Impact of Digital Transformation of Engineering Enterprises on Enterprise Performance Based on Data Mining and Credible Bayesian Neural Network Model

Zhenfan Liu

1Fujian Whole Process Engineering Management Co., Ltd., Fuzhou 350000, China
2Université du Québec, Quebec G7H 2B1, Canada

Correspondence should be addressed to Zhenfan Liu; 201712530021020@ctgu.edu.cn

Received 18 August 2022; Revised 16 September 2022; Accepted 19 September 2022; Published 29 December 2022

1. Introduction

At present, the scientific and technological revolution has been carried out rapidly, and certain achievements have been made, among which the digital information technology has accelerated its breakthrough. DT (digital transformation) of enterprises is the foundation of digital economy. Accelerating the development and construction of digital economy and continuously promoting the integration of real economy and digital economy has become a hot spot in the market [1]. Engineering is an important part of China’s economic development, and its degree of development directly determines China’s economic situation. Nowadays, digital economy has been integrated into the daily activities of modern society, and DT has become a challenge that enterprises must face and a threshold that they have to cross. Nearly all business types, across all industries, see creative projects as crucial to the expansion of their companies. Product maintenance and technical upgrading, which users could previously carry out on their own or through other service organizations, became increasingly difficult to carry out with low cost and high efficiency, forcing them to rely on the expert services offered by product manufacturers to resolve issues. This was due to the deepening product complexity in engineering enterprises, which was hampered by the high degree of product specialisation.

The future market is global, and because it is global, consumers need to acknowledge and support it. China has released a number of regulations and legislation to safeguard the future of China’s building machinery with national coercive force in order to increase market competitiveness and long-term consideration. In order to increase the financial income of businesses, it is imperative that all departments comprehend the true wants of clients and develop and sell things with this perspective in mind. By encouraging the intelligent upgrading of businesses and completing the
DT, Chinese engineering companies can accomplish the transformation of old and new kinetic energy and advance towards a stage of high-quality development. Under the tide of digital economy, digital technology gives engineering enterprises a great opportunity to innovate and upgrade [2]. Service innovation has broken through the traditional strategic logic of product innovation relying on technological development, and under the competitive logic of service-led competitive paradigm and customer value as the core, it has become an important strategic goal of engineering enterprises [3]. Therefore, more and more managers try to endow the supply chain with higher digital performance through the extensive use of digital technology.

To achieve in-depth mining of information with high value and potential value, DM (data mining) technology must be applied. DM can be seen as a collection of various complex mathematical models, techniques, and systems. However, there are currently only a small number of publications on the effect of DT on enterprise performance from the perspective of the digital economy. Thus, the study of the DT of engineering enterprises in this paper is beneficial for providing relevant departments with empirical evidence in the microfield to further improve the digital economy policy. Next, relevant departments are urged to establish a full set of DT policy-oriented system to provide guidance for the DT and upgrading of manufacturing enterprises.

The innovation of this paper is mainly reflected in two aspects:

1. The effect of DT on business performance is examined in this research. Based on the already available research on the financial effects of big data application, intelligence, and networking, it shows the trajectory of the influence of DT on business performance.

2. In this paper, a performance prediction method of innovative project portfolio based on Bayesian network and PLS is proposed. It not only expands the application of Bayesian network in performance prediction, but also provides decision support for innovative project portfolio management of enterprises.

2. Related Work

2.1. Research on DT of Enterprises. DT is the future development direction of manufacturing enterprises, and it is also an inevitable trend. Digitalization will make people’s lives change dramatically, and then transform into a new way. Ng and Yee believe that digitalization aims to improve the operational status by using advanced digital tools, including data analysis, cloud computing, artificial intelligence, and other industries [4]. Chen and Metawa found that the relationship between DT and business process management is inseparable [5]. According to Saedti et al., the general trend of global industrial reform and the key to boosting overall national strength is reshaping the value chain structure of traditional industries, enabling traditional industries to be equipped with advanced science and technology, and then bringing new growth vitality to generate [6].

According to Chaillet and Dumont the use of digital technology can improve the synergy effect of each operation link within an organization, streamline the entire business process, and increase the effectiveness of the overall management and control, effectively lowering various costs of the organization [7]. According to Liang’s perspective, the first stage in a company’s DT is to develop its digital thinking culture and extend it to all of its employee’s creative ideas [8]. Esau et al. believe that DT can be divided into three stages from low level to high level, namely, information digitization stage, business digitization stage, and digitization stage. These three stages overlap and develop together [9]. Ma et al. think that if an enterprise wants to gain stronger vitality through transformation, it needs to use the Internet thinking mode and integrate digital technology to reconstruct the enterprise from four aspects: business model, capital model, management model, and mental model [10].

