Research Article

Evaluation Model of Enterprise Lean Management Effect Based on Data Mining

Rui Chang
Hebei Petroleum University of Technology, Hebei, Chengde 067000, China

Correspondence should be addressed to Rui Chang; changrui@cdpc.edu.cn

Received 9 November 2021; Revised 17 December 2021; Accepted 12 January 2022; Published 23 February 2022

Academic Editor: Gengxin Sun

Copyright © 2022 Rui Chang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article discusses the task allocation decision of single management ability and multiple management ability in view of the operation stage of the specific project of enterprise lean management service supply chain. First of all, the basic process of the management service integrator’s task allocation to the functional management service provider is given. Based on factors such as business satisfaction and other factors, we construct a management task allocation model based on service income price discounts in the case of fuzzy demand. On this basis, we further dig out the external and internal related factors of supply risk, use enterprise research and data mining methods to verify theoretical hypotheses and identify related factors of supply risk, analyze the impact of the characteristics of cooperation with different types of suppliers in data mining to determine the key factors of supply risk for different types of supply and demand cooperation, and determine the focus of supply risk control for different types of suppliers for data mining. In the process of simulation experiment, the method of maximum satisfaction of fuzzy multiobjective programming is adopted, and the influence of factors such as the preference of decision-makers and price discounts on the result of task allocation is analyzed in combination with specific examples. Finally, on the basis of fully considering the characteristics of management services, we build a comprehensive evaluation index system for management service providers in terms of service capabilities, service quality, informatization capabilities, and cooperation risks and comprehensively consider the uncertain factors in the decision-making process. The experimental results show that the proposed interval data mining solution method based on the degree of possibility can realize the ranking of management service providers by combining the subjective wishes of decision-makers with the objective evaluation of data mining.

1. Introduction

With the continuous development of the economy, the Internet and e-commerce have been widely used, which has brought about profound changes in the way enterprises operate [1]. At the same time, the supply relationship between enterprises and it has become more and more complex. The formation of global information networks and global markets and the acceleration of technological changes have prompted enterprises to continuously seek ways to reduce costs, increase benefits, and improve market competitiveness. These problems have caused lean managers and theoretical workers in enterprises to redefine and think about product supply relationships and models to adapt to changes in the economic environment. Material enterprise lean is an important part of enterprise resources, and enterprise lean management is also an important part of supply chain management. Therefore, it is closely watched by senior managers and theoretical researchers of today’s enterprises. How to apply information technology to enterprise lean management and fully integrate with enterprise information management system is one of the major tasks facing enterprises [2–6].

Enterprise lean management is one of the research hotspots in the field of modern enterprise management, and it is an effective means for enterprise resource decision-making and improving core competitiveness. Based on the analysis of the classic enterprise lean management theory and the enterprise lean management theory under the supply chain environment, the paper puts forward a generalized enterprise lean management model and proposes an indicator system for a comprehensive supplier evaluation model.
The model integrates analytic hierarchy process, group decision-making method, and association rule method into the model, which not only emphasizes the importance of expert knowledge and domain knowledge, but also takes into account the frequency of purchase transactions and the analysis of historical transaction data, effectively avoiding subjective for the risk of sexual evaluation, and discovers the inherent laws and potential knowledge of the supplier’s supply. A mutually independent and interrelated system has been formed, so as to obtain more scientific strategic supplier selection results. Compared with the decision-making results of strategic supplier selection based on the analytic hierarchy process and the association rule method, the model is more scientific and effective [8–10].

The model in this paper embodies the multifield collaborative cooperation relationship, which can scientifically and effectively respond to and predict the needs of end customers, and can grasp accurate customer information in time, providing a systematic basis for determining the final strategic supplier. At the same time, the system structure and related technologies of enterprise lean management are studied, and the enterprise lean management system framework based on the second generalized enterprise lean management mode is constructed. Third, based on the research results of supplier classification and supply risk-related factor identification, the dummy variables of load supplier category characteristics and supply risk influencing factors are integrated as an indicator system, and data such as support vector machine (SVM) and decision tree (DT) are used. Furthermore, the combination coefficients are solved by the principle of minimizing the error square sum, and it is constructed based on the data mining and DT linear combination evaluation model of supply risk. The evaluation results show that the system errors of the above models are relatively small, but the linear combination model shows the highest evaluation accuracy and the best stability, and the evaluation effect is better than independent data mining and DT models.

