Empirical study on the enterprise financial management and decision based on the decision-tree algorithm

Jingjing Wu

College of Information Engineering, Fuyang Teachers’ College, Fuyang 236000, China

Abstract

Enterprise financial management is the management on assets investment, financing and the distribution of the profit obtained from the working fund in the enterprise operation under the unified goal. Enterprise financial management is an important part of the enterprise management, and under the background of the market economy, the refined management has become the main trend of modern enterprise development, attaching more and more importance on financial management. With the enterprise development, the financial data quantity constantly expands, which has raised higher requirements for the enterprise financial management and decision. Therefore, business intelligence, data warehouse, data mining technology and decision support system have been consecutively used in the enterprise financial management and decision, playing an important role in pushing the traditional enterprise financial management and decision level. In this article, the study is carried out on the enterprise financial management and decision by using the decision algorithm in combination with the data mining and analysis technology based on a large amount of financial data in the enterprise financial management, with the construction of the enterprise financial management and decision system based on the decision tree algorithm, thus providing effective support for daily operation of the enterprise financial management.

Keywords: Decision Tree, Enterprise Decision, Financial Management.

1. STUDY OVERVIEW

1.1 Study background

Financial management is the basic unit of the enterprise organization, and is the core of the enterprise management and the significant guarantee of maintaining normal implementation of daily work of the enterprise, capable of linking the departments, business and even every employee. According to the related survey, one medium-sized enterprise can collect the information of about $2.25 \times 10^{21}$ every year, which is constantly improved at the growth rate of 60%. How to reasonably analyze and use the mass data, and play a counterreaction on the enterprise financial management has been one of the key problems to be widely studied by the enterprise financial management working personnel, experts and scholars. In China, the large-sized enterprises have laid sufficient emphasis on the mass data and informatization construction, the enterprise financial management information system building has been gradually improved, and traditional use of financial statement and empirical adoption of the management and decision mode cannot meet the demand of modern enterprise financial management, and cannot conform to the goal of the enterprise development in the mass data era. Therefore, it is required to make full use of the enterprise financial mass data, analyze the hidden information included in the financial mass data, discover the problems and financial risks in enterprise financial management and decision, raise the targeted solution or evasive measures in a timely manner, which play the important coping and reference role for the enterprise managers, and have the important fundamental effect on normal development of the enterprise financial management and have the far-reaching significance.

1.2 Literature overview

Enterprise financial management shall observe the following rules: first, risk-income weighing rule, that is, the compensation of additional profits shall be made for additional risks; second, currency-time value rule, that is, the same amount of money will be more valuable today than in the future; third, the value is measured by mainly considering the cash, instead of profit; fourth, incremental cash flow rule, that is, only the increment is related to the enterprise financial management; fifth, in the competitive market, there is no project with high
profit; sixth, effective capital market is flexible and the price is reasonable; seventh, agency problem, that is, the management personnel have different interests with that of the owner; eighth, the enterprise tax has direct impact on the business decision; ninth, some of the enterprise financial risks are avoidable and some are unavoidable (Zhang and Hu, 2016). Enterprise financial management mainly has the following four basic goals, first, maximization of the enterprise, that is, obtaining the maximum financial income with the minimum cost fund; second, maximization of the shareholder wealth. The enterprise is funded by the shareholder, and the fundamental reason for the funding of the shareholder is to hope more income from the enterprise. So it is required to attach importance on the maximization of the shareholder wealth in the enterprise development; third, maximization of the enterprise value. The maximization of the enterprise value is to guarantee long-term sustainable development of the enterprise through reasonable financial management and decision of the enterprise, make full use of the time value of the fund and the relations between the risk and reward; fourth, the maximization of the stakeholder value. The enterprise is the geometry of the related stakeholders, including shareholder, operator, employee, creditor, customer, supplier, competitor and the country. Pushing the enterprise development is to attach importance on the development of the stakeholder; finally, maximization of social value. The modern enterprises will unavoidably undertake the social responsibility and have the social value. The maximization of social value can provide the important motive for the sustainable development of the enterprise (Wen and Huang, 2016). BI software is also called business intelligence software, and is tool converting the existing data in the enterprise into knowledge information and helping the enterprise to engage in the management and decision making. In the business intelligence system, the data mainly comes from other business system with relatively high accuracy. At present, many enterprises have started the building of business intelligence system, playing the important role in pushing the use of the enterprise data information resources, performance appraisal and employee analysis report (Cheng, 2016).