2.2. Research on Enterprise Performance. The current business environment is witnessing the emergence of digital innovation and opportunities, which has fundamentally changed the business environment. At the micro level, the DT driven by technological change will obviously affect the management and operation of enterprises, while the application of big data, intelligence, and networking, as different manifestations of digitalization, will also affect the transformation of production and operation modes of enterprises, and then affect the financial performance of enterprises.

Merkenstein and Lindeque found that big data has a significant effect on enterprise financial risk prediction, and it is one of the effective ways for enterprise financial risk early warning [11]. Freathy and Thomas studied the relationship between social responsibility of food processing enterprises and enterprise performance from the perspective of social responsibility, and found that food processing enterprises that actively undertake social responsibility can obtain considerable economic benefits [12]. Bradford et al. used PLS (partial least squares) to analyze the financial risks of food processing industry and their impact on small and medium-sized enterprises [13].

Shad et al. used network capability, strategic flexibility, and organizational duality to experiment the relationship between these three factors and strategic performance, indicating that network capability has a substantial impact on strategic performance [14]. Moshesh et al., through perfecting the traditional performance appraisal system, divided the strategic performance appraisal system into several related aspects, and considered that the strategic performance system is a recyclable system [15]. Jones and Williams think that investing in R&D can increase the competitiveness of products and thus improve the company’s performance. Compared with those enterprises that do not invest in R&D, the profit rate of R&D investment is nearly 3% higher [16].
Yudianto et al. deeply studied the impact of DT on the production and operation of enterprises and the specific mechanism [17]. It is found that enterprises have changed organizational structure, reconstructed production mode, established a new business model, and broken the previous organizational boundary by using digitalization, which has a significant impact on all links in the closed loop of the whole value chain of enterprises and promoted the improvement of performance.

3. Methodology

3.1. Path Analysis of the Influence of Engineering DT on Enterprise Performance. More and more companies are in a state of rapid growth, and networking also makes the competition among various companies increasingly fierce.

Digital technology connects the ending of enterprise activities and all the links in the middle of activities, so that enterprises can quickly respond to market changes, and the better the digital technology, the faster the response speed. The development of enterprises promotes the optimization of organizations. The transformation of organizational structure has promoted the core competitiveness of enterprises. The promotion of organizational structure transformation on the performance of enterprise groups has been significantly strengthened.

Digital economy is a brand-new system, which transforms the economy and politics in society into a new situation. The borderless nature of the Internet effectively accelerates the global industrial transformation and flattens the organizational structure, which is conducive to the interaction and transmission of information. On this basis, the digital economy also derives the characteristics of economy, inclusiveness, and sustainability. DT is a typical change in the era of digital economy [18, 19], which creates a goal-oriented business model by contacting the core business of enterprises. Enterprises can try to collect data through DT, and use digital methods to increase communication and contact between enterprises, so as to share information. It provides timely, accurate, and intuitive data for production managers. Moreover, enterprises are increasingly relying on “cloud services.” Cloud services provide a steady stream of power for enterprises, which is a necessary condition for enterprises to survive in the digital society. It will also make full use of data to do market and customer research.

The research of DT and effectiveness belongs to the category of strategic management research, and the related theories of strategic management constitute the basic theoretical basis of the research, which has reference significance for the research of enterprise digitalization. Physically, the resources owned by enterprises can be divided into three categories, namely, tangible resources, intangible resources, and capabilities. By using these resources, enterprises can develop and form competitive advantages. From the perspective of the environment itself, high environmental complexity will increase the speed and frequency of enterprise strategic change. From the perspective of the enterprise itself, the higher the openness of the enterprise, the higher the environmental exposure, the more complex the environment it faces, and the greater the disturbance of the environmental uncertainty to the enterprise operation. The technical level pays attention to the characteristics of technology itself and the relationship with the organization, which affects the adoption process of the organization. The organizational level describes that some elements of the organization will affect the adoption of innovative technologies. This paper selects variable R&D intensity and digital investment as technical factors, endogenous financing capacity, enterprise scale and knowledge, technology intensity as organizational factors, environmental complexity as environmental factors, and constructs a research model (see Figure 1) that affects the DT and effectiveness of enterprises.

