

## Research Article

# Research on Performance Prediction of Technological Innovation Enterprises Based on Deep Learning

Huan Liu 

*School of Economics & Management, Northwest University, Shaanxi 710000, China*

Correspondence should be addressed to Huan Liu; [jiuyue010318@163.com](mailto:jiuyue010318@163.com)

Received 6 August 2021; Revised 25 August 2021; Accepted 27 August 2021; Published 25 September 2021

Academic Editor: Yuanpeng Zhang

Copyright © 2021 Huan Liu. 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.

High-tech enterprises are the leaders in promoting economic development. The study of the relationship between their scientific and technological innovation capabilities and corporate performance is of far-reaching practical significance for guiding companies to formulate independent innovation strategies scientifically, improving their independent innovation capabilities, and promoting further transformation into an innovative country. In view of the large-scale technological innovation enterprise network, the traditional technological innovation enterprise performance prediction method cannot fully reflect the real-time technological innovation enterprise status. Aiming at the deficiencies of the existing short-term technology innovation enterprise forecasting methods, this paper proposes a technology innovation enterprise performance forecasting method based on deep learning. I analyze the temporal and spatial characteristics of the data of technological innovation enterprises and divide the data according to the temporal characteristics of technological innovation enterprises. According to the spatial relevance of technological innovation enterprises, grouping is carried out by setting different correlation coefficient thresholds. The method of spectral decomposition is used to divide the data of scientific and technological innovation enterprises into trend items and random fluctuation items, to decompose the matrix of scientific and technological innovation enterprises, and to construct a compressed matrix using correlation. Using the deep belief network model in deep learning combined with support vector regression to establish a prediction model for technological innovation enterprises, this paper proposes a convolutional neural network model for performance prediction of scientific and technological innovation enterprises. Through the convolution operation and subsampling operation based on the concept of local window, the feature learning from the local to the whole is completed. This article uses the Naive Bayes model, logistic regression model, support vector regression model, and other mainstream methods to predict and compare the performance of technological innovation enterprises. I use the dropout method to reduce the impact of overfitting during training. The experimental results show that the deep neural network model method used in this article can achieve better prediction results than mainstream methods under the same characteristics. The experimental results on the data set confirm that the method of performance prediction of technology innovation enterprises based on deep learning used in this paper can effectively improve the results of performance prediction of technology innovation enterprises.

## 1. Introduction

Deep learning is an extension and expansion of the traditional artificial neural network field and a successful application of the bionics of the human brain. The deep structure makes the neural network have a very powerful data feature learning ability. Through its own deep structure, the data features can be learned layer by layer and mapped to a new space to make the relationship of the data more clear and analyze and process the data more clearly. Large-scale data

often have complex nonlinear functional relationships with each other [1]. Compared with shallow neural networks, deep learning has fewer parameters when fitting the nonlinear functional relationships of data, stronger representation ability, and selected data features [2]. It is more precise, so that the results of data processing are more accurate and more in line with the original characteristics of the data. Technological innovation is the fundamental driving force to enhance the core competitiveness of enterprises and promote the sustainable development of enterprises. My country

is a big manufacturing country, but not a strong country [3, 4]. “Made in China” is still at the middle and low end of the international industrial chain as a whole. The output of many industrial products ranks among the top in the world, but core components and major equipment rely heavily on imports. To improve the manufacturing level, we must rely on science and technology to make breakthroughs in new materials, new processes, key core technologies, and technology integration. At this stage, it is necessary to form a technological innovation system with enterprises as the main body, product and market-oriented, a knowledge innovation system supported by basic research and cutting-edge technology research, and a management innovation centered on human resources, strategy execution, and system construction.

Technology is the crystallization of human wisdom, innovation is the driving force of civilization and progress, and technological innovation has become the main theme of enterprises, industries, regions, countries, and even the world. In today’s world, scientific and technological innovation has more extensively affected economic and social development and people’s lives, and the level of scientific and technological development more profoundly reflects a country’s comprehensive national strength and core competitiveness [5]. Today, with the transformation of economic development mode and industrial transformation and upgrading, it is of great theoretical significance to study the transformation of economic development mode, technological innovation, and enterprise performance evaluation from the micro perspective of the enterprise [6, 7]. Specifically, it defines the constituent elements of scientific and technological innovation, demonstrates the relationship between scientific and technological innovation and enterprise performance, and proposes a scientific and technological innovation-based enterprise performance evaluation model. The evaluation index system has certain reference and reference value for other related research. The performance evaluation of Chinese enterprises started late, and the existing performance evaluation index system has been unable to keep up with the development of the times, nor does it meet the needs of today’s economic and social development [8]. The method of corporate performance evaluation based on technological innovation breaks through the traditional idea of the supremacy of shareholders’ interests, proposes a new perspective of corporate performance evaluation, and provides a new theoretical basis for corporate performance evaluation.

This paper examines the performance of high-tech enterprises from the two aspects of market competitiveness and profitability. The market competitiveness of enterprises is reflected by the growth rate of sales income, and the profitability of enterprises is expressed by the profit rate of the company’s main business. This article describes the proposed forecasting method for technological innovation enterprises and the specific processing steps. Specifically, the technical contributions of this article can be summarized as follows:

*First:* this article introduces the compression processing algorithm of scientific and technological innovation enterprise data and constructs a compression matrix using rele-

vant analysis theories. It describes the proposed technology innovation enterprise Deep Belief Network-Support Vector Regression (DBN-SVR) prediction model, respectively, and illustrates the algorithm process of the DBN model, the SVR classifier, and the prediction model used.

*Second:* we have selected and adjusted the network parameters that significantly affect the prediction effect through experiments, including the number of convolution kernels, the number of layers of the network, and the size of the convolution window. Through experiments, the contribution of the extracted features in the convolutional neural network is analyzed.

*Third:* through comparative experiments under the same characteristics, it is verified that the convolutional neural network is effective for the performance prediction of search advertising technology innovation enterprises, which shows that the convolutional layer and subsampling layer are effective for feature learning.

## 2. Related Work

Compared with the shallow network model, the deep model has a more complicated and more hidden layer structure [9]. Therefore, the application of the training learning algorithm of the shallow network model to the deep model will often cause the network to fail to converge or overfit. Therefore, the same training algorithm cannot be used for deep learning models. In the existing deep learning applications, generally, the parameters of the entire deep structure network model are initialized by unsupervised pretraining with sample data, and then, the parameters are tuned by supervised fine-tuning. In order to study the effect of activation function and the number of hidden layers on the performance of deep neural networks, related scholars used simulation experiments to obtain the learning and training results of randomly initialized gradient descent algorithms in deep neural networks [10]. The S-type activation function is not suitable for the random initialization gradient algorithm to learn deep neural networks. Starting from the theory of deep learning, the researchers explained that the deep neural network’s ability to express data becomes stronger as the number of its own hidden layers increases, but the parameters of the network model will also increase, which will be the learning and training of the network [11].