2. Related Work

The current theoretical research on manufacturing supply risk issues is mainly based on the theory and methods of risk management, placing supply risk in the research system of supply chain risk, and proceeding from the perspectives and processes of risk identification, risk assessment, and risk control. Research on supply risk identification: one type is to decompose supply risk through perceptual knowledge based on the law of supply risk itself and personal experience, such as Probability Tree Analysis (PTA) and Fault Tree Analysis (FTA) methods. The other is to rely on quantitative methods such as objective statistics to find the law of risk, such as questionnaire survey method and environmental analysis method [11–13].

Alfian [14] established the wholesale price contract and revenue sharing contract in the management service supply chain from the countermessures, analyzed the overall optimal service ordering volume of the supply chain, and established management subcontractors and management service providers on this basis of Nash equilibrium model. Albahri [15] found that under the condition of uncertain demand, a variety of management ability task allocation models are established that aim at the smallest integrator cost, the largest satisfaction of functional service providers, the smallest penalty intensity, and the best possible match between different service capabilities and discussed the influence of uncertainty and relational cost coefficient on the result of task assignment. Baek [16] studied the quantity coordination problem of two-level management service supply chain capacity cooperation based on Stackelberg’s master-slave decision-making and constructed a cost model of management service integrators and functional management services with/without capacity matching constraints under uncertainty. The profit model of the provider and the coordination result under the master-slave decision is given. Li [17] established a capacity coordination model for supply chain option contracts based on Stackelberg’s master-slave game, studied the ordering and investment decision-making problems of the management capabilities of integrators and subcontractors, and obtained the optimal strategy of both parties through the solution of the planning model and verified the effectiveness of the coordination mechanism.

By analyzing the evolution of supplier relationship management in the petrochemical industry, Jalali [18] takes a refinery as an example; based on the Kraljic model, petrochemical companies’ procurement materials are classified as suppliers based on the criteria of strategic importance and supply risk, and the petrochemical industry is determined by three representative types of enterprise supplier relationships: strategic cooperation, reverse marketing, and reverse auction. According to the characteristics of the petrochemical industry, scholars construct a corresponding supplier evaluation index system and apply the Analytic Hierarchy Process (AHP) to classify and manage suppliers from their corporate capabilities and importance. In view of the large transaction data between buyers and suppliers, the researcher proposed a k-means subspace clustering algorithm to classify suppliers, which solved the problem of data mining and analysis of high-dimensional sparse samples and verify the effectiveness of the method through examples [19]. Researchers use the AHP method to classify the domestic and foreign suppliers of supporting equipment for shipbuilding enterprises and apply the fuzzy clustering method to the classification research of manufacturing industry suppliers [20].

3. Construction of an Evaluation Model for the Effect of Enterprise Lean Management Based on Data Mining

3.1. Data Mining Hierarchy. Data mining, also known as Knowledge Discovery in the Database (KDD), refers to the nontrivial process of extracting implicit, unknown, and potentially useful information from a database or data warehouse. Data mining technology can extract from a large amount of incomplete, noisy, ambiguous, and random data
the hidden, unknown, but potentially useful information and knowledge. Figure 1 is the hierarchy of data mining.

