

Retraction

Retracted: Prediction Algorithm of Collaborative Innovation Capability of High-End Equipment Manufacturing Enterprises Based on Random Forest

Mobile Information Systems

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

Copyright © 2023 Mobile Information Systems. 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This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation. The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

 Z. Xiao, J. Shi, R. Tan, and J. Shen, "Prediction Algorithm of Collaborative Innovation Capability of High-End Equipment Manufacturing Enterprises Based on Random Forest," *Mobile Information Systems*, vol. 2021, Article ID 8378274, 6 pages, 2021.



Research Article

Prediction Algorithm of Collaborative Innovation Capability of High-End Equipment Manufacturing Enterprises Based on Random Forest

Zhenhong Xiao,¹ Jianbang Shi ⁽¹⁾,¹ Rui Tan,^{1,2} and Junyi Shen³

¹School of Economics and Management, Harbin Engineering University, Harbin 150001, China
²College of Continuing Education, Beijing University of Technology, Beijing 100124, China
³School of Economics and Management, Harbin University of Science and Technology, Harbin 150001, China

Correspondence should be addressed to Jianbang Shi; shijianbang@hrbeu.edu.cn

Received 22 November 2021; Revised 6 December 2021; Accepted 9 December 2021; Published 21 December 2021

Academic Editor: Ateeq Rehman

Copyright © 2021 Zhenhong Xiao et al. 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 paper studies the competitiveness of listed companies in high-end equipment manufacturing industry by using random forest. Random forest is a supervised machine learning algorithm that is actually based on the regression and classification. It takes some important decisions that are always based upon the set of samples. It counts majority for the classification purposes while it takes an average for the regression. For empirical analysis, 88 listed companies are selected. It is found that there are great differences in comprehensive competitiveness among industries. Enterprise scale accounts for a high proportion in the comprehensive competitiveness, and its score often affects the comprehensive strength; and the gap between companies in the same industry is also obvious. The empirical evaluation results of this paper provide three enlightenments for enterprises to improve their comprehensive competitiveness, such as seizing the strategic opportunity to expand the market, expand the scale of enterprises, improve asset management, and narrow the industry gap.

1. Introduction

With the gradual deepening of the scientific and technological revolution, high-tech equipment such as smart home, driverless, and intelligent robot are filled with all aspects of our daily life. The competition in the high-end equipment manufacturing industry is becoming more and more intense. The report of the 19th National Congress is proposed to accelerate the construction of a manufacturing power and promote the sustainable development of high-end equipment manufacturing industry, which is the core step to enhance the key competitiveness of industry. High-end equipment manufacturing industry is an emerging industry with rapid development at present. The industrial facilities and manufactured equipment use advanced production technology. Moreover, the professional knowledge required is not limited to a certain range, so as to achieve knowledge intensive and technology leadership. The products of this

industry use many high-precision and cutting-edge technologies, so it is the core of China's industrial development industrial chain. As for the definition of the scope of highend equipment manufacturing industry, the Ministry of Industry and Information Technology of China divided eight areas during the 13th Five Year Plan Period: aerospace equipment, marine engineering equipment and high-tech ships, advanced rail transit equipment, high-end CNC machine tools, robot equipment, modern agricultural machinery equipment, high-performance medical machinery, and advanced chemical complete sets of equipment. These eight industries come down in one continuous line with the manufacturing power strategy proposed by [1].

The development of high-end equipment manufacturing industry is an important link that determines the core competitiveness of China's industrial value chain. Its strategic layout reflects the key points of future economic development and scientific progress. Generally speaking, the development of listed companies in high-end equipment manufacturing industry is still very good, and their proportion in industrial production is also gradually increasing, which is the key driving force to accelerate China's highquality economic growth. Through the analysis of the competitiveness of listed companies in high-end equipment manufacturing industry, we can find their advantages and disadvantages and put forward corresponding solutions; promoting the healthy development of high-end equipment manufacturing industry can also provide the corresponding reference for other industries, so as to contribute to accelerating the construction of manufacturing power in China.

The following section discusses the related work and literature review. In Section 3, the proposed methods are described that include random forest, its parameter optimization, and design for organizational collaborative capabilities. In Section 4, the evaluation case study is presented. Finally, the current research study is concluded in Section 5.