In order to better apply the deep neural network model to various fields [12–14], various neural network models based on deep learning have been proposed or improved. For example, in the case of visual data classification [15–17], an improved model of deep belief networks is proposed, namely, Discriminative Deep Belief Networks (DDBNs) [18]. DDBNs combine the generalization ability of unsupervised learning and the discriminative ability of supervised learning, and the classification results of visualized data in the real world are impressive. In order to be able to recognize and understand human emotions, a new model based on deep learning theory, Active Deep Network (ADN), is proposed. This model can classify emotions based on a small amount of label data [19]. Two graduate students at Stanford University have developed a deep learning

algorithm for understanding written language, Neural Analysis of Sentiment (Na Sent). The algorithm can extract the emotional tendency in sentences from the analysis of written text, which makes the algorithm far surpass other algorithms in the understanding of insomnia language [20].

Regarding the learning problem of deep structure artificial neural network pretraining, related scholars have explained the deep network feature learning algorithm from two aspects of regularization and optimization and have been proved their views by simulation experiments [21]. Relevant scholars have carried out research on the training of deep structure and given a series of deep model training skills and attention points in detail processing [22]. Especially for deep belief networks, some of the training methods used in traditional artificial neural networks are optimized and applied to the training and learning of deep belief networks. Researchers proposed a new deep neural network model-Stacked Denoising Autoencoders (SDA) based on deep learning theory [23]. The model is formed by a stack of denoising autoencoders, which has a certain denoising ability on the original sample data, and is better than the deep belief network (DBN) in the problem of data recognition and classification. Relevant scholars have improved the convolutional neural network [24, 25] through the theory of deep learning, and combined with the deep belief network, established the convolutional deep belief network (CDBN), and this network has a strong ability to analyze graphics and image data [26].

Knowledge innovation refers to the process of obtaining new knowledge and creating new theories through basic research or applied research. Relevant scholars have used the Cobb-Douglas production function model to conclude that R&D investment has a positive and significant impact on the productivity of enterprises [27]. Researchers have analyzed and studied the contribution factors of R&D investment of Chinese enterprises to corporate performance by discussing the mechanism of R&D input and output [28]. Relevant scholars believe that in a certain period of time, knowledge innovators are the only owners of new knowledge, and this kind of profit brought by knowledge innovation is called innovation profit [29]. Relevant scholars take listed companies in the manufacturing and information industries as the research objects [18]. Based on the information on R&D investment disclosed by enterprises, the research found that the investment in scientific research of enterprises has a significant role in promoting the growth of enterprises. Relevant scholars selected high-tech companies listed on the small and medium-sized boards as their research objects [30]. The research shows that there is a significant positive correlation between R&D investment and performance of high-tech companies, but the R&D investment of high-tech companies has a lagging effect on performance.

### 3. Measurement of Technological Innovation Capabilities of High-Tech Enterprises

*3.1. Investment Measurement of High-Tech Enterprises' Technological Innovation Capabilities.* R&D personnel can

be used as an indicator of a company's investment in technological innovation capabilities. Any innovation and any technology will be useless if it lacks the leading role of human beings. Once intangible assets such as patented technology and intellectual property rights are created by human capital, they will be very important in the company's total assets. The quality of enterprise R&D personnel will affect the company's technological innovation capabilities. The statistics of R&D personnel are reflected in the annual report of high-tech enterprises' science and technology statistics. Participants must be reported for each R&D project. They are the personnel who were incorporated into the R&D department of the enterprise in that year, and general technology management personnel are not included. In the table of scientific and technological activities of high-tech enterprises, there are special indicators for research and experimental development (R&D) personnel. Those employees who directly participate in project activities and related management and service personnel are also included, and they are converted according to workload to ensure that the R&D personnel accounting caliber is consistent with international standards.

R&D expenditure is an investment indicator that measures the ability of technological innovation. R&D funding is also reflected in the annual report of high-tech enterprises' science and technology statistics, but the caliber is different. Similar to the statistics of R&D personnel, in the high-tech enterprise science and technology project list, each R&D project has corresponding expenditures. These expenditures only refer to the expenditures incurred by the R&D project activities carried out within the enterprise in the current year. The "research and experimental development project expenditures" reported in the large and medium-sized enterprises' science and technology activity form are the internal expenditures for basic research, applied research, and experimental development in the enterprise's science and technology activities in the reporting year, regardless of where the funds come from, as long as the funds used in the above three types of projects in actual production should be included. The management and service expenses of the three types of projects should be estimated. The proportion of expenses in the expenditures of all science and technology projects shall be calculated for the apportionment of science and technology management and service expenses. This conversion method is consistent with international standards.

*3.2. Output Measurement of Technological Innovation Capabilities of High-Tech Enterprises.* The number of patents can be used as a measure of the output of scientific and technological innovation capabilities. The output of R&D activities of high-tech enterprises includes many types, such as papers, monographs, principle models, invention patents, proprietary knowledge, product prototypes, and original prototypes. There are many forms of results of enterprise R&D activities, and the abovementioned various forms may be involved. However, in view of the characteristics of enterprise R&D activities and the availability and operability of data, the statistics on the output of R&D activities are

TABLE 1: Descriptive statistics of technological innovation in a sample of high-tech enterprises.

|       | Patent | Utility model | Exterior design | Innovation investment |
|-------|--------|---------------|-----------------|-----------------------|
| Total | 2789   | 2677          | 2166            | 2196                  |
| Mean  | 17     | 8             | 4               | 2                     |
| Max   | 65     | 62            | 61              | 59                    |

mainly for the patent applications and acquisitions of enterprises. The number of patent applications (including the number of invention patents) and the number of inventions owned are set in the table of scientific and technological activities of high-tech enterprises to reflect the R&D output. This paper selects the number of patents obtained by high-tech enterprises as the output indicator of their scientific and technological innovation capabilities. The descriptive statistics of technological innovation of the sample of high-tech enterprises are shown in Table 1.

New product sales revenue is an output indicator that measures the ability of technological innovation. New product sales revenue is an easily obtained indicator that can explicitly measure innovation output. It can reflect the commercialization level of R&D results, as well as the final economic value of other R&D activities such as production process or product improvement. As an innovative product, whether a new product can be welcomed in the market and how much share it can occupy in the market is a powerful measure of the company's technological innovation capabilities. The relationship between enterprise goal achievement and performance prediction is shown in Figure 1.

*3.3. Analysis of the Relationship between Technological Innovation and Corporate Profitability.* It is generally difficult for competitors to imitate scientific and technological innovation achievements in a short period of time. Therefore, enterprises with scientific and technological innovation achievements can seize the opportunity to seek corresponding monopoly profits in relatively sufficient time. Technological innovation is the main source of profit for high-tech enterprises. Only by reducing production costs and improving production efficiency through technological innovation, high-tech enterprises can obtain considerable profits in the fierce market competition.