2. ANALYSIS AND IMPROVEMENT OF DECISION TREE ALGORITHM

2.1 Decision tree theory

Decision tree algorithm is the important part of the prediction model algorithm, with which the mass data can be classified with certain constraints. Then the information with high value chosen from the classification is provided to the decision maker so as to play the copying and reference role in the decision-making behavior of the decision maker. The decision tree algorithm is the inductive learning algorithm in principle, and the advantage is that it has high accuracy and can display the important decision-making attributes (Ding et al., 2015).

Decision tree algorithm is the key technology in the data mining, and can be classified according to the training data set and the water nature in the cluster label, and according to the corresponding algorithm model, to obtain new information. There are multiple classification techniques, that is, genetic algorithm, Bayesian algorithm, decision tree algorithm and rough set algorithm. The core of decision tree algorithm is to construct the decision tree model (Si, 2015), with the detailed flowchart as shown in Figure 1.

![Decision Tree Algorithm](image)

Figure 1. Decision Tree Algorithm

From the above figure, it may be known that the root node of the decision tree algorithm is also the data source provided in the businesses of the enterprise, the data is preprocessed for the concentrated training and included
as the decision tree. The information data in the decision tree shall be processed according to the classification rules, and the residual parts shall be pruned and then added into the decision tree, while the part with high value will be centralized in the rule. The information data obtained from the rule centralization, that is, the data with relatively high value, can be used in the enterprise financial management and decision, to form the reference opinion or scheme (Wang, 2015). In the business system, the decision tree model is shown in Figure 2.

![Decision Tree Diagram](image)

**Figure 2. Intention Map of Purchasing Goods**

In the decision tree algorithm, good tree structure can highly summarize the supplied data sources and integrate them into the reference scheme of the financial decision making, thus realizing correct identification of the sample category. For some new specimens, compared with traditional neural network algorithm, genetic algorithm or Bayesian algorithm, the decision tree algorithm has high speed, and the results are easier for understanding and have higher accuracy. But the decision tree algorithm has certain defects, which are reflected non-uniform data or data loss, resulting in significant discount of the accuracy rate (Qin, 2014). In addition, after the scanning and sequencing for times, the calculation efficiency will be significantly reduced.

### 2.2 Entropy of information and measurement method

#### 2.2.1 Entropy of information

In physics, entropy mainly refers to the degree of disorder for the measurement in the thermodynamic system. In the information technology, entropy is mainly used to weight the inaccuracy of the information data. In the decision tree algorithm, in combination with Shannon information theory, the test attribute is chosen with the entropy of information as the standard, and the training entity set is classified, to construct the decision tree, and predict how to finish all examples in the classification of the test attribute space. In the construction of the decision tree, it is required to choose an attribute to classify the data, so that the class value of the data on the sub-node is basically same (Wu et al., 2014). If the data on a node is uniformly distributed, the entropy at the node is the largest, vice versa. The definition formula of the entropy of information is shown below:

\[
Entropy(P_1, P_2, \ldots, P_n) = -P_1 \log_2 P_1 - P_2 \log_2 P_2 - \cdots - P_n \log_2 P_n
\]  

In the above formula, the minus signal mainly refers to the logarithm of fraction, if the value of the entropy of information is positive number. Generally, the relation formula is shown below,

\[
\text{Info}\{C_1, C_2, \ldots, C_n\} = \text{entropy}[E_S]
\]