Data mining technology is the product of the combination of database technology and artificial intelligence technology. It solves the problem of large amounts of data but lack of useful information in the development of information technology and completes the conversion from business data to decision-making information. The knowledge discovery (KDD) process in the database can be roughly divided into several steps: data selection, data preprocessing, data transformation, data mining, and result interpretation and evaluation.

\[
W_{st}^{m+n} = \begin{bmatrix}
W_{11}^{m+n} & S_{12}^{m+n} & r_{11}^{m+n} \\
W_{21}^{m+n} & S_{22}^{m+n} & r_{22}^{m+n} \\
W_{31}^{m+n} & S_{32}^{m+n} & r_{32}^{m+n}
\end{bmatrix}
\]  \hspace{1cm} (1)

Regarding the comparison and sorting of interval numbers, there have been many discussions in the literature, which mainly include two categories: one is to give the absolute size relationship (greater than, less than, or equal to) of two interval numbers, which is called deterministic ranking method.

\[
f(m) = \frac{P \times C(m, t)}{\sum_{i=1}^{n} c(i, t)} - \frac{n \times f(m, t)}{\sum_{i=1}^{n} f(i, t)} = 0.
\]  \hspace{1cm} (2)

This type of method is too rough, loses a lot of information, and is difficult to sort multiple interval numbers, so its applicability is not high; the other type is based on the ambiguity of people’s understanding of things, giving one interval number greater than another interval number. The degree is called a degree-based sorting method.

\[
g(m, t) = n \times \frac{h(m, t)}{k(i, t)} = \sum_{i=1}^{n} \frac{h(i, t)}{k(i, t)}
\]  \hspace{1cm} (3)

\[
\max \sum_{i=1}^{N} a_i = \frac{1}{m} \sum_{i=1}^{N} \sum_{j=1}^{N} a_j y_i a_j y_j \times k(x_i^2, x_j^2) = 0.
\]

The applicability of the NSG possibility degree method is relatively strong. For any two interval numbers, the degree that one is greater than or equal to the other can be given. If \( p(a \geq b) > 0.5 \), it is considered that the possibility of mouth being greater than or equal to \( b \) is high; when \( p(a > b) < 0.5 \), it is considered that the possibility of mouth being greater than or equal to \( b \) is small; when \( p(a \geq b) = 0.5 \), the probability that the mouth is greater than or equal to \( b \) is 0.5.

Calculating the new single-objective linear programming model, the optimal solution obtained is the fuzzy optimal solution (satisfactory solution) of the original fuzzy multiobjective programming model, and then this optimal solution is brought into each objective function of the original model, and each objective is the optimal objective function value under the condition that the original problem obtains the optimal solution.

\[
g(x) = \begin{cases}
f(x, d) - f(x, m) & x > d \\
\frac{f(x)}{f(x) - f(x - 1)} & x \leq d
\end{cases}
\]

\[
\sigma(x, t) \times w(\xi) - \min \{\delta w_{FR} + w_{F1} + f(\xi, t)\} = 0.
\]  \hspace{1cm} (4)

Data mining can be described as the process of extracting patterns from data. It repeatedly uses a variety of Did algorithms to determine patterns or suitable models from observation data. (1) Model: includes model functions (such as classification or clustering, etc.) and basis functions. The model is formed after the parameters are determined from the data. (2) Priority criterion: due to the inaccuracy of the data, will lead to the diversity and incompleteness of the model. Therefore, it is necessary to have a deterministic model or parameter priority mechanism, which is generally a function of the degree of conformity of a certain form of model to the problem. (3) Search algorithm: including various search algorithms for models, parameters, or priority criteria.

3.2 Enterprise Lean Management Process. The enterprise lean management process is divided into two stages: the first is the working stage, at which time the neural network keeps the connection weight of each node unchanged. The calculation starts from the input layer, and the output is calculated node by node until the calculation of each node in the output layer is completed, which is also called the forward propagation process.

At this time, the output of each node of the neural network remains unchanged. Starting from the output layer, learning starts from the output layer and calculates the modification amount of the weight of each connection node by node to modify the weight of each connection until the input layer, also known as the back propagation process. If the output of the network output layer differs greatly from the expected output value, the backpropagation process will start, and the weight of each connection between the network nodes will be modified according to the signal error between the network output and the expected output, so as to reduce the network for the error between the output signal and the expected output. The BP network is through such continuous calculation process of forward propagation and backward propagation, which finally makes the output value of the network output layer and the expected value converge.