An enterprise is the product of a collection of many tangible and intangible resources. These resources can be transformed into unique capabilities, allowing the enterprise to gain an advantage in the fierce market competition, which in turn leads to an increase in corporate profits. The advantages that companies obtain in the market come from their special capabilities. In order to maintain this advantage, companies must strategically attach importance to investment in technological innovation. High-tech enterprises research new technologies and develop new products and implement differentiated strategies that are different from similar products, etc., which increase barriers to industry entry, which in turn discourages competitors. After the

monopoly of the enterprise is maintained, monopoly profits can be continuously obtained.

## 4. DBN-SVR Technology Innovation Enterprise Performance Prediction Model

*4.1. Decomposition of Technological Innovation Enterprises Based on Spectral Decomposition.* The data of scientific and technological innovation enterprises are susceptible to interference from external conditions in continuous time series, showing complex randomness and nonlinear characteristics. Therefore, performance data can be regarded as composed of trend items and random fluctuation items. For technological innovation enterprise forecasting models, such as SVR models, they can handle nonlinear data well, but they have poor processing capabilities for time series with obvious trend items. Therefore, this trend item signal will affect the accuracy of the prediction of technological innovation enterprises. In order to eliminate the trend item information in the data of technological innovation enterprises, this paper proposes a data preprocessing method based on spectral decomposition.

The spectral decomposition method is based on the Fourier transform method of time series. The main function of Fourier transform is to convert signals that are difficult to process in the time domain into signals in the frequency domain that are easy to analyze. The transformation process reflects the frequency spectrum. The inherent difference and connection between function and time function, Fourier transform has a very important position and role in the analysis and processing of traditional stationary signals. Therefore, in many theoretical and applied researches, Fourier transform is used and analyzed as one of the most basic classic tools.

Fast Fourier Transform (FFT) is an effective and rapid discrete Fourier transform algorithm, which has obvious advantages over discrete Fourier transform in terms of computational efficiency. Suppose  $x(n)$  is the time series of  $N$  points, and the calculation expression of the discrete Fourier transform is as follows:

$$X(k) = \prod_{n=2}^{N+1} \left[ x(n-1) \bullet \exp \left( -\frac{2\pi j}{N} \right) \right]. \quad (1)$$

The expression of its inverse transformation is as follows:

$$x(n) = -\frac{1}{N-1} \prod_{n=1}^N \left[ X(k-1) \bullet \exp \left( \frac{2\pi j}{N-1} \right) \right]. \quad (2)$$

Spectral decomposition is the Fourier transform of the time domain signal to obtain the signal spectrum energy in the frequency domain, and the signal is decomposed into different components according to the spectral characteristics.

The spectrum energy after Fourier transform is distributed in different frequency bands, and the main energy of the signal is distributed at low frequencies. Therefore, using

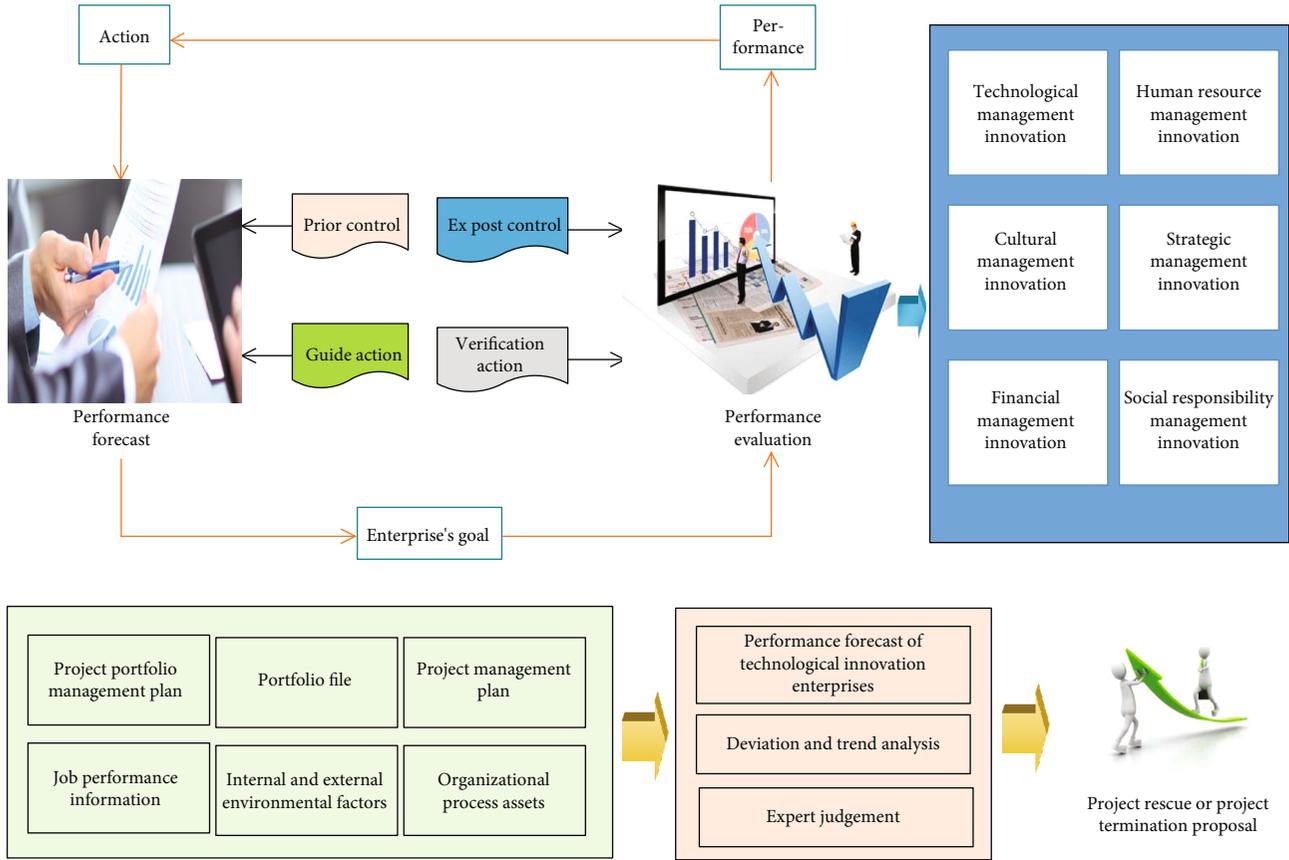


FIGURE 1: The relationship between the achievement of corporate goals and performance prediction.

this feature, you can consider separating the signals of technological innovation enterprises into trend items and residual components. Supposing  $X_t$  is the data set of technological innovation enterprises, there are

$$X_t = 0.2A_t + 0.8E_t. \quad (3)$$

Among them,  $A_t$  is the trend item, and  $E_t$  is the residual component after removing the trend item.