Every training set \(E_S\) is classified, and with attribute A, entropy[\(E_S\)] will be finally reduced. Supposing new expectation information quantity as shown below:

\[
\text{new}_-\text{entropy}(E_S, A) = \sum_{\text{value}(A)} \left| \frac{|E_{S|A}|}{|E_S|} \right| \text{entropy}[E_S]
\]

Where, entropy[\(E_S\)] is the reduced quantity, that is, Gain[\(E_S, A\)] wined from the information of the training set \(E_S\) for attribute A, and if the information in certain quantity can win more attributes, and it is more beneficial for the training set, with the formula shown below:
According to the definition of the information of entropy, it may be known that the information of entropy, in principle, is the average information quantity after the elimination of the redundant information, representing the disorder degree of the information, with the calculation formula shown below:

\[
I(S_1, S_2, \ldots, S_m) = -\sum_{i=1}^{m} P_i \log_2(P_i) \tag{5}
\]

2.2.2 Information gain

The measurement of the information quantity of the attributes in the classification is subject to the standard of the size of the information gain, the decision tree construction of this node will be decided according to the calculated information gain value, and the bigger information gain indicates bigger contribution of classification, otherwise with smaller contribution of classification. If the study goal is of the unified attribute, it is only required to calculate the entropy of information (Liu et al, 2013). The information gain is used to choose the type attribute of phenol, and the information gain obtained after the branch A is shown below,

\[
Gain(A) = I(s_1, s_2, \ldots, s_m) - E(A) \tag{6}
\]

Where, \(I(s_1, s_2, \ldots, s_m)\) is the entropy of information expected from the given sample \(S\), and the calculation formula of \(E(A)\) is shown below:

\[
E(A) = -\sum_{j=1}^{v} \frac{s_{ij} + \ldots + s_{mj}}{s} I(s_{ij} + \ldots + s_{mj}) \tag{7}
\]

Supposing that in attribute A, there are different values \(\{a_1, a_2, \ldots, a_j\}\), the attribute A classifies the sample \(S\) into \((s_1, s_2, \ldots, s_j)\), where, \(s_j\) is the number of the \(C_i\) samples in \(s_j\).

2.3 Construction of the decision tree

According to the above contents, it is possible to construct a complete decision tree model of enterprise financial management and decision. First, supposing that the total sample training set is \(E\). If there is no data in the training set \(E\), for the returned \(F\) single node, all results are failure values. While if the data in the same attribute constitutes the training set \(E\), it will return the single node with Category \(C\) mark. If there is one set \(A\) without category, but containing continuous attribute and the set \(A\) does not have the data information, then the returned attribute value will be considered as the attribute value with the most samples in the sample training set \(E\). After checking all information data in the set \(A\), if the set \(A\) has one element \(j\) with continuous attribute, it is required to let the maximum value in \(A_j\) as \(B_1\), and maximum value as \(B_2\) (Xie, 2013). Implement the for cycle, set the initial value of the data \(j\) as 2, and after one execution, add the initial value of data \(j\) by 1, till \(j=n-2\). The formula is shown below:

\[
B_n = B_i + j * (B_i - B_n) / n \tag{8}
\]

Supposing that \(B_i\) is equal to the maximum information gain attribute value of element \(A_i\), and the maximum attribute value of information in the set \(A\) is \(X\). Build the related mapping relation set \(\{X_j\}_{j=1, 2, 3, \ldots, m}\) of the set \(X\), and construct the node according to the values in the above set, where the nodes are expressed as \(x_1, x_2, x_3, x_4, \ldots, x_m\). Finally, repeat the above process to construct other subtrees (Wang et al, 2011).

3. ENTERPRISE FINANCIAL MANAGEMENT AND DECISION MODEL BASED ON THE DECISION TREE ALGORITHM
Enterprise financial management and decision system theory is a multidisciplinary and comprehensive theory, and is an important part integrated from the enterprise risk management, project investment management, information technology, and dynamic digital modeling of multidisciplinary knowledge. The emphasis of the enterprise financial management and decision system is the enterprise financial index fluctuation with high trust rules. When the index fluctuation exceeds certain range, it proves that there is certain problem with the enterprise financial management and decision, and it is required to find and resolve the problems in a timely manner, thus help the enterprise to evade or prevent the enterprise financial risks (Xu, 2011).

