ISSN 2522-3488 Vol. 6, Issue 1: 40-43, DOI: 10.25236/IJNDES.2022.060107

Research on cost prediction of civil engineering construction based on multi-dimensional data mining

Cui Lixiang

Zhengzhou Business University, Henan, Gongyi, China
clx15085904838@163.com

Abstract: *The accuracy of the cost prediction of civil engineering is an important step in the refined management of modern civil engineering. Therefore, a research on the cost forecast of civil engineering based on multi-dimensional data mining is proposed, which combines data mining technology and cost forecasting. Based on the results of data mining, the predictive data is preprocessed, and the factor that has the greatest impact on the cost is screened out, which is the installation cost of the building in the project. And design a fitting function to preprocess the data. Establish a civil engineering cost prediction model. The first step is to determine historical indicators, search for corresponding historical data based on actual engineering data, and transform the prediction characteristics of the data, and finally complete the civil engineering cost prediction, and finally prove it through test experiments. The method's engineering cost prediction accuracy is relatively high.*

Keywords: *multidimensional data; data mining; civil engineering; construction engineering; engineering cost; cost forecast;*

1. Introduction

The main source of funds for civil engineering projects is construction project investment. Investment is the source of ensuring the normal operation of the project and an important support for the normal operation of the enterprise. The most important thing for investment is to make cost forecasts. Correct investment decisions are based on the premise of accurate cost forecasts. Accurate cost forecasts can bring considerable benefits to enterprises. The plan design of the investment decision of the construction project is carried out in the early stage of the engineering construction design. The prerequisite for the accurate estimation of the project cost is that the project information is clear [1], and the market price of each project material is mastered in order to make a more stable Accurate cost prediction. Only when the error of cost prediction is within a reasonable range, can it be helpful for civil engineering to make investment decisions. The cost prediction of civil engineering is an important part of project planning and construction. Therefore, when predicting the cost of the project, try to understand the situation in the project as much as possible.

Improving the accuracy of cost estimation is also conducive to quoting the construction unit, especially in the context of modernization [2], the quality of the technology and raw materials mastered by the engineering team can be guaranteed, and the level of cost is directly related to competition. The size of the force, in the absence of time for the preparation of bid documents. Enhancing one's own competitive strength through the accuracy of cost forecasting. This article proposes the use of multi-dimensional data mining to predict the cost of civil construction projects, which can quickly model the factors that affect the cost and finally realize the estimate of the cost. Improve the speed and stability of civil engineering cost prediction.

2. Preprocessing of prediction data based on data mining

From a market perspective, the project cost includes the total cost of all actions to complete the project, including land prices, equipment purchases and regular lease prices, purchases of raw materials, project contracting, and technology upgrades. Project investors hope to use reduced investment to build projects with high economic returns. Therefore, the principle of data processing for project costs is to deal with unnecessary expenses. In civil engineering, the project that consumes the most capital is the installation cost [3]. Therefore, the cost of the most consuming project should be minimized when

ISSN 2522-3488 Vol. 6, Issue 1: 40-43, DOI: 10.25236/IJNDES.2022.060107

making cost estimates. The composition of the total cost of the installation project is as follows:

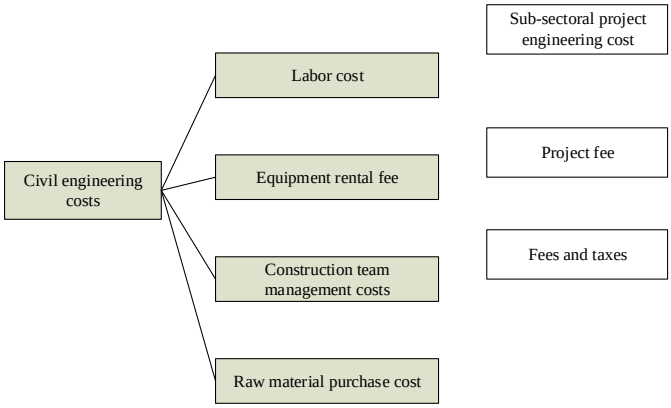

Figure 1 Components of installation costs

It can be seen from Figure 1 that in the civil construction project cost forecast designed in this paper, there are many participants involved in the project. In the entire life cycle of the building, the early construction cost estimation is more important, and data mining is used to screen the forecast data. In [4], it is possible to establish a screening model for civil engineering cost prediction, and establish a corresponding estimation function. Classify and mine data of different concepts, and predict potential factors and demands from a multi-dimensional perspective. Non-linear data filtering when supporting vectors has unique advantages in small sample predictions. Therefore, the following fitness functions can be used for data processing:

$$y = a \bullet b(x) + c \quad (1)$$

In the above formula, y is the degree of fit function, a is the data feature for data mining, b is the slope of the function, x is the predictive vector value of the cost, and c is a constant [5]. The closer the fit curve is to a straight line, the higher the fit, and the predicted data factor is a factor that has a higher correlation with the cost.

3. Establish a model for predicting the cost of civil engineering construction

3.1. Determination of historical indicators

Because the cost prediction is not only the need to mine and process the existing data of the project, but also the feature extraction of the historical data. The determination of the feature index should be the actual data as the source data of the data processing, and the historical data comes from the national construction project cost. Monitoring platform, in order to avoid excessive calculation of the model, it is necessary to delete unrelated data [6], duplicate data and wrong data, and ensure the accuracy of the original data through computer verification.