2. Related Work

Through in-depth research on Japan's electronics industry, transportation, and other equipment industries, it is found that Japan's equipment manufacturing enterprises mainly adopt the way of cooperation to integrate the technical knowledge inside and outside the enterprise, so as to realize the technological innovation of the enterprise [2]. It is proposed that technological innovation plays a very important role in the equipment manufacturing industry in the United States and Britain, and the cooperation between industry, university, and research institutions established by equipment manufacturing enterprises, universities, and scientific research institutions can effectively improve the overall technical level of the equipment manufacturing industry [3]. It is proposed that in developed countries, the government attaches great importance to the development of domestic decoration manufacturing industry and supports and encourages equipment manufacturing enterprises to improve their level of technological innovation from the perspective of policies, so as to establish the competitive advantage of equipment manufacturing industry [4]. Taking telecom equipment manufacturing enterprises as the object, it is proposed that joining the technical committee of the whole industry can obtain relevant market information advantages, which is very beneficial to formulate and adjust the enterprise's innovation plan and effectively improve the enterprise's innovation level [5]. An empirical study was carried out on 144 equipment manufacturing enterprises in Spain [2]. It is proposed that the technological innovation of enterprises can improve their integration efficiency of innovation resources and play a positive role in improving the performance of enterprises. It is also proposed that the domestic equipment manufacturing industry has not built a systematic industrial chain, so collaborative innovation cannot be realized, which is one of the reasons for the gap between the domestic equipment manufacturing industry and the equipment manufacturing industry in developed countries. The domestic equipment manufacturing industry

should develop in the direction of clustering, horizontal, and vertical integration in the future [3]. It is proposed that improving the ability of technological innovation can effectively strengthen the competitiveness of the equipment manufacturing industry [6]. Through the analysis of the equipment manufacturing industry in Liaoning, he proposed to optimize the innovation capability of the equipment manufacturing industry from the aspects of innovation environment, innovation policy, and innovation platform. This paper analyzes the current demand and environment of domestic equipment manufacturing industry and puts forward that implementing open innovation is an effective way to enhance the innovation ability of equipment manufacturing industry [4]. It is proposed that technological innovation is an important factor affecting the innovation promotion of equipment manufacturing industry, and the innovation level of equipment manufacturing enterprises is enhanced by enhancing the core strength of innovation team [7].

3. Methodology

In this section, we discuss the techniques of random forest. For evaluating enterprise collaborative capabilities, the design of the model based on random forest is explained. Further, parameter optimization of the random forest model is deliberated as follows.

3.1. Random Forest. Random forest fundamentals are that random forest uses the Bootstrap resampling technique by randomly selecting k samples from the original training samples N to generate a random forest, and the final classification is determined by the voting results of the decision trees [3]. The random forest algorithm can be regarded as an integration and improvement of the original decision tree algorithm [6], where a single decision tree achieves node splitting through feature selection to obtain the final prediction category [8], and the random forest systematically combines n decision trees together to determine the category to which the data belong through the majority voting mechanism [7]. The random forest flowchart is shown in Figure 1.

The process of training the random forest is the same as that of training each independent decision tree [9]; the random forest training process is parallel, which mainly depends on how each tree is trained independently, thus building the model faster.

3.2. Design of a Random Forest-Based Model for Evaluating Enterprise Collaborative Capabilities. The random forest model can avoid overfitting on the basis of its good fitting ability [10]. The random forest algorithm can deal with highdimensional data, and the food safety inspection data have a variety of specific failure categories for the same major category of failure indicators alone. In the training process, it is possible to detect the interactions and importance of features, which has certain reference significance [11]. The ID3 and C4.5 decision tree algorithms mine the data set far



FIGURE 1: Random forest flowchart.

as much information as possible, which leads to a complex decision tree model, while the classification and regression tree (CART) algorithm is used for different purposes such as it can simplify the decision tree model by pruning, optimize the decision tree generation efficiency [2], and facilitate the training of each independent decision tree on the divided food inspection sample set, so this study uses the CART algorithm as the base generation construction algorithm for each independent decision tree in the random forest. The CART algorithm uses the Gini coefficient as a measure of feature selection and splitting and constructs a generative decision tree by assuming that the current feature Xi contains *T* categories, and the probability of the *t*th category is P_t . The Gini coefficient can be calculated by the following equation:

Gini =
$$\sum_{t=1}^{T} P_t (1 - P_t) = 1 - \sum_{t=1}^{T} P_t^2$$
, (1)

where Gini represents the Gini coefficient, T represents the number of feature categories, and P_t represents the probability that the food indicator is the current feature. Equation (1) shows that the Gini coefficient is the probability of inconsistency between two samples taken from the data and the category sign of the samples. The smaller the value of Gini coefficient, the higher the classification purity of the model, and therefore the better the classification accuracy of the data [12]. All the feature vectors of the decision tree are traversed, and the feature vector with the smallest Gini coefficient is selected as the splitting feature of the node until the decision tree is constructed, and finally the category of the leaf node is used as the category of the input data.