The success of the DT of the supply chain of an enterprise depends on its environment, and the digital environment is the premise of the organization’s DT of the supply chain. DT of supply chain requires not only a clear digital strategy, but also digital organization and culture. Organizational culture is similar to the attitude of employees. If the established culture is unwilling to change, the result will be a split organization. Intelligent operation focuses on the ability of enterprises to obtain useful information from massive data, which can effectively provide the basis for enterprise decision-making, continuously improve customer experience, strengthen current core business, make full use of information flow, logistics, and capital flow to realize efficient cooperation among supply chain members, improve the visualization and agility of supply chain, effectively respond to the rapidly changing market environment, minimize inventory, and effectively reduce costs [20].

The introduction of numerous digital platforms has reduced the time required for gathering, classifying, analyzing data and information, increased information transparency, decreased information asymmetry, and encouraged the integration and utilization of resources. Through in-depth data analysis, businesses may communicate with customers more effectively with the aid of corporate digital platforms, create a productive window of communication with clients, offer them a novel consumption experience, and raise customer happiness. The effectiveness with which business funds are used is increased, the ability of accounts receivable to be collected, and concurrently, the turnover rate, profit rate, and enterprise development are all accelerated by the corresponding fixed assets, such as inventory. In order to encourage the sharing of data between the supply chain’s upstream and downstream nodes, the supply chain must possess sufficient data processing capabilities to mine user data, comprehend user wants, direct businesses to carry out innovative activities, and deliver individualized services. It is clear that the DT of the supply chain can not only guarantee the quality of goods and services, but also successfully save costs. It can also quickly adapt to the environment of the market, which is changing at a rapid rate, and enhance an organization’s ability to provide goods.

3.2. Enterprise Performance Prediction Based on DM. Among the components of quality cost, prevention cost, appraisal cost, and external quality assurance cost are inputs
to ensure product quality, which can be classified as quality assurance cost. This kind of cost input is the premise of reducing quality loss, increasing quality income and improving the competitiveness of enterprises. By measuring the present value of the company’s future cash flow and considering intangible assets such as ownership and goodwill, the company’s performance can be reflected more comprehensively. By analyzing the value chain of an enterprise, the key achievement areas and key performance evaluation indicators of the enterprise can be determined, and through layer-by-layer decomposition, a three-level key performance indicator system of enterprise, department, and post can be formed.

For engineering enterprises, it is a key asset to "own" the customer interface through their own service organizations. Engineering enterprises can have a deeper understanding of customer’s purchasing center, membership relationship, internal policy situation, etc., through representatives of service organizations who frequently contact with customers, and then put forward customer solutions that match customer’s real needs. First-line service employees implanted in the customer’s site work together with customer employees to establish friendship, share knowledge, and exchange information in the process of solving problems together through frequent communication and long-term interpersonal interaction. This behavior mode can meet the needs of long-term stable relationship between engineering enterprises and customers, and can support the implementation of service strategy of engineering enterprises. However, if engineering enterprises want to create customer value through service transformation, they need to understand the process, resources, and practice of customer management. Therefore, when implementing the customer-centered service strategy, engineering enterprises tend to choose the behavior mode of employee implantation to support the realization of the strategy. Accordingly, it is likely that equipment implantation will not be adopted because it is difficult to grasp the information of customer’s whole operation process.

In the field of retail trade, the classification algorithm can classify and personalize the customers by classifying and counting the data of commodity sales and customer's purchasing power in various regions, thus making the distribution strategy of stores more targeted. At present, there are many mathematical structured models for association rule algorithms, which can be classified according to their application objectives, and can be divided into sequence association model, quantity association model, causal association model, and so on.

The following is a definition of the dynamic neighbourhood radius adaptive density’s attainable distance:

\[ R_A = R \frac{A_i}{A_{i+1}} \]  \hspace{1cm} (1)

where \( R \) stands for the starting density’s attainable distance, and \( A_i \) and \( A_{i+1} \) stand for the density values of two sequentially determined cluster’s density attraction sites. Let \( S[X] \) be the set of \( N \) data sets \( t_1, t_2, \ldots, t_N \) projected to the attribute set \( X \). The distance measurement formula of \( S[X] \) is as follows:

\[ d(S[X]) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} \delta_X(t_i[X], t_j[X])}{N(N - 1)} \]  \hspace{1cm} (2)