3.2. Carry out data prediction feature conversion

The data for establishing the cost forecast of civil engineering and construction projects is processed with normalized characteristics, and qualitative indicators are transformed into quantitative indicators. The processing formula is:

$$I = \frac{x_i - x_1}{x_2 - x_1} \quad (2)$$

In formula (2), x2 is the maximum construction cost of the project, and x1 is the minimum construction cost of the project. Xi is the most suitable cost, the index benefit relationship between the three is I, and the feature conversion of data is to reduce the amount of noise. Therefore, the model established after the feature conversion of the data is very necessary [7-8]. Through the calculation of this formula, the best benefit relationship can be obtained. The data not only uses the actual data of the current project, but also uses the historical data after data processing as the data source of the model, so the calculated cost is more accurate.

ISSN 2522-3488 Vol. 6, Issue 1: 40-43, DOI: 10.25236/IJNDES.2022.060107

4. Test experiment

In order to verify the practicability of the cost prediction method designed in this paper, a test experiment is designed to compare the cost prediction method designed in this paper with the traditional cost prediction method, and compare the prediction accuracy of the traditional method and the design method in this paper.

4.1. Experiment preparation

The actual data of a certain project is used as the sample data of the experiment. There are 300 sample data in total, and the 300 data are divided into 200 training data and 100 test data. In the model training, 200 pieces of training data are used to train the prediction model for cost estimation, in order to avoid fitting problems. Normalized sample training is performed on the data, and the training results are shown in Table 1:

Table 1 Normalized training samples

Sample set number	Cost index	Project area	structure type	Basic type	Engineering standards	Proportion of raw material steel bars	Unilateral cost
A1	-0.6120	-1.09	-0.1	1.21	1	-0.91	0.02
A2	-0.3371	-2.03	-0.1	1.37	1	-0.13	0.05
A3	-0.5221	-2.07	-0.2	1.56	0	-0.14	0.02
A4	-0.5541	-2.41	-0.2	1.43	0	-0.57	0.02

Table 1 is the training results of the normalized samples. After the samples are trained, the next step can be tested.

4.2. Test Results

After five sample trainings, the cost estimation method designed in this paper and the traditional BP civil engineering cost prediction method and the actual cost of fit test results are shown in Figure 2:

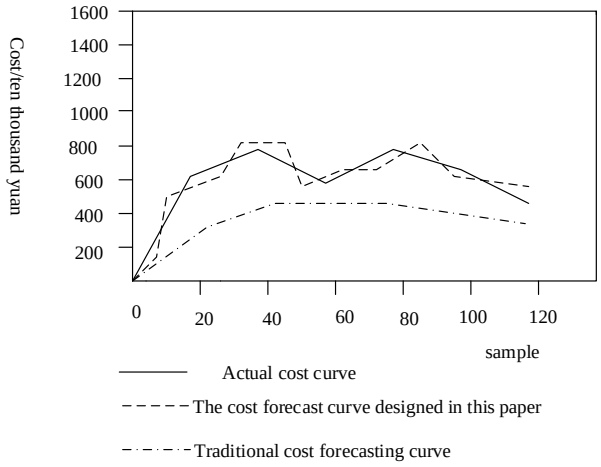

Figure 2 Test results

The experimental results are shown in Figure 2. From the curve in the figure, it is obvious that the curve fit between the method designed in this paper and the actual project cost is higher, and the prediction accuracy of the method designed in this paper is higher.

5. Conclusions

This article is based on multi-dimensional data mining technology to predict the cost of civil engineering. Compared with the traditional cost prediction method, the prediction accuracy is higher, indicating that the prediction method of this article has a certain value. In terms of collection, the indicators can be further adjusted to ensure the accuracy of data collection. Because the sample project data is limited, it is hoped that in the future research, we can refine the deficiencies and contribute to the accuracy of the project cost.

ISSN 2522-3488 Vol. 6, Issue 1: 40-43, DOI: 10.25236/IJNDES.2022.060107

References

[1] Yuan Fang. Construction project cost prediction method based on time series[J]. Construction and Budget, 2021(11): 35-37.
[2] Wang Yibin, Zou Xiaohong. Research on the rationality review of construction engineering cost based on FCM and LSSVM[J]. Journal of Engineering Management, 2021, 35(04): 25-29.
[3] Dong Na, Lu Sihua, Xiong Feng. ABC-SVM-based construction project cost prediction under the background of big data[J]. Technical Economy, 2021, 40(08): 25-32.
[4] Chen Yuehua, Zheng Simin, Liu Wenlu. Research on industrial construction cost prediction based on machine learning [J]. Journal of Wuhan University of Technology (Information and Management Engineering Edition), 2021, 43(04): 314-321.
[5] Liu Weijun, Li Nian. Based on MEA-BP neural network construction project cost index prediction [J]. Engineering Economics, 2021, 31 (01): 23-27.
[6] Teng Lingyun. Research on construction of construction project cost prediction model based on BP neural network [J]. Housing Industry, 2020(12): 110-113.
[7] Wang Zheng, Li Ping, Tan Xiaorong. Discussion on the application of BIM technology in construction engineering cost management[J]. Ju She, 2020(07): 154.
[8] Zhang Lange. Design of rapid estimation model of project cost based on BIM technology[J]. Journal of Changchun Institute of Technology (Natural Science Edition), 2019, 20(03): 5-8.