The nodes of the input layer usually just perform some simple normalization processing on the input sample data. The real network calculation processing starts from the second layer to the output layer. Figure 2 shows the data output distribution of the enterprise lean management network.
Risk assessment is an important factor in the selection of partners for integrated enterprises. It mainly includes the assessment of corporate reputation, financial status, and collaborative development capabilities. E-commerce transactions under the virtual Internet environment are fundamentally different from direct transactions, and reputation is essential to ensure the quality of the management services they provide. Reputation indicators are mainly measured by indicators such as historical transaction records and credit rating certification. The historical delivery record is measured by the ratio of the number of bad service records of the service provider to the total number of services; the credit rating certification mainly refers to the country’s classification of the credit rating of the management enterprise.

The financial status is mainly measured by the debt ratio and profit margin of the company’s assets. Aiming at the problem to be studied in this chapter the assessment of the supply risk of nuclear power equipment manufacturers, it is necessary to estimate the dependence between the variables based on the known samples and establish the regression between the dependent variable and the independent variable on the basis of the obtained observation data. Functions, research methods, and implementation techniques used to describe the regularity between variables so as to predict, judge, and evaluate unknown or unmeasured data.

3.3. Management Effect Evaluation Indicators. The three links of data selection, data preprocessing, and data transformation belong to the steps of the management effect evaluation index data preparation stage. Data selection is to determine the operation object, that is, to extract the corresponding data from the original data according to the user’s needs. Data preprocessing is used to eliminate noise, derive and calculate missing data, eliminate duplicate records, and complete data type conversion. Data transformation is used to extract the really useful data features from the original features to reduce the workload of data mining, that is, to reduce the data dimension. The first step in the data mining stage is to determine the mining tasks, such as

Figure 1: Data mining combination of database technology hierarchy.

Figure 2: Data output distribution of enterprise lean management network.

Network data for enterprise lean management

- A circle with a radius of 10
- A circle with a radius of 5
- 9.9 from the origin

3.3. Management Effect Evaluation Indicators. The three links of data selection, data preprocessing, and data transformation belong to the steps of the management effect evaluation index data preparation stage. Data selection is to determine the operation object, that is, to extract the corresponding data from the original data according to the user’s needs. Data preprocessing is used to eliminate noise, derive and calculate missing data, eliminate duplicate records, and complete data type conversion. Data transformation is used to extract the really useful data features from the original features to reduce the workload of data mining, that is, to reduce the data dimension. The first step in the data mining stage is to determine the mining tasks, such as
data summarization, classification, clustering, association rule discovery or sequence pattern discovery, etc.

Two factors should be considered when choosing a mining algorithm: data of different nature should use algorithms related to its characteristics and user requirements for discovery results (for example, some users want easy-to-understand and descriptive knowledge, and some users want accurate predictive knowledge with as high a degree as possible). If the discovered knowledge cannot meet the user’s requirements, the entire discovery process needs to be returned to the discovery stage before reselecting data, adopting new data transformation methods, setting new data mining parameter values, or even changing a mining algorithm. Figure 3 is the hierarchical division of management effect evaluation indicators.

The basic idea of the data mining method is to treat each evaluated unit as a DMU, and then constitute the evaluated group by many DMUs, and use the weight of each input and output index of the DMU as a variable, through a comprehensive analysis of the input and output ratios. We carry out evaluation calculations to determine the effective production frontier, so as to determine whether each DMU is effective in data mining. At the same time, the projection method can also be used to point out the reasons for the effectiveness of nondata mining or weak data mining and the direction and degree to be improved.

At present, the commonly used evaluation method is to analyze, synthesize, correlate, and aggregate the security data existing in the network infrastructure by using related technologies and procedures. Because data mining methods do not need to estimate parameters in advance, they can avoid subjective factors, simplify operations, and reduce errors. Given a data set with N tuples or records, construct K groups (K < N), and each group represents a cluster. Moreover, these K groups meet the following conditions: (1) Each group contains at least one data record. (2) Each data record belongs to only one group (it can be relaxed in some fuzzy clustering algorithms). (3) For a given K, an initial grouping method is given first, and then the grouping is changed through repeated iterations, so that the grouping scheme after each improvement is better than the previous one.