The tuple distance is indicated by the letter \( \delta \) in the formula above. The variance of its projected data set to attribute set \( X \) increases with increasing distance from \( S[X] \). The fuzzy c-means algorithm, which is an extension of the K-means algorithm, is based on fuzzy set theory. It is an approach to data clustering that relies on objective function optimization. Currently, this method is frequently utilised and quite well-liked. Flexible fuzzy partitions, such as the FCM technique, are best suited for grouping data sets with ellipsoidal distribution. If a certain k-itemset is called frequent itemset, the minimum weighted support degree is \( \mu \min_{\text{sup}} \), to meet the following condition:

\[ \left( \sum_{t_j \in X \cup Y} \mu_j \right) \sup(X \cup Y) \geq \mu \min_{\text{sup}} \]  \hspace{1cm} (3)

If the association rule \( X \Rightarrow Y \) is interesting and \( X \cup Y \) is a frequent itemset, then its confidence is not lower than the lowest confidence threshold \( \mu \min_{\text{sup}} \).
We calculated the weighted fuzzy support degree \( WFSup(R_{ik}) \) of each fuzzy partition item.

\[
WFSup(R_{ik}) = \frac{\sum_{i=1}^{n} (w_{ijk} \times \mu_{i}^{\theta}(x_j))}{n}
\]  

(4)

We arranged all fuzzy partitions \( R_{ik} \) according to the weight \( w_{ijk} \) from big to small, and generate candidate 1-itemset \( C_1 \).

In the association rules, the support degree of \( Y \) derived by \( X \) represents the ratio of \( X \) and \( Y \) in all records, and the formula is as follows:

\[
\text{Support}(X \longrightarrow Y) = \frac{\text{count}(X \cup Y)}{|D|}.
\]  

(5)

Whether they want to enhance the competitiveness of enterprises, expand the scale of enterprises or get rid of the risk of wearing hats, they are all undergoing DT. At this time, it is very important for them that the performance after DT can be improved. Therefore, we put forward a new forecasting technology, that is, making full use of DM technology to realize the forecasting process, processing the quantitative financial index data of enterprises, and establishing models, so as to provide timely and efficient forecasting support for enterprise managers. The specific process is shown in Figure 2.

We used the method of adding a few classes to make the data set in a balanced state. Down sampling, on the other hand, is a state in which the majority sample and the minority sample are balanced by reducing the majority sample. The selection is based on the following considerations: the choice of one or two methods can be compared and analyzed; secondly, the selected classifier has good generalization ability, which is helpful to the improvement of the model. We can easily solve the selected probability of each branch in the new environment. The results are compared and analyzed to determine which model has a relatively high prediction accuracy when dealing with the binary dependent variables. To determine how comparable the test data set and training data set are, we employ the closest neighbourhood retrieval approach in this process. This method’s calculation formula is as follows:

\[
D_{xy} = \| x - y \| = \sqrt{\sum_{j=1}^{n} (x_j - y_j)^2},
\]  

(6)

where \( D_{xy} \) represents the Euclidean distance between two cases, \( x \) and \( y \) represents two cases, respectively, \( j \) represents the total number of attributes, and \( x_j \) and \( y_j \) represents the attributes corresponding to the two cases, respectively.

In this paper, parameter learning is used to predict the performance through the Bayesian network model of innovative project portfolio prediction, and then the rationality of the model design is proved. Bayesian network parameter learning refers to the conditional probability of each node variable obtained by learning and training the sample data after determining the Bayesian network structure.

Bayesian estimation is to learn all possible values of parameter \( \theta \) with prior knowledge under the network structure \( G \) and sample set \( X \). Bayesian estimation of the expression for estimating \( \theta \) is as follows:

\[
\hat{\theta} = \arg \max_{\theta} P(\theta|D, E).
\]  

(7)

PLS is a form of structural equation model, which has obvious advantages in simulating complex causal relationship among latent variables. Structural equation model is widely used in social science research. The learning rate is not a constant and needs to be automatically adjusted in the process of model training. The automatic adjustment formula is as follows:

\[
\eta(t) = \eta(t-1) \times \exp\left(\frac{\log(\eta_{\text{low}}/\eta_{\text{high}})}{d}\right).
\]  

(8)

In the formula, \( \eta_{\text{low}} \) and \( \eta_{\text{high}} \) represents the minimum value and the maximum value of the learning rate, respectively, and \( d \) is the attenuation.

A cluster \( C \) of \( X \) on the attribute set should be equal to or less than the density threshold \( d_{0}^{X} \) and equal to or more than the frequency value \( s_{0} \), that is,

\[
d(C_{X} [X]) \leq d_{0}^{X}, \quad C_{X} \geq s_{0}.
\]  

(9)

In this process, the prediction function we use is as follows:

\[
H(x) = \arg \max \sum_{i=1}^{T} h_i(x_i).
\]  

(10)

In this way, we can build BaggingX prediction model.