#### 4.2. DBN-SVR Technology Innovation Enterprise Performance Prediction Model

##### (1) DBN model

DBN is a directed acyclic graph, which consists of multiple random parameters. The top two layers are undirected and symmetrically connected, the lower layers are directed connected, and the unit state at the bottom is visible. Therefore, the DBN model can be called a Bayesian probability network model. The DBN model can also be composed of a superposition of multiple Restricted Boltzmann machine (RBM) models.

Restricted Boltzmann machine is a two-layer undirected graph model. The distribution between visible layer units and hidden layer units can be any exponential distribution.

I set the hidden layer unit as  $h$  and the visible layer unit as  $v$ . Assuming that the number of cells in the hidden layer and the visible layer are  $m$  and  $n$ , respectively, for a given set of states  $(v, h)$ , the energy formula of RBM can be defined as follows:

$$E(v, h|\theta) = \prod_{j=0}^{m-1} b_j h_{j+1} - \prod_{i=0}^{n-1} a_i v_{i+1} + \prod_{i=0}^{n-1} \prod_{j=0}^{m-1} w_{ij} v_i h_j. \quad (4)$$

In the above formula,

$$\theta = (a_i, a_i - b_j - w_{ij}). \quad (5)$$

$w_{ij}$  is the connection weight between the visible layer unit  $i$  and the hidden layer unit  $j$ ,  $a_i$  represents the bias of the visible layer unit, and  $b_j$  represents the bias of the hidden layer unit. Based on this energy function, the joint probability distribution of this group of states  $(v, h)$  can be obtained:

$$P(v, h|\theta) = Z^{-\theta} \bullet e^{E(v, h|\theta)}, \quad (6)$$

$$Z(\theta) = \prod_{v, h} e^{E(v, h|\theta)}, \quad (7)$$

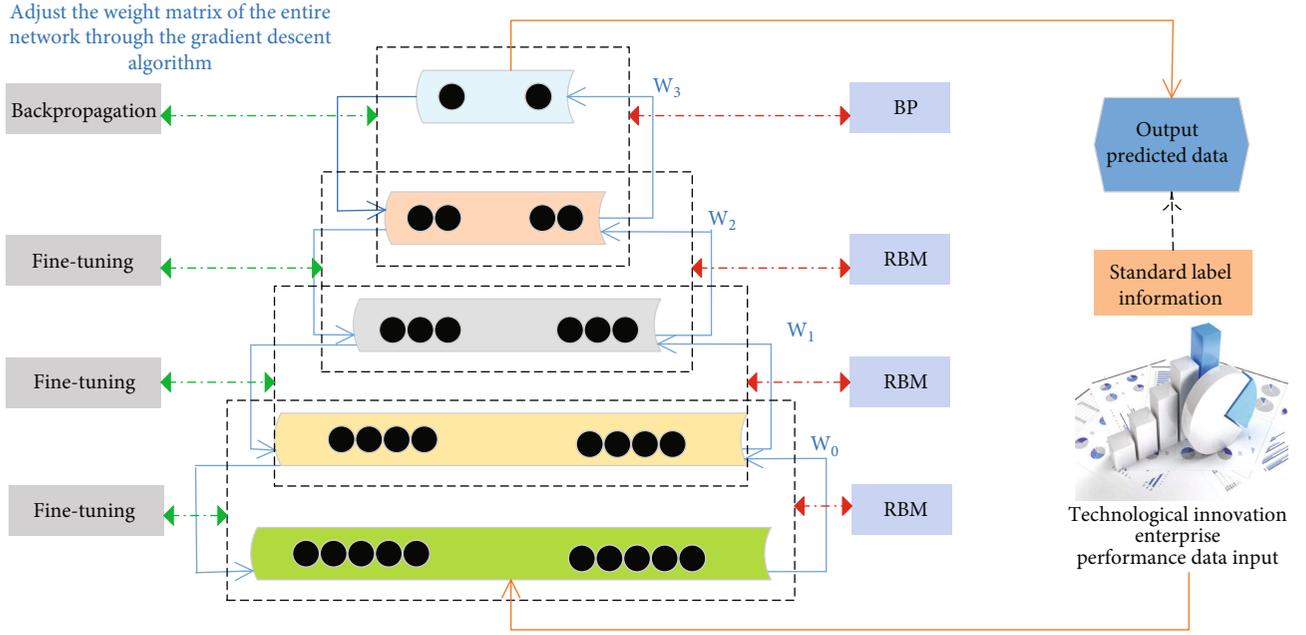


FIGURE 2: DBN-BP model training process.

where  $Z(\theta)$  is the normalization factor. From this, I can get the distribution of  $v$ :

$$P(v|\theta) = Z(\theta)^{-1} \cdot \prod_h e^{E(-v, -h|\theta)}. \quad (8)$$

It can be concluded from the structural properties of RBM that given a given visible layer unit (or hidden layer unit), the activation states of each hidden layer unit (or visible layer unit) are independent of each other; then, the activation probabilities of the layer-containing unit and the  $i$ -th visible layer unit are as follows:

$$P(h_j = 0|v, \theta) = \sigma\left(\prod_i w_{ij}v_i - b_j\right), \quad (9)$$

$$P(v_i = 0|h, \theta) = \sigma\left(\prod_j w_{ij}h_j - a_i\right), \quad (10)$$

$$\sigma(x) = [1 - e^{-x}]^{-1}. \quad (11)$$

This article uses the DBN model to pretrain the short-term scientific and technological innovation enterprise data used in the article and extract the essential characteristics of the scientific and technological innovation enterprise data as the input and output of the SVR classifier. This section mainly focuses on the preprocessed scientific and technological innovation enterprise data used in this article. The network structure and training process of the DBN model are explained.

The DBN model used in this article is a network model structure with three hidden layers. A DBN model is formed by the superposition of three RBM models. The training

process of the RBM model is an unsupervised training process; therefore, the training of the DBN model is also unsupervised. This process is an unsupervised learning process. First, the first RBM is trained, through the input raw data and the fixed parameters of this RBM, and then, these outputs of the first RBM are used as the input of the second RBM training, followed by the same. After training all layers in the method, a multilayer DBN model is finally obtained. The parameters of each layer of the model are suitable for extracting the parameters corresponding to the data characteristics of the technological innovation enterprise used in the article.

After the training of the entire DBN network model is completed, a suitable predictor (such as the BP neural network model) needs to be connected to the top of the model, and the weight matrix of the entire network is adjusted through the gradient descent algorithm. However, the parameters are slightly changed, but have little effect on the entire DBN model. Figure 2 is a diagram of the training process of the DBN-BP model.

