1C-05

Machine Learning Application for Plant Cost Engineering Support System

Masafumi Furukawa, Yusuke Nakaya
Faculty of Information Technology and Business
Cyber University

1. Introduction

Plant cost engineering consists of basic designs, specification preparations, inquiries for quotation, proposals for inquiry and contracts. Recently, the requirements for international procurements have been increasing because of the increase of global plant investments. These activities require well-trained professional engineers and certain working periods. To minimize these labor fees and time, a server side application, ***autoQuota***, is developed.

2. Functions of the System

There are three functions of this system. First, ***autoQuota*** rationalizes cost engineering and basic design works in plant engineering projects. ***autoQuota*** completed plant cost engineering transactions 20 times faster than the existing transaction procedures.

Second, ***autoQuota*** collects engineering data through the transactions into a relational database management system (hereinafter RDBMS). The test operation was conducted for 40 days. The data obtained were 147 of 4 kinds and 21 categories equipment data from 48 Chinese makers. This result is extremely shorter in operation time than the existing data collecting procedures. The cost reduction is estimated to be approximately US$1,600- per one transaction.

Third, ***autoQuota*** estimates costs by a machine learning algorithm. This report focuses on the details of the machine learning algorithm and its theories beneath.

3. Machine Learning Algorithm

The plant equipment price and its capacity factor, i.e. Motor kW, have the following relation.

$$P_0 = bW_0^a \quad \cdots \quad (1)$$

Where:
P_0: Price in US dollar.
W_0: Motor capacity in kW.
a,b: Constants.

Equation (1) is linearized to derive the following equation to calculate the constants a and b by the linear regression algorithm.

$$\log P_0 = a \log W_0 + \log b \quad \cdots \quad (2)$$

Cost estimation starts with a buyer side specification. ***autoQuota*** receives, modifies and implements the specification, and then stores specification data in RDBMS. ***autoQuota*** selects significant data from the database, then starts linear regression calculations to obtain the constants a and b by the least-square method.

If the correlation factor is less than the value of 5% level of significance or 0.7, the worst fitting data is removed from the data list. The linear regression is repeated with the revised data to satisfy the correlation factor requirements.

At the same time, the motor capacity, which is the variable to estimate the cost, is calculated by the following equation.

$$W_0 = \frac{Q_0 H}{c\eta} \quad \cdots \quad (3)$$

Where:
Q_0: Fluid capacity (m^3/h)
H: Total head (m)
c: Constant
η: Equipment efficiency

The equipment price is calculated by Equation (2) with the motor kW in Equation (3). The simplified flowchart of

the algorithm is shown as Figure 1.

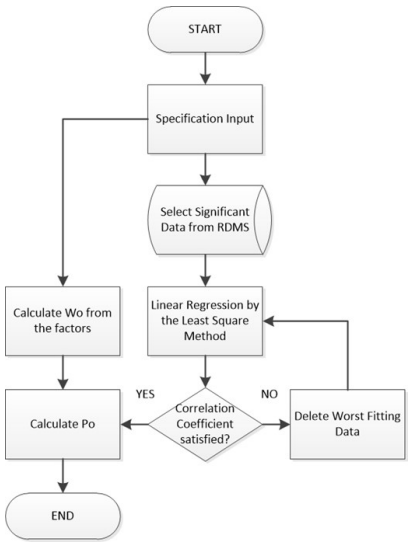

Figure 1 Algorithm Flowchart

4. Results and Considerations

To verify ***autoQuota*** estimation, the results of autoQuota is compared with a manual cost calculation. There are three compared items, which are R^2 values, the error over price percentages by the scatter plot diagram (Figure 2), and the comparisons to the actual quoted prices (Table 1).

The sample equipment is a small size carbon steel centrifugal pump with $5m^3/h$ capacity and 15m head.

R^2 value manually calculated was 0.3384. On the other hand, ***autoQuota*** obtained 0.5426. The calculated equations are as follows.

Manual Calculation:

$$\log P_0 = 0.6192 \log W_0 + 2.8035 \quad \cdots \quad (4)$$

autoQuota:

$$\log P_0 = 0.6098 \log W_0 + 2.8656 \quad \cdots \quad (5)$$

The error over price diagram also proved that ***autoQuota*** can obtain better results than the manual calculations. The manual calculation reached over 300% errors, while ***autoQuota*** does 200%.

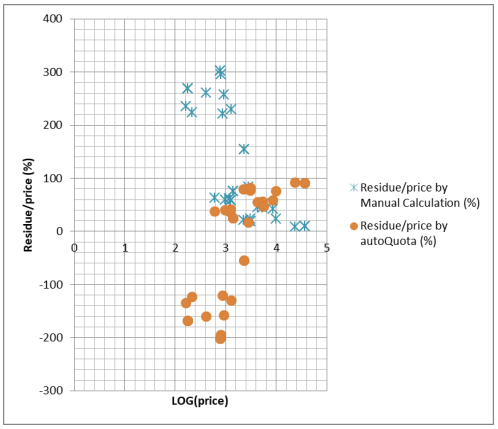

Figure 2 Scatter Plot Diagram

The comparison of the calculated prices and the actual quoted prices are shown in Table 1. There are 6 actual prices, i.e. 162, 606, 970, 1250, 2250 and 3100 US dollars.

Table 1 Calculated Prices vs Actual Prices (USD)

	Manual Calculation	autoQuota	autoQuota (+2, -0.5)	Acutual Data
Upper	-	910	1009	3100
Middle	481	481	481	1110
Lower	-	383	387	162

All the results above lead the conclusions that ***autoQuota*** can estimate more accurate prices than manual calculations.

5. Future Developments

This database can be analyzed by further sophisticated algorithms to construct an automatic plant design system. If ***autoQuota*** is applied to a simulator like ASPEN®, it will enable to build a total plant design system.

References:

[1] Kenneth K. Humphreys: Planning, Estimating, and Control of Chemical Construction Projects, Marcel Dekker Inc., 2001.
[2] R. Keith Mobley: Plant Engineer's Handbook, Butterworth-Heinemann, 2001.
[3] John S. Page: Conceptual Cost Estimating Manual, Gulf Professional Publishing, 1996.