3.3. Parameter Optimization of Random Forest Model. In order to optimize the hyperparameters in the random forest algorithm and improve the performance and effectiveness of the random forest model, this study uses the hyperparameter grid search algorithm [13] to select a set of optimal hyperparameters for the random forest to avoid manually retesting the model sequentially. In this study, the number of decision trees in the random forest is set to n_estimators which is very small for underfitting and very large for significant model improvement, so the values of n_{esti-} mators are set to [50, 100, 200, 300, 500, 800]; the maximum depth of each decision tree in the random forest is set to max_depth, which is in the range of [2, 3, 4, 5]. The above two parameters are tuned to optimize the whole model parameters. The Grid Search CV function in the sklearn library is used to systematically iterate through multiple parameters and determine the optimal combination of model parameters through 10-fold cross-validation. The data set is divided into 10 parts by 10-fold cross-validation, and 9 of them are used as training data, and the remaining 1 is used as test data for experiments, and a corresponding correct rate is obtained for each experiment.

4. Evaluation Case Study

This section proposes the evaluation model, sample data processing, learning and training of random forest, and establishment of the random forest prediction model.

4.1. Evaluation Model. In Section 3, the idea of constructing the random forest model has been discussed, and based on this, the innovation capability evaluation model diagram of equipment manufacturing enterprises based on random forest is made, as shown in Figure 2.

4.2. Sample Data Processing. In the evaluation system, the evaluation indicators of each innovation ability are different in terms of measurement units, and there is a very large gap between different index values, so there is no comparability between various indicators, which makes the random forest model difficult in the specific input and application of data [14]. In order to solve the problem of incompatibility of each index data, we must carry out standardized processing on the obtained original data when applying the random forest method, so as to ensure the comparability of the sample index data.

In this paper, the proportional conversion method is used to standardize the sample data, and the conversion formula is as follows:

$$T = \frac{X - X_{\min}}{X_{\max} - X_{\min}},\tag{2}$$

where *T* is the processed data (target data), *X* is the raw data, X_{\min} is the minimum raw data, and X_{\max} is the maximum raw data.

4.3. Learning and Training of Random Forest. Through the research on the relevant theories of statistics, the selection of appropriate kernel function is the most key link to ensure the successful mapping of samples from low-dimensional space to two-dimensional space. If different kernel functions are selected, the algorithms applied by random forest will be different in the whole operation process [15]. It is not that the final result can be obtained once the kernel function is determined. Therefore, the parameter selection of kernel function should be fully considered, which will affect the



FIGURE 2: Evaluation model of innovation capability of equipment manufacturing enterprises based on random forest.

final operation result of the random forest model to varying degrees. Considering various factors such as sample size, dimension size, and characteristics of kernel function itself, this paper selects radial basis function, which has a wide convergence domain, and the function has universal applicability in sample size and dimension size.

In the actual training of radial basis function, three very important parameters need to be set in advance, namely: first, penalty coefficient *C*; second, nuclear width σ ; and third, insensitive coefficient ε . These three functions are not fixed. In the whole training, they need to be adjusted appropriately according to the specific situation, so as to ensure the best training effect of the function. The experimental simulation results are shown in Figures 3 to 6.

When the kernel parameter and error both tend to a stable state at the same time, the parameters show the state of optimal combination. At this time, the results are penalty coefficient C = 89 and kernel width $\sigma = 0.97$. Insensitivity coefficient $\varepsilon = 0.002$ with a minimum error of 0.00138.

4.4. Establish Random Forest Prediction Model. In the previous paper, three optimized parameters have been obtained, namely, penalty coefficient C = 89, kernel width $\sigma = 0.97$, and insensitivity coefficient $\varepsilon = 0.002$, and then an intelligent evaluation model is established that can evaluate the innovation ability of equipment manufacturing enterprises. The specific operation process is as follows: when the error is the smallest, get a model and save it in the software. The model is the best prediction model we need. If it is necessary to re-evaluate the innovation capability of a specific enterprise in an industry, just open the trained model that has been successfully saved in matlab2015b software, input the standardized data of equipment manufacturing enterprises into the model, and refer to the evaluation grade table for the final prediction results obtained through the model, and then we can understand the strength of the enterprise's innovation ability.

The processed five sample data are selected and input into the model for prediction. The results are shown in Table 1.

It can be seen from Table 1 that after the data are predicted by the random forest model, the relative error remains within the range of 0.02%, reaching the required







FIGURE 4: Nuclear parameter values and cycle times.

standard of accuracy. Therefore, in the random forest intelligent evaluation model, Gaussian radial basis function is regarded as kernel function to meet the requirements of rationality and feasibility. This shows that on the premise of



FIGURE 5: Insensitivity coefficient and number of cycles.



FIGURE 6: Mean square error and number of cycles.

TABLE 1: Analysis of test sample results.