4. Experiment and Results

This paper takes W engineering enterprise as the research object. In order to obtain financial data easily, this paper used SPSS software to analyze the related financial data of this company from 2010 to 2021 and get the relationship between quality cost and enterprise performance. By reading the annual report of the company, the company’s official website, relevant industry research, etc., the enterprises that significantly mention the DT field or digital strategy for further confirmation is selected.

The first-hand information is mainly obtained through in-depth interviews. According to the research questions, the interview outline was drawn up in advance, and the members of the research team conducted in-depth communication with the managers of the case enterprises through on-the-spot visits and video conferences. In addition, during the post-processing of interview records, the interview team kept in touch with the interviewees. When analyzing the solvency of W enterprise, the selected indicators are current ratio, quick ratio, and asset-liability ratio as shown in Table 1. Data visualization of Table 1 is shown in Figure 3.
It can be found that the turnover ratio of W enterprise before and after the transformation is in a relatively normal range, indicating that the enterprise has strong liquidity. From the perspective of long-term solvency, enterprises have invested a lot of money in introducing talents, R&D costs, fixed assets, etc. According to the data in the table, from 2011 to 2015, the asset-liability ratio of W enterprises was relatively stable, but all of them were in a high position. It also guarantees the company's long-term debt repayment ability.

Taking 2015 as the dividing line before and after the transformation, this paper makes a longitudinal comparative analysis of a number of financial indicators before and after the transformation of W enterprise and preliminarily judges whether the DT has improved the performance of the enterprise, what specific aspects have been improved, and whether the operating conditions of W enterprise have improved during the whole transformation process. Table 2 shows the changes of operating profit margin, gross profit margin, and net profit margin of W enterprise from 2010 to 2021. Figure 4 shows the data visualization results of Table 2.

Compared with before the transformation, the period expense ratio of W enterprises has remained at around 17.3% in the period from 2010 to 2021, and this indicator of W enterprises has been rising year after year since 2017. In

<table>
<thead>
<tr>
<th>Year</th>
<th>Liquidity ratio</th>
<th>Quick ratio</th>
<th>Asset-liability ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>3.762</td>
<td>1.199</td>
<td>58.496</td>
</tr>
<tr>
<td>2011</td>
<td>1.043</td>
<td>1.23</td>
<td>58.386</td>
</tr>
<tr>
<td>2012</td>
<td>3.366</td>
<td>1.156</td>
<td>58.092</td>
</tr>
<tr>
<td>2013</td>
<td>3.124</td>
<td>1.044</td>
<td>59.517</td>
</tr>
<tr>
<td>2014</td>
<td>1.472</td>
<td>1.127</td>
<td>58.549</td>
</tr>
<tr>
<td>2015</td>
<td>2.889</td>
<td>1.22</td>
<td>58.107</td>
</tr>
<tr>
<td>2016</td>
<td>3.025</td>
<td>1.094</td>
<td>61.072</td>
</tr>
<tr>
<td>2017</td>
<td>2.517</td>
<td>1.187</td>
<td>59.477</td>
</tr>
<tr>
<td>2018</td>
<td>4.202</td>
<td>1.033</td>
<td>61.154</td>
</tr>
<tr>
<td>2019</td>
<td>1.578</td>
<td>1.031</td>
<td>58.905</td>
</tr>
<tr>
<td>2020</td>
<td>2.593</td>
<td>1.154</td>
<td>58.896</td>
</tr>
<tr>
<td>2021</td>
<td>2.166</td>
<td>1.252</td>
<td>60.411</td>
</tr>
</tbody>
</table>

Figure 2: DT performance prediction process.
the whole range, the financial expenses of W enterprise are almost in a state where the interest income is greater than the interest expenditure because of the smooth development of its own financial business. Even if there is a certain amount of financial expenses, its proportion and influence are almost negligible. The company’s sales expenses mainly come from product installation and after-sales, advertising expenses, logistics and transportation, and the salary of sales staff, among which the salary of employees is the lowest. DT is characterized by changes and transformations, which are driven and built on the basis of technology. DT of enterprises must rely on technology, and it is changeable and innovative. This paper counts the R&D investment intensity of W enterprises from 2010 to 2021 (see Figure 5).