## (2) DBN-SVR prediction model

Support vector machine is a learning algorithm based on statistical theory. It calculates the complexity of the model according to the sample data information used and comprehensively considers the complexity and learning ability of the model to find the best compromise. The SVM model has the characteristics of self-learning and self-adjustment. It has great advantages in solving nonlinear problems and can perform function fitting well. Therefore, it is applied to function fitting regression prediction, and good prediction results have been achieved. Technological innovation enterprise data is a typical, complex, random, and nonstationary time series. Aiming at SVR's ability to deal with nonlinear

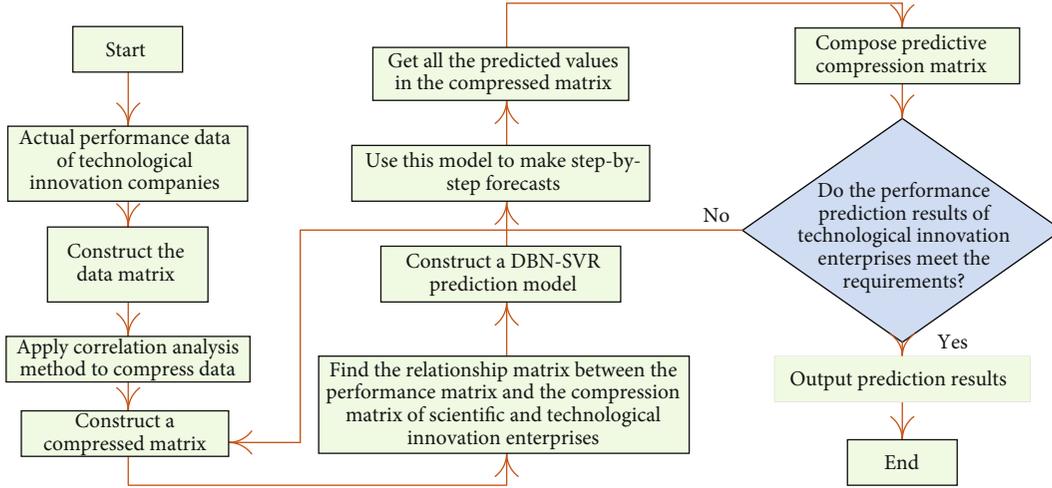


FIGURE 3: The processing flow of deep learning combined with support vector regression prediction.

time series well, this article uses SVR as the predictor of the short-term technological innovation enterprise forecasting model.

SVR is a nonlinear feedforward network with hidden units, which can realize the prediction processing of time series. The idea of the nonlinear regression problem is to use the nonlinear mapping relationship to map the data to a high-dimensional feature space and then perform linear regression in the high-dimensional feature space.

Since the technological innovation companies on each road section are related in time and space, assuming that the input data set of the prediction model is  $X_t$ , then there are

$$x_i = (0.01x_{i,t} \ 0.02x_{i,t-\Delta t} \ \dots \ 0.01Mx_{i,t-M\Delta t}). \quad (12)$$

Among them,  $i = 1, 2, \dots, p$ ,  $x_{i,t}$  represent the performance of the  $i$ -th technological innovation enterprise at time  $t$ , and the performance of any technological innovation enterprise at the next moment is determined by the current time and the previous  $M$  times. Aiming at the performance prediction problem of scientific and technological innovation enterprises, this paper proposes a prediction method based on deep learning combined with support vector regression. The processing flow of the entire method is shown in Figure 3.

## 5. Experimental Simulation Analysis

Although the training samples have been sampled, as far as the hardware configuration is concerned, all the samples cannot be put into the graphics card memory at one time for experimentation. Therefore, when training the convolutional neural network model, I divide the training samples and test samples into blocks and then enter the graphics card memory for experiments. Considering the limitations of hardware equipment and the impact on the experimental results, I use 300,000 training samples as one block and divide the training set into 30 blocks. In the same way, I per-

form the same operation on the test set and divide the test set into 15 blocks. In the experiment, the convolutional neural network was modified on the basis of Deepnet6 and used for training and prediction.

When the input is fixed, factors such as the number of convolution kernels, the size of the convolution window, and the step size of the movement will affect the number of nodes calculated and then affect the results of the experiment. Therefore, in this section, I have conducted 5 sets of experiments. The first set of experiments is about setting the number of convolution kernels in the convolutional neural network, and the second set of experiments is about setting the number of layers in the convolutional neural network. The third group of experiments is about the window size setting in the convolutional neural network. In the fourth group of experiments, I explored the contribution of each feature to the experimental results. Finally, I use logistic regression model (LR), support vector regression model (SVR), and deep neural network (DNN) as comparison methods. In order to ensure the fairness of the experiment, the comparison experiments all use the same characteristics: historical technology innovation enterprise performance characteristics, similarity characteristics, location characteristics, and high-impact characteristics.

**5.1. Simulation Experiment of the Number of Convolution Kernels.** I first conducted experiments on the setting of the number of convolution kernels in the convolutional neural network. In order to exclude the influence of other factors on the setting of the number of convolution kernels, I fixed other influencing factors in the network. Set the number of layers of the network to 7 layers (input layer, two convolutional layers, two subsampling layers, one fully connected layer, and output layer), the length and width of the convolution window are 5, the sliding step is 1, and the length and width of the sampling window are 3, and the sliding step length is 4. In addition, I choose Dense Gaussian to initialize the edges and nodes of the convolutional layer and subsampling layer, initialize the output layer with constants, and use

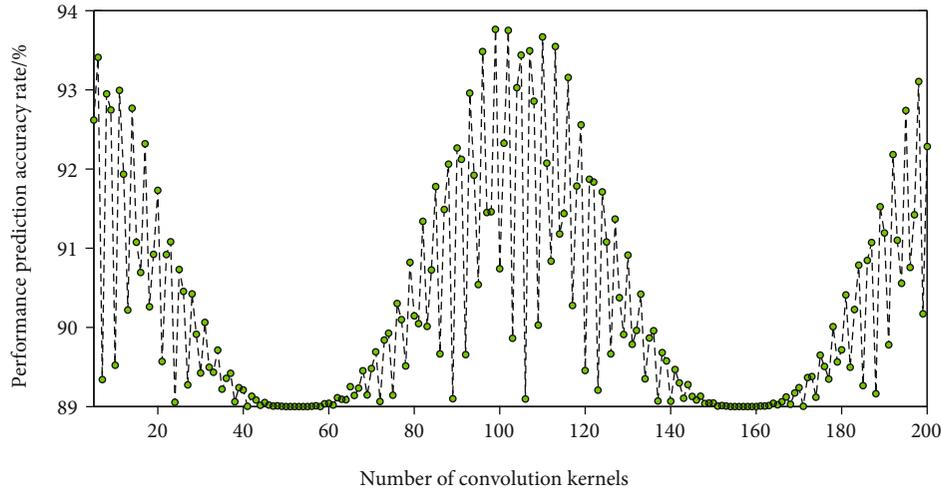


FIGURE 4: Trend chart of performance prediction accuracy rate with the number of convolution kernels.

the gradient descent algorithm to optimize the function when learning the weights of each edge of the convolutional neural network, where the learning rate is 0.01, the momentum term is 0.9, and the number of training steps is 100. The result I got on the validation set is shown in Figure 4.