Sample number	Estimate	Actual value	Error	Evaluation grade
16	0.9218	0.9207	0.0011	Excellent
17	0.8891	0.8898	0.0007	Good
18	0.7469	0.7456	0.0004	Good
19	0.6319	0.6316	0.0003	Like
20	0.5132	0.5123	0.0009	Like

ensuring that the basic data can be collected, any equipment manufacturing enterprise can evaluate the innovation ability of equipment manufacturing enterprises with the help of this model.

5. Conclusion

In this paper, support vector machine and its related principles are briefly introduced. On the premise of considering the complexity of function and data accuracy, an innovation capability evaluation model of equipment manufacturing enterprises based on random forest is established, so that the innovation capability of enterprises can be well evaluated and extended to other applications. With the help of support vector machine theory, the intelligent evaluation model of random forest is established, and the model is studied and trained, so as to obtain the evaluation model of innovation ability of equipment manufacturing enterprises and put forward the necessity of improving equipment from many aspects. There exists some of the relative error but along with that even this proposed evaluation model can function properly.

5.1. Future Work. This work can be extended further up to a satisfactory mark. This idea can be adopted and implemented practically at a major level. This evaluation model can be really helpful in many aspects of life to assess the capabilities of the projects. This assessment shall be helpful in choosing a project that is much better than others in terms of innovation capabilities.

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 there are no conflicts of interest.

References

- S. W. Jang, "Efficient covering of target areas using a location prediction-based algorithm," *The Journal of Supercomputing*, vol. 77, pp. 1–18, 2020.
- [2] R. Lu and M. Lu, "Stock trend prediction algorithm based on deep recurrent neural network," *Wireless Communications* and Mobile Computing, vol. 2021, no. 7, 10 pages, Article ID 5694975, 2021.
- [3] C. S. Shi and D. P. Meng, "Supply risk prediction model of high-tech equipment manufacturing industry based on random forest," *Xi Tong Gong Cheng Yu Dian Zi Ji Shu/Systems Engineering and Electronics*, vol. 32, no. 8, pp. 1667–1671, 2010.
- [4] F. Liu, M. Huang, Q. Yang, and Y. Wang, "Research on sustainable development performance evaluation of China's high end equipment manufacturing enterprises," *IOP Conference Series: Earth and Environmental Science*, vol. 632, no. 5, Article ID 052027, 2021.
- [5] Y. Ai and D. Peng, "Innovation model of China's high-end equipment industry: Do social capital and dynamic capabilities matter for the COVID-19 crisis?" *Frontiers in Public Health*, vol. 9, Article ID 683120, 2021.
- [6] W. Li, Y. Chai, F. Khan et al., "A comprehensive survey on machine learning-based big data analytics for IoT-enabled smart healthcare system," *Mobile Networks and Applications*, vol. 26, pp. 1–19, 2021.
- [7] H. Li, J. Zhao, and F. Yang, "Algorithm of intelligent fault diagnose technology for network of manufacturing enterprise," *Journal of Beijing Jiaotong University*, vol. 41, no. 2, pp. 54–58, 2017.

- [8] W. Peng and C. Shi, "The study on the high-end equipment manufacture enterprise organizational innovation path forming influence factors static impact mechanism," *Journal* of Applied Sciences, vol. 13, no. 22, 2013.
- [9] M. A. Jan, F. Khan, R. Khan et al., "Lightweight mutual authentication and privacy-preservation scheme for intelligent wearable devices in industrial-CPS," *IEEE Transactions* on *Industrial Informatics*, vol. 17, no. 8, pp. 5829–5839, 2020.
- [10] J. Li, J. Cai, F. Khan et al., "A secured framework for sdn-based edge computing in IoT-enabled healthcare system," *IEEE Access*, vol. 8, pp. 135479–135490, 2020.
- [11] R. Skinner, "Character education education week," PS: Preventive Maintenance Monthly, vol. 130, no. 5, pp. 595–599, 2004.
- [12] L. Xin, Z. He, X. You, and C. L. Philip Chen, "A novel joint tracker based on occlusion detection," *Knowledge-Based Systems*, vol. 71, no. 1, pp. 409–418, 2014.
- [13] Y. Lu and X. Xu, "Cloud-based manufacturing equipment and big data analytics to enable on-demand manufacturing services," *Robotics and Computer-Integrated Manufacturing*, vol. 57, pp. 92–102, 2019.
- [14] H. Y. Xing, Q. Zhang, and W. Xu, "Hybrid algorithm for weak signal detection in chaotic sea clutter," *Acta Physica Sinica*, vol. 64, no. 4, pp. 40506–040506, 2015.
- [15] L. Wang, C. Zhang, Q. Chen et al., "A communication strategy of proactive nodes based on loop theorem in wireless sensor networks," in *Proceedings of the 2018 Ninth International Conference on Intelligent Control and Information Processing* (*ICICIP*), pp. 160–167, IEEE, Wanzhou, China, November 2018.