It can be seen from this that from 2010 to 2021, the R&D expenditure and R&D investment intensity of W enterprises increased overall, especially the R&D investment intensity increased continuously except in 2015, which shows that the independent innovation capability of enterprises is constantly improving and the investment in R&D is also increasing. The measurement data of portfolio performance and its influencing factors mainly come from questionnaire

![Figure 3: Visualization of flow ratio and quick ratio data.](image-url)
Table 2: Changes in operating profit margin, gross profit margin, and net profit margin of W enterprise from 2010 to 2021.

<table>
<thead>
<tr>
<th>Year</th>
<th>Operating profit margin</th>
<th>Gross profit margin of sales</th>
<th>Net profit margin on sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>5.448</td>
<td>17.344</td>
<td>3.713</td>
</tr>
<tr>
<td>2011</td>
<td>5.73</td>
<td>17.675</td>
<td>3.571</td>
</tr>
<tr>
<td>2012</td>
<td>5.845</td>
<td>17.598</td>
<td>6.62</td>
</tr>
<tr>
<td>2013</td>
<td>8.278</td>
<td>20.17</td>
<td>6.421</td>
</tr>
<tr>
<td>2014</td>
<td>7.656</td>
<td>21.135</td>
<td>6.197</td>
</tr>
<tr>
<td>2015</td>
<td>9.481</td>
<td>21.22</td>
<td>7.044</td>
</tr>
<tr>
<td>2016</td>
<td>9.782</td>
<td>24.386</td>
<td>8.659</td>
</tr>
<tr>
<td>2017</td>
<td>9.381</td>
<td>23.334</td>
<td>8.279</td>
</tr>
<tr>
<td>2018</td>
<td>11.837</td>
<td>23.979</td>
<td>5.24</td>
</tr>
<tr>
<td>2019</td>
<td>11.397</td>
<td>27.208</td>
<td>4.992</td>
</tr>
<tr>
<td>2020</td>
<td>11.217</td>
<td>27.512</td>
<td>8.598</td>
</tr>
<tr>
<td>2021</td>
<td>12.278</td>
<td>27.249</td>
<td>9.182</td>
</tr>
</tbody>
</table>

Figure 4: Data visualization results.

Figure 5: R&D innovation of W enterprise from 2010 to 2021.
The more samples are surveyed, the more time, money, and manpower are consumed. As shown in Figure 6, the average accuracy of verification samples is 0.652, while that of training samples is 0.726 when the sample count is 150. Additionally, sample sizes fewer than 100 will not be taken into account because they cause the Bayesian network optimised by PLS to change in structure.

The newly gathered sample data is used as the test set, with the distinction that all previously unbalanced data were used as the training set and that those previously unbalanced data were used as the training set after being balanced, allowing for a comparison and analysis of the two outcomes, as depicted in Figure 7.

Whether it is a single model, a cluster mixed model or a cluster fusion model, the results of each model based on unbalanced data as training set are better than those of balanced data. On the other hand, the result of true negative rate is just the opposite, that is, the balanced data as a training set is better than the unbalanced data as a whole. This may be because the sample of DT failure in the selected research object is small, which makes the results of each model tend to be consistent. Therefore, from the comparison of this set of data, we can see that case-based reasoning and its classifier combined with integration and clustering are more suitable for unbalanced data sets, which not only accords with the actual law, but also has certain practical value. DT has become an inevitable choice for companies. Only by defining the strategic goal of DT of enterprises and fully understanding the existing valuable resources of the company, such as capital, technology, talents, etc., can we find an efficient path suitable for enterprises in the development. The company can achieve the purpose of improving efficiency, organizational management ability, and long-term sustainable development through DT, but the company cannot put all the hopes of improving enterprise performance in the DT, thus ignoring the control of other processes.

5. Conclusion

More and more companies are in a state of rapid growth, and networking also makes the competition among various companies increasingly fierce. In the era of digital economy, DT has become the common choice of most enterprises. Taking W enterprise as an example, this paper studies the path of its DT affecting its performance, and makes targeted optimization suggestions for the enterprise in combination with its own development. In order to support enterprise managers with timely and effective forecasting, we therefore propose a new forecasting technology that fully utilises DM technology to realize the forecasting process. This new technology involves processing quantitative financial index data of enterprises and establishing models. The findings indicate that as the number of samples rises, the average accuracy of training samples gradually declines while the average accuracy of testing samples gradually rises. When there are 150 samples, the average accuracy of the training samples is 0.726 and the average accuracy of the verification samples is 0.652.

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 conflicts of interest.
References