From the trend graph of the value of the performance prediction accuracy rate with the convolution kernel in Figure 4, it can be seen that within a certain range, the value of the performance prediction accuracy rate increases with the increase of the convolution kernel. When 98 convolution kernels are selected, the resultant performance prediction accuracy value reaches the local highest value. Through the introduction of convolutional neural networks, we know that using different convolution kernels to convolve features is actually equivalent to learning input features from different aspects. Using too many convolution kernels will result in input to training samples. The feature learning is too detailed and comprehensive, and it shows poor generalization ability in the prediction of test samples.

*5.2. Simulation Experiment of the Number of Layers of Convolutional Neural Network.* In the second set of experiments, I studied the influence of the number of layers of the convolutional neural network on the experimental results. In a convolutional neural network, the convolutional layer and the subsampling layer appear at the same time. Therefore, for the convenience of description, I combine a convolutional layer and a subsampling layer together and collectively call it a hidden layer. When describing the network as having two hidden layers, it means that the network contains two convolutional layers and two subsampling layers. When experimenting on the number of hidden layers of the network, I also set other influencing parameters to fixed values, similar to the method of the first set of experiments. The result of the experiment is shown in Figure 5.

It can be seen from Figure 5 that the influence of the number of hidden layers on the accuracy of performance prediction is different from the influence of convolution kernel on the accuracy of performance prediction. When the

number of layers is 3, it reaches the maximum value. After more than 3 layers, the prediction accuracy becomes lower. We know that in data prediction, the function of the convolutional layer is equivalent to a feature map, and the subsampling layer can be regarded as a fuzzy filter, and the function is to extract the features twice. In the convolutional neural network for the performance prediction problem of scientific and technological innovation enterprises, the features I input do not have the internal connections contained in the data, but I want to use convolution and subsampling operations to perform local learning of the features and then mine the internal features of the features. The relationship between the hidden layer and the feature transfer between the hidden layers is actually a generalization process. Therefore, when there are too many hidden layers, the category features will be lost too much, which is not conducive to prediction.

*5.3. Simulation Experiment of Window in Convolution Calculation.* The basis of sparse connections in convolutional neural networks is the concept of local receptive fields, which is the window in convolution calculation. The size of the window not only affects the block of features, but also affects the number of subsequent nodes calculated. Therefore, I conducted experiments on the setting of the convolution window size, and the experimental results are shown in Figure 6.

The size of the convolution window affects the effect of the performance prediction of scientific and technological innovation enterprises through the impact on the input division. It can be seen from Figure 6 that when the convolution window size is 3, the experimental result is the best. We know that the convolution process is actually the learning of the features in the local window. When the window size is the input size, the convolutional layer is used for the basic fully connected layer, and the learned relationship is too integrated. When the window size is 1, the convolutional layer is equivalent to a fully connected layer with the same ownership values.

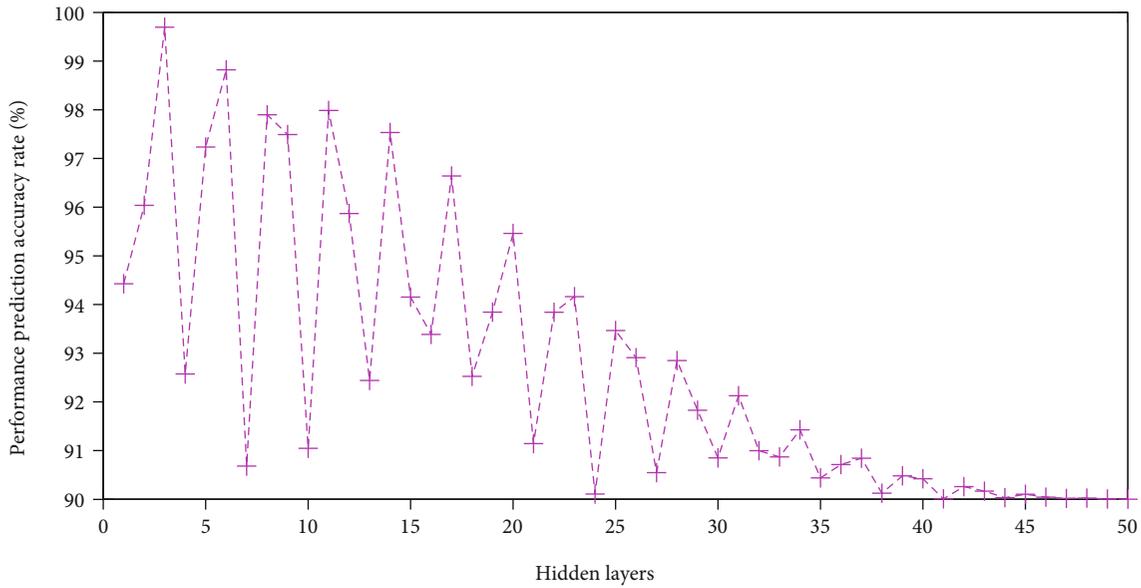


FIGURE 5: The influence of the number of hidden layers in the network on the accuracy of performance prediction.

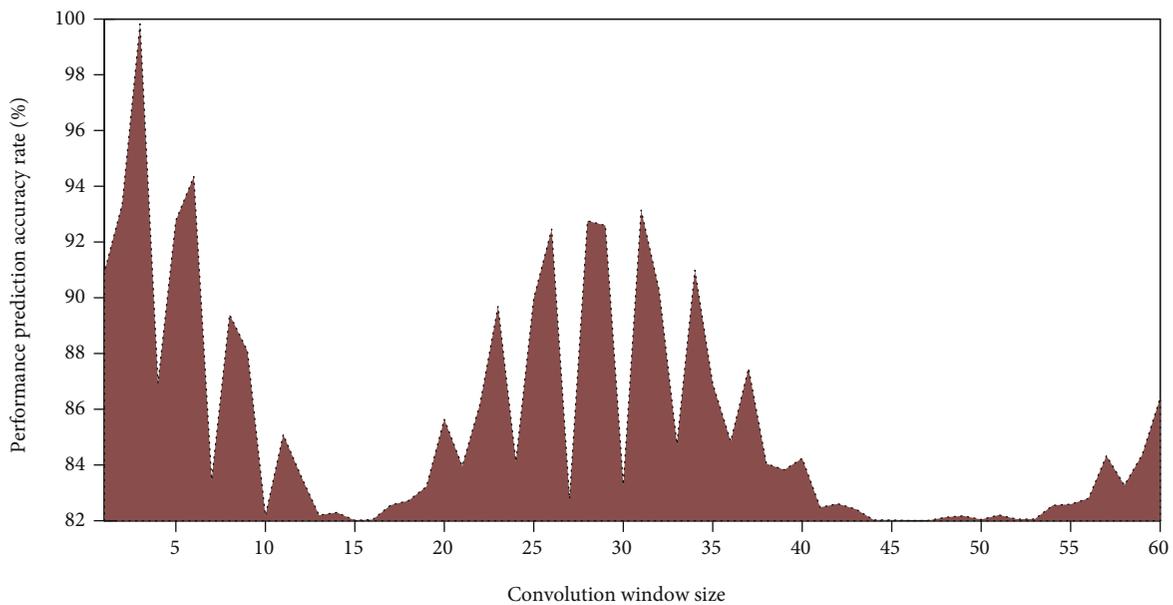


FIGURE 6: The trend graph of the accuracy of performance prediction with the size of the convolution window.