3.4. Model Evaluation Index Weight Update. The model evaluation index time series mode is to describe and model frequently occurring laws or trends based on time or other series. The sequence mode combines the correlation mode and the time series mode, focusing on the correlation between the data in the time dimension. The choice of the following three parameters has a great impact on the results of sequence pattern mining: (1) The duration t of the sequence, which is the effective time of a certain time sequence or a time period selected by the user; (2) the time folding window w (w ≤ t), events that occur within a certain period of time, can be regarded as occurring simultaneously; and (3) the description is mainly used to discover and describe the patterns that people can explain in the data; the descriptive pattern is to make a standardized description of the facts in the current data.

Prediction includes the use of variables or domains in the database to predict the unknown or future values of other variables of interest; the predictive model uses time as the key parameter, and for time series data, it is based on its historical and current values to predict its future value. For specific data mining applications, the relative importance of prediction and description may differ. Figure 4 shows the weight distribution of model evaluation indicators.

The system construction phase is a management process in which the supply and demand parties evaluate each other and determine the initial cooperation willingness. Under the dual effect of both parties’ cooperation capabilities and willingness to cooperate, the high and low attributes of supplier supply risk r is BS, they obey the binomial distribution, where B represents the supplier with high supply risk, and S represents the supplier with low supply risk. Before the completion of the supply task, the purchaser cannot fully understand the supplier’s supply (cooperation) performance, and the high risk probability of the supplier is determined to be p = p1; the low risk probability is 1− p = p2. Further, we analyze the cooperation situation in the system construction stage: in the system construction stage, because nuclear power equipment manufacturing companies are difficult to understand the supplier’s true supply performance in detail beforehand, they tend to evaluate performance and capabilities through the materials submitted by the supplier.

Therefore, suppliers expecting to cooperate with nuclear power equipment manufacturing enterprises may hide information that the purchaser does not know and is not conducive to cooperation in order to obtain the cooperation contract. This is usually manifested as follows: suppliers with relatively high supply risks whitewash cooperation willingness and exaggerate cooperation capabilities.

4. Application and Analysis of the Evaluation Model of Enterprise Lean Management Effect Based on Data Mining

4.1. Data Mining Feature Extraction Processing. Supply risk assessment based on data mining is to establish SVR models and determine regression equations with the help of Lib data mining technology. In order to ensure the balance of data information characteristics and achieve cross-comparison and inspection, this paper will normalize the 60 valid questionnaires (60 samples) of the survey data, and according to the results of the supplier classification, according to the data mining four types of samples in the total proportion of the sample: 4 supplier-oriented samples, 19 strategic samples, 8 demand-side-oriented samples, and 9 loose samples were selected.

To infer the exact threat level, type of attack, and related or interdependent risks, a total of 40 samples were used as the first set of data, and the remaining 20 samples are used as the second set of data. Based on the two sets of data, two sets of SVR evaluation models are established: (1) The first set of data is used as the training sample to evaluate the supply risk of the second set of samples. (2) The second set of data is used as the training sample to evaluate the first set of data. In
the field of regression research, data mining and the selection of its parameters and kernel functions and its parameters have not yet formed a unified guidance model. The selection of optimal data mining algorithms, kernel function types, and their parameters can only be compared or used by a large number of experiments. The interactive inspection function is provided by the software package for optimization. Lib data mining is based on a large-scale search and attempt, trying to find the optimal parameters through the interactive inspection function of the system to achieve the minimum MSE of the system. Figure 5 is the data mining feature extraction distribution.

