Q, Gu D. Day-Ahead Electricity Prices Forecasting Using Artificial Neural Networks [C]// International Conference on Artificial Intelligence & Computational Intelligence. IEEE, 2010.