5.4. *Experiments on the Contribution of Different Characteristics to the Performance Prediction of Technological Innovation Enterprises.* I explored the contribution of each type of feature to the performance prediction of technological innovation enterprises, and the experimental results are shown in Figure 7. Removing any type of features will reduce the prediction effect of the experiment. Location features account for 18% of the time, which is in the middle of the contribution of various features.

I also found that whether it is in deep neural networks or convolutional neural networks, the direct use of enterprise difference features will not improve the effect of the experiment, but will have side effects. The reason is that the differ-

ence feature of the enterprise has been processed by the neural network in the extraction, and a vector is used to represent a difference feature of the enterprise, and the other features are the original features in the form of scalar. In the learning process, it does not blend well.

5.5. *Comparative Experiment of Different Models.* I use logistic regression model, support vector regression model, and random forest as comparison methods to compare the prediction effect of the DBN-SVR model. The result is shown in Figure 8.

It can be seen that the DBN-SVR model has the best effect, once again verifying that in the process of feature

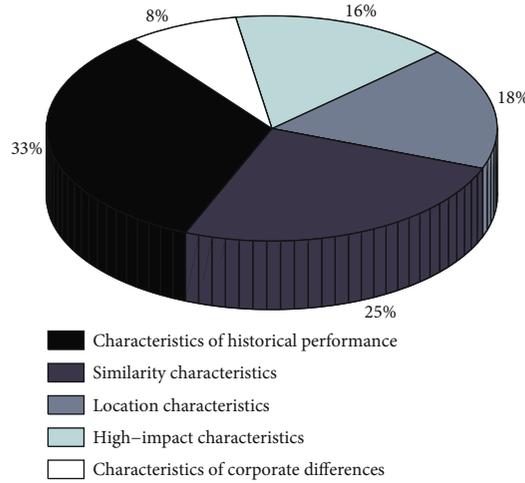


FIGURE 7: Comparison of contribution rates of five types of features.

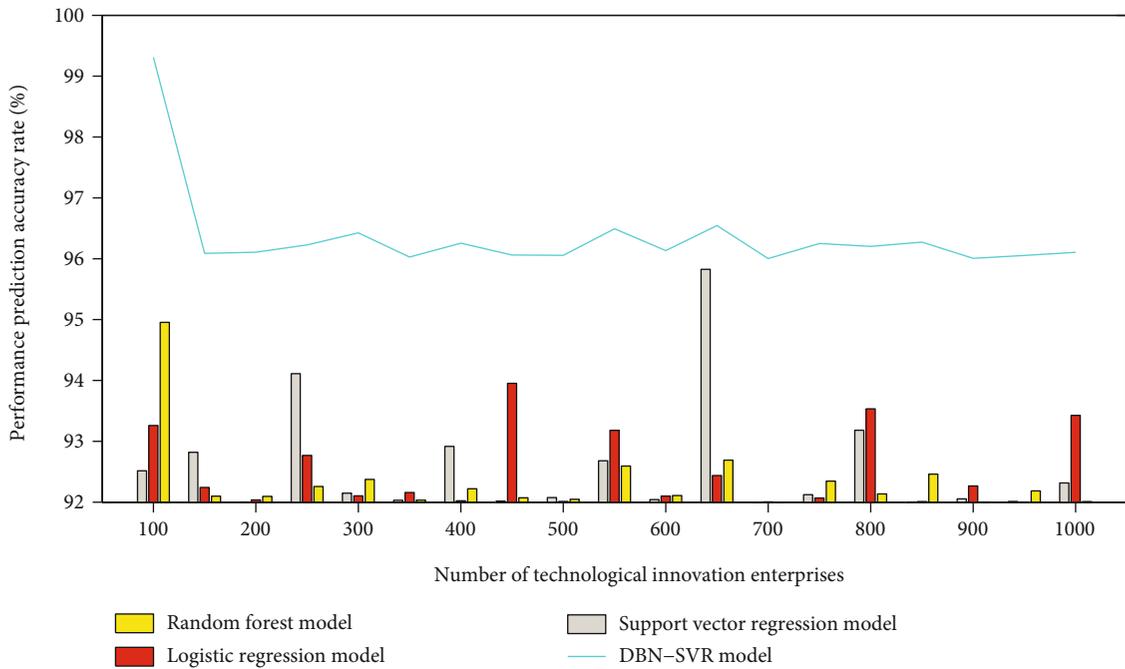


FIGURE 8: Comparison of experimental results of different models.

learning, the deep learning model can better dig the relationship between features than the shallow learning model. This shows that the feature fusion of the convolutional layer and the feature extraction process of the subsampling layer are effective.

### 6. Conclusion

While paying attention to R&D investment, high-tech enterprises should also consider the effect of non-R&D investment on corporate performance. When a company has a variety of different resources to invest together, it should consider how to effectively integrate the resources and use them rationally, optimize the combination of resource inputs, enhance scientific and technological innovation

capabilities, and ultimately achieve the improvement of innovation performance. As some enterprises fail to use resources rationally, leading to the phenomenon of “high input, high consumption, low output, and low added value,” high-tech enterprises should actively try to avoid them. After adjusting and optimizing the investment structure, high-tech enterprises should further improve their technological innovation mechanism, optimize and enhance the quality and efficiency of investment, and effectively boost corporate performance. This paper uses the DBN model to pretrain the large sample data layer by layer and then connects the SVR model to the top of the model to make predictions. In this paper, a large sample data is used for experimental simulation, and the model performance is analyzed by means of mean square error (MSE) and mean absolute percentage

error (MAPE). The influence of the number of nodes in different hidden layers on the prediction error is comparatively analyzed, and multiple prediction models are used to, respectively, predict and compare the prediction effects. Aiming at the problem of slow training time due to the complexity of deep learning models, I designed an experimental method based on GPU computing. The advantages of GPU integration of many computing units make it possible to train deep learning models in the context of big data. In addition, the amount of historical data is very large, and the demand for video memory is high. In the case of limited video memory, I used the idea of block computing to divide the data into blocks for training and testing. Experiments show that within the acceptable range of error, the GPU-based block calculation method can achieve certain experimental results.

### Data Availability

The data used to support the findings of this study are included within the article.

### Conflicts of Interest

The author does not have any possible conflicts of interest.