According to the influencing factors of the integration degree of the supply chain, the integration intensity of the evaluation index of the integration degree in the SNA is used for reference to reflect the closeness of the core enterprise’s connection with suppliers and customers. According to the characteristics of the whole supply chain, the use of “supply chain enterprise relationship” reflects the closeness of the connections between enterprises in the whole supply chain. The main factor affecting the integration of the entire supply chain is the closeness of the relationship between the core enterprise and other enterprises. Although there are contacts between enterprises, if the frequency of contact is not high, the exchange of information is incomplete, or the company’s positioning of partnerships is not high; the integration of such a supply chain is not high. The closeness of the relationship is related to many factors, such as the degree of trust, the depth of communication between the two parties, the frequency of contact, and the degree of interdependence of resources. Close cooperation, mutual benefit, and more interactive relationship is a strong relationship; the opposite
is a weak relationship. If the relationship between the core enterprise and other node enterprises is higher, the integration degree of the supply chain will be higher.

4.2. Evaluation and Simulation of the Effect of Enterprise Lean Management. DT is a relatively common method of regression analysis for the evaluation of lean management effects in enterprises. Its processing process is to divide the data into several branches according to a tree structure. Each branch contains the category attributes of the data tuple and further extracts useful information from each branch. Information forms rules, establishes the relationship between all indicators of the risk assessment system and the degree of supply risk, and evaluates the supply risk of suppliers who are in supply and demand cooperation (or have not yet cooperated, but can obtain indicator observations). This indicator describes the ratio of the variability of the observation value to the variability of the evaluation value, and it reflects the consistency of the variation between the observation value and the evaluation value.

The purpose of network security situation assessment is to obtain the network security situation at a macro level. When $CD = 1$, it means that the observed value and the estimated value have the same variation; when $0 < CD$, it means that the model cannot fully describe individual extreme values, which is suitable for data with low variability. Through the integrated arc fan diagram, you can intuitively understand the overall status of the entire supply chain; we find out the main reasons leading to low integration, clarify the direction to improve integration, and strive to improve, through scientific selection and evaluation of core enterprises and upstream supply chain and downstream customers to achieve effective synergy and overall value growth of various enterprises, forming an optimal full-process supply chain. Table 1 is the evaluation of the effect of enterprise lean management.

Because of the large amount of data, the establishment of the decision number is usually realized through computer programming. Due to the limitation of the length of this article, and the number and characteristic information of the historical suppliers of the company, the research object of this article, only the number of historical suppliers (10) is collected, and the characteristics of each are used to explain the whole algorithm’s truth and ideas. Figure 6 is the MSE value distribution of the data mining evaluation model.

Comparing the MSE values of the 4 groups of evaluation models, the MSE of the linear combination model is 0.0076, which is the minimum value of all models. At the same time, in addition to the MSE value of 0.034 for the second group of
4.3. Example Application and Analysis. This paper presents the data tuple taken from the supplier’s special meal information database as the training set. The class label attribute class has three different values (representing elimination type, growth type, and strategic type, respectively, so there are three different types). Suppose class L1 is assigned to manual, class L2 corresponds to CZ, and class L3 corresponds to ZL, class L1 has 3 samples, class L2 has 5 samples, and class L3 has 2 samples. The fixed asset turnover rate is the proportional relationship between sales revenue and fixed assets in a certain period of time. The operating efficiency of fixed assets, and to a certain extent, reflects the production flexibility of the enterprise. In this sense, the value of this indicator is not that the larger the better, but the moderate; that is to say, there must be a certain degree of flexibility.

This needs to capture a large amount of network security information data and then analyze the correlation between this information. The calculation formula is net sales revenue/average net value of fixed assets = (sales revenue-sales discount-sales discount-sales return)/((opening net fixed assets ten ending net fixed assets)/2). In a broad sense, data analysis can be divided into Verification-Driven Data Analysis and Discovery-Driven Data Mining. Among them, multidimensional query and OLAP can easily observe the actual situation of the system in order to determine whether a certain hypothesis is true, so it belongs to the category of verification. Data mining is to discover knowledge from the unknown in a large amount of data, so it belongs to the category of mining analysis. Figure 7 is the distribution of data tuples in the data mining information database.