First, build the enterprise financial management and decision system theory, including, the definition of enterprise financial crisis, construction of overall planning roadmap, development trend of enterprise lifecycle theory, financial crisis enterprise financial crisis warning procedure.

Second, build the index system of the enterprise financial management and decision system. The index system of the enterprise financial management and decision system not only includes the financial risk analysis indexes, but also includes the evaluation indexes of the current enterprise financial management and decision level. At present, since the study on the financial management and decision system has been complete, it will not be introduced and deeply studied in this study, and the emphasis will be laid on the study of the related investment project indexes, such as cash flow and financing amount. The data is restructured to obtain the data set with proper data mining, find out the law among the financial indexes, and predict the development trend of enterprise financial management and decision system (He, 2010).

3.1 Enterprise financial management and decision model

Enterprise financial management and decision model in the decision tree algorithm generally consists of the nodes and edges. The nodes are mainly classified into three types, first, root node, that is, there is no entering edge, but including zero or multiple leaving edges, that is, no data input, but with or without the data output. Second, the internal node, that is, there is one entering edge, and there are more than two leaving edges, that is, the data comes from the root data and is transmitted through the internal node into other part. Third, leaf node, also called ending node, that is, one entering edge and no leaving edge. The entering edge means that the internal node is transmitted through the information, and there is no leaving edge because the leaf node is itself the output data, available for the management and decision of the decision maker (Zhang et al, 2010).

3.2 Building of the database three-layer model

In the enterprise financial management and decision database, the three-layer model is shown in Figure 3:

At first, in conceptual model, the acquired information shall be analyzed with respect to the unit, time and budget index for the enterprise financial management and decision database, and the information data will be analyzed and studied with the decision tree algorithm. From the budget analysis of the departments and businesses, it is possible to analyze the management mechanism, personnel configuration and fund inventory of the enterprise from another aspect. The concept is shown below,

First, the boundary of the enterprise financial budget analysis model is the boundary of the database;

Second, in the enterprise financial budget and decision, the required data includes basic unit information data, business budget data, fund use plan data, project income data, initial balance data, current balance data and so on;
Third, the main performance indexes mainly include the income and expenditure information in the financial budget;

Fourth, the definition dimensions mainly include the unit information, time, budget income and expenditure;

Fifth, the category definitions of the above dimensions mainly include budget code, time, amount and name and so on.

Secondly, with regard to the logic model, the logic model mainly undertakes the extraction and analysis of the data sheet, relation formula and data information. In the storage model in the logic model level, mainly in the star shaped, fact constellation and snowflake models (Xue, 2015).

Finally, in the physical model, the physical model mainly undertakes the function of storage and organization of the information data in the data warehouse and is the key part of the whole data warehouse. So constructing the physical model is the process of constructing a physical database to certain extent.

REFERENCES


He Y. (2010). Strengthening the combination of enterprise financial management and asset management is the key to improve economic efficiency, finance and economics circles (Academic Edition), (08), 88-89.


Qin G. (2014). Ways and methods to promote enterprise financial management under the new situation -- Taking the financial management of construction enterprises as an example, China business, (26), 62-63.


Wen J.W., Huang J.B. (2016). Influencing factors and improvement strategies of financial management system for farmers' Professional Cooperatives -- Taking Jiangsu Province as an example, agricultural modernization research, 37(02), 304-309.


Xu G. (2011). Strengthen financial management, improve the level of internal supervision. Investigation and Reflection on the current financial management mode of state-owned enterprises, reform and opening up, (06), 86-87.

Xue G. (2015). Learn from the experience of financial management in American first-class universities to improve the financing ability and management level of universities in China, Journal of the National Institute of education administration, (01), 76-81.