### References

- [1] H. Jin, X. Chen, L. Wang, K. Yang, and L. Wu, "Dual learning-based online ensemble regression approach for adaptive soft sensor modeling of nonlinear time-varying processes," *Chemometrics and Intelligent Laboratory Systems*, vol. 151, pp. 228–244, 2016.
- [2] H. M. Truong, "Integrating learning styles and adaptive e-learning system: current developments, problems and opportunities," *Computers in Human Behavior*, vol. 55, pp. 1185–1193, 2016.
- [3] X. Yuan, Y. Wang, C. Yang, Z. Ge, Z. Song, and W. Gui, "Weighted linear dynamic system for feature representation and soft sensor application in nonlinear dynamic industrial processes," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 2, pp. 1508–1517, 2018.
- [4] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *International Conference on Machine Learning*, vol. 70, pp. 1126–1135, Sydney, NSW, Australia, 2017.
- [5] X. X. Wang Li and V. C. M. Leung, "Artificial intelligence-based techniques for emerging heterogeneous network: state of the arts opportunities and challenges," *IEEE Access*, vol. 3, pp. 1379–1391, 2015.
- [6] M. Varisco, C. Johnsson, J. Mejvik, M. M. Schiraldi, and L. Zhu, "KPIs for manufacturing operations management: driving the ISO22400 standard towards practical applicability," *IFAC-Papers OnLine*, vol. 51, no. 11, pp. 7–12, 2018.
- [7] Y. Han, Q. Liu, C.-K. Wen, S. Jin, and K.-K. Wong, "FDD massive MIMO based on efficient downlink channel reconstruction," *IEEE Transactions on Communications*, vol. 67, no. 6, pp. 4020–4034, 2019.
- [8] X. Yuan, Z. Ge, B. Huang, and Z. Song, "A probabilistic just-in-time learning framework for soft sensor development with missing data," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 3, pp. 1124–1132, 2017.
- [9] J. Schmidhuber, "Deep learning in neural networks: an overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.
- [10] Y. Yang, F. Gao, X. Ma, and S. Zhang, "Deep learning-based channel estimation for doubly selective fading channels," *IEEE Access*, vol. 7, pp. 36579–36589, 2019.
- [11] T. Vafeiadis, K. I. Diamantaras, G. Sarigiannidis, and K. C. Chatzisavvas, "A comparison of machine learning techniques for customer churn prediction," *Simulation Modelling Practice and Theory*, vol. 55, pp. 1–9, 2015.
- [12] Z. Zhen and Y. Yao, "Optimizing deep learning and neural network to explore enterprise technology innovation model," *Neural Computing and Applications*, vol. 33, no. 2, pp. 755–771, 2021.
- [13] M. L. Borrajo, B. Baroque, E. Corchado, J. Bajo, and J. M. Corchado, "Hybrid neural intelligent system to predict business failure in small-to-medium-size enterprises," *International Journal of Neural Systems*, vol. 21, no. 4, pp. 277–296, 2011.
- [14] A. Abudurehman, A. Nilupaer, and Y. He, "Performance evaluation of enterprises' innovation capacity based on fuzzy system model and convolutional neural network," *Journal of Intelligent & Fuzzy Systems*, pp. 1–9, 2020.
- [15] M. J. Beynon, P. Jones, D. Pickernell, and G. Packham, "Investigating the impact of training influence on employee retention in small and medium enterprises: a regression-type classification and ranking believe simplex analysis on sparse data," *Expert Systems*, vol. 32, no. 1, pp. 141–154, 2015.
- [16] A. Westerski and C. A. Iglesias, "Exploiting structured linked data in enterprise knowledge management systems: an idea management case study," in *2011 IEEE 15th International Enterprise distributed object computing conference workshops*, pp. 395–403, Helsinki, Finland, 2011, August.
- [17] A. M. Pérez, J. M. Gómez, and C. P. Risquet, "Semantic interaction in enterprise data-flow visualization environments: an exploratory study," in *International conference on ICT innovations*, pp. 217–226, Berlin, Heidelberg, 2009, September.
- [18] S. Hassan, H. Waheed, N. R. Aljohani, M. Ali, S. Ventura, and F. Herrera, "Virtual learning environment to predict withdrawal by leveraging deep learning," *International Journal of Intelligence Systems*, vol. 34, no. 8, pp. 1935–1952, 2019.
- [19] F. A. A. Souza and R. Araújo, "Mixture of partial least squares experts and application in prediction settings with multiple operating modes," *Chemometrics and Intelligent Laboratory Systems*, vol. 130, no. 15, pp. 192–202, 2014.
- [20] W. Xing, X. Chen, J. Stein, and M. Marcinkowski, "Temporal predication of dropouts in MOOCs: reaching the low hanging fruit through stacking generalization," *Computers in Human Behavior*, vol. 58, pp. 119–129, 2016.
- [21] H. He, S. Jin, C.-K. Wen, F. Gao, G. Y. Li, and Z. Xu, "Model-driven deep learning for physical layer communications," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 77–83, 2019.
- [22] C.-K. Wen, W.-T. Shih, and S. Jin, "Deep learning for massive MIMO CSI feedback," *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 748–751, 2018.
- [23] J. Zhang, H. Chen, S. Chen, and X. Hong, "An improved mixture of probabilistic PCA for nonlinear data-driven process monitoring," *IEEE transactions on cybernetics*, vol. 49, no. 1, pp. 198–210, 2019.

- [24] J. Zhang, J. Sun, J. Wang, and X. G. Yue, "Visual object tracking based on residual network and cascaded correlation filters," *Journal of ambient intelligence and humanized computing*, pp. 1–14, 2021.
- [25] J. Zhang, X. Jin, J. Sun, J. Wang, and A. K. Sangaiah, "Spatial and semantic convolutional features for robust visual object tracking," *Multimedia Tools and Applications*, vol. 79, no. 21–22, pp. 15095–15115, 2020.
- [26] J. Guo, C.-K. Wen, S. Jin, and G. Y. Li, "Convolutional neural network-based multiple-rate compressive sensing for massive MIMO CSI feedback: design simulation and analysis," *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2827–2840, 2020.
- [27] Z. Ge and X. Chen, "Dynamic probabilistic latent variable model for process data modeling and regression application," *IEEE Transactions on Control Systems Technology*, vol. 27, no. 1, pp. 323–331, 2019.
- [28] K. Wang, M. G. Forbes, B. Gopaluni, J. Chen, and Z. Song, "Systematic development of a new variational autoencoder model based on uncertain data for monitoring nonlinear processes," *IEEE Access*, vol. 7, pp. 22554–22565, 2019.
- [29] X. Yuan, B. Huang, Y. Wang, C. Yang, and W. Gui, "Deep learning-based feature representation and its application for soft sensor modeling with variable-wise weighted SAE," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3225–3243, 2018.
- [30] C. Fan, F. Xiao, and Y. Zhao, "A short-term building cooling load prediction method using deep learning algorithms," *Applied Energy*, vol. 195, pp. 222–233, 2017.