The first-level model dictionary mainly describes the main body of the model. The first-level model dictionary describes all models, including detailed description of each model’s number, name, type, main function, model remarks, etc., and clearly defines the data type of each content. The second-level model dictionary describes the specific conditions of each model. It uses the model number as a key word to describe in detail the model number, model parameters, model notes, and data types involved in each model. For multiojective optimization problems, since there is more than one objective function, and there are certain conflicts and incommensurability, it is often impossible to achieve the optimal value of these objective functions at the same time.

Applying the idea of fuzzy mathematics, the degree of satisfaction of the decision-maker for each goal is expressed by the value of the membership function and transformed into regarding the new single-objective optimization problem of membership degree. This indicator indicates the degree to which the sum of all evaluation values deviates from the sum of all observation values. When CRM > 0, it means that the overall evaluation value is less than the observation value; when CRM = 0, it means that the sum of the two is equal (however, the indicators may not correspond to each other equal). Figure 8 shows the optimal value distribution of the data mining objective function.
The knowledge base is a special data mining database used for knowledge management to facilitate the collection, sorting, and extraction of knowledge in various fields. The knowledge in the knowledge base is mainly divided into two categories: fact knowledge and rule knowledge, which are stored in the fact base and the rule base, respectively. Management integrators select service providers for multiple management processes.

There are $p$ candidate service providers, some of which can provide multiple management services; management service integrators can objectively and truly master each. The information of a candidate service provider: the management service integrator is in a strong position in task allocation, and the service provider is in a weak position; each service provider has service capacity limitations, and service shortages are not allowed within the capacity; in a certain period of time
within a certain tolerance range, the customer’s demand for management services is vague. Generally speaking, business intelligence systems are built on data warehouse and data mining technologies. By collecting, sorting and analyzing various data inside and outside the enterprise, the enterprise can deepen its understanding of customers and the market, and use certain tools to make reasonable evaluations and predictions on business operating conditions, customer needs, and market dynamics, and provide scientific decision-making basis for business management.

5. Conclusion

This paper discusses the application of data mining methods in the management of management service providers and uses fuzzy clustering analysis in the dynamic classification of management service providers. First, the topological space generated by fuzzy concepts under the APS theory is given. Based on the neighborhood generated by related concepts, a new measurement method is defined to determine the distance between samples; the new distance is applied to replace the Euclidean distance. In the traditional fuzzy clustering method, the research in this paper is a new application of data mining technology. The strategic supplier management evaluation model enables enterprises to scientifically and reasonably evaluate strategic suppliers. The evaluation provides a reliable basis for decision-making for business managers. Finally, the evaluation model of strategic suppliers is verified by an example based on a specific enterprise. The model uses rough sets as the main method of data analysis and single rule generation and integrates the prior knowledge and information gain of domain experts as the final weight of conditional attributes. It not only emphasizes the importance of experts’ prior knowledge and domain knowledge, but also takes into account objective data analysis effectively avoids the risk of subjective evaluation, and realizes the comprehensive integration of qualitative analysis to quantitative analysis: the analysis and processing of a single rule through multifeature combination mode mining, giving full play to the advantages of data and making it a decision-making rule. Through the test of international public data, the clustering results are compared, which proves that the improved clustering method has good accuracy; the improved fuzzy clustering analysis method is used in the evaluation and classification of management service providers, through examples discuss the classification of suppliers under simple and complex concepts, and provide a scientific reference for actual decision-making.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The study was supported by Hebei Petroleum University of Technology.

References

[1] N. Kunnathuvallappil Hariharan, "Applications OF data mining IN FINANCE,” International Journal of Innovations in Engineering Research and Technology, vol. 5, no. 2, pp. 72–77, 2018.
[2] V. Suma and S. M. Hills, "Data mining based prediction of demand in Indian market for refurbished electronics,” Journal of Soft Computing Paradigm (JSCP), vol. 2, no. 02, pp.101–110, 2020.
[3] F. Wang, L. Ding, H. Yu, and Y. Zhao, "Big data analytics on enterprise credit risk evaluation of e-Business platform,” Information Systems and E-Business Management, vol. 18, no. 3, pp. 311–350, 2020.
[4] S. Q. Salih, A. Sharafati, K. Khosravi et al., “River suspended sediment load prediction based on river discharge information: application of newly developed data mining models,” Hydrological Sciences Journal, vol. 65, no. 4, pp. 624–637, 2020.
[5] J. Dou, A. P. Yunus, D. Tien Bui et al., “Evaluating GIS-based multiple statistical models and data mining for earthquake and rainfall-induced landslide susceptibility using the LiDAR DEM,” Remote Sensing, vol. 11, no. 6, p. 638, 2019.
[6] S. Bharara, S. Sabitha, and A. Bansal, “Application of learning analytics using clustering data Mining for Students’ disposition analysis,” Education and Information Technologies, vol. 23, no. 2, pp. 957–984, 2018.
[7] R. P. Schumaker, "Machine learning the harness track: a temporal investigation of race history on prediction,” Journal of International Technology and Information Management, vol. 27, no. 2, pp. 2–24, 2018.
[8] Y. Liu, H. Cai, G. Li, and H. Zhao, "Research on application of data mining in hospital management,” in Proceedings of the Robots & Intelligent System, pp. 374–376, Changsha, China, May 2018.
[9] N. A. Yassein, R. G. M. Helali, and S. B. Mohomad, “Predicting student academic performance in KSA using data mining techniques,” Journal of Information Technology & Software Engineering, vol. 7, no. 5, 2017.
[10] S. Hlee, J. Lee, S.-B. Yang, and C. Koo, "The moderating effect of restaurant type on hedonic versus utilitarian review evaluations,” International Journal of Hospitality Management, vol. 77, pp. 195–206, 2019.
[11] J. Pei, K. Zhong, J. Li, J. Xu, and X. Wang, "ECNN: evaluating a cluster-neural network model for city innovation capability,” Neural Computing & Applications, vol. 9, pp. 10–13, 2021.
[12] S. Huber, H. Wiemer, D. Schneider, and S. Ihlenfeldt, "DMME: data mining methodology for engineering applications - a holistic extension to the CRISP-DM model,” Procedia Cirs, vol. 79, pp. 403–408, 2019.
[13] T. Saeed, "Knowledge management, innovation, technology and direct marketing as antecedents of data mining: the mediating role of direct marketing,” Invertis Journal of Science & Technology, vol. 13, no. 2, pp. 88–99, 2020.
[14] G. Alfian, J. Rhee, H. Ahn et al., "Integration of RFID, wireless sensor networks, and data mining in an e-pedigree food traceability system,” Journal of Food Engineering, vol. 212, pp. 65–75, 2017.
[15] A. S. Albahri, R. A. Hamid, J. K. Alwan et al., “Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (COVID-19): a systematic review,” Journal of Medical Systems, vol. 44, pp. 10-11, 2020.

[16] H. Baek, M. Cho, S. Kim, H. Hwang, M. Song, and S. Yoo, “Analysis of length of hospital stay using electronic health records: a statistical and data mining approach,” PLoS One, vol. 13, no. 4, Article ID e0195901, 2018.

[17] Z. Li, Y. Wang, and K.-S. Wang, “Intelligent predictive maintenance for fault diagnosis and prognosis in machine centers: industry 4.0 scenario,” Advances in Manufacturing, vol. 5, no. 4, pp. 377–387, 2017.

[18] S. M. J. Jalali, S. Moro, M. R. Mahmoudi, K. A. Ghaffary, M. Maleki, and A. Alidoostan, “A comparative analysis of classifiers in cancer prediction using multiple data mining techniques,” International Journal of Business Intelligence and Systems Engineering, vol. 1, no. 2, pp. 166–178, 2017.

[19] S. A. Janabi and M. A. Mahdi, “Evaluation prediction techniques to achievement an optimal biomedical analysis,” International Journal of Grid and Utility Computing, vol. 10, no. 5, pp. 512–527, 2019.

[20] D. Shi, J. Guan, J. Zurada, and A. Manikas, “A data-mining approach to identification of risk factors in safety management systems,” Journal of Management Information Systems, vol. 34, no. 4, pp. 1054–1081, 2017.