Journal of Combinatorial Optimization (2022) 44:3778–3791
https://doi.org/10.1007/s10878-022-00911-9

Performance prediction and optimization for healthcare enterprises in the context of the COVID-19 pandemic: an intelligent DEA-SVM model

He Huang[1] · Liwei Zhong[2] · Ting Shen[2] · Huixin Wang[2]

Accepted: 15 September 2022 / Published online: 12 October 2022

Abstract
The coronavirus disease (COVID-19) pandemic has caused significant changes in the external environment of enterprises, resulting in tremendous negative impacts. Accordingly, the irregular fluctuation of business data poses a critical challenge to traditional approaches. Therefore, to combat the effects of the COVID-19 pandemic, an effective model is required to proactively predict an enterprise's performance and simultaneously generate scientific performance optimization solutions. Consequently, at the intersection of artificial intelligence algorithms, operations research, and management science, an intelligent DEA-SVM model, which has a theoretical contribution, is developed in this study. The capabilities of this model are verified through sufficient numerical experiments. On the one hand, this model outperforms traditional algorithms in prediction accuracy. On the other hand, effective performance optimization solutions for low-performance enterprises are obtained from the input–output perspective. Moreover, the application value of this model is reflected in its successful implementation in the healthcare industry. Thus, it is a user-friendly tool for realizing the stable operation of enterprises in the context of the COVID-19 pandemic.

Keywords SVM algorithm · Data envelopment analysis · Healthcare management · Performance prediction · Performance optimization

1 Introduction

Since the end of 2019, the coronavirus disease (COVID-19) pandemic has raged worldwide. The pandemic has seriously hindered social and economic development in most

✉ Liwei Zhong
zhongliwei19720311@163.com

1 Business School, University of Shanghai for Science and Technology, Shanghai, China

2 Shanghai Municipal Hospital of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China

countries (Shorfuzzaman 2021). It is undeniable that enterprises have been plagued by many problems such as rising costs, declining revenue, and low operating efficiency. Minimizing the negative impact of the pandemic on enterprises has become a research hotspot (Alraja et al. 2022; Zamani et al. 2022). Hence, investigating the issues of performance prediction and optimization for enterprises in this mutant environment is of highly consequential. Specifically, an effective model is required to predict the performance in advance, and the corresponding performance optimization solution is generated simultaneously.

However, the issues of performance prediction and optimization are more complex in the context of the COVID-19 pandemic. Scholars have studied different techniques, such as statistical methods and artificial intelligence (AI) approaches. In the era of big data, the latter has rapidly developed in recent years. It is well known that the algorithms based on machine learning theory simulate human learning behavior, obtain new knowledge, and reorganize the existing knowledge structure to continuously improve performance (Huang et al. 2021a, b). These AI algorithms not only deal with large-scale data with nonlinear relationships but also obtain more accurate results. However, it is worth noting that reliable data guarantee the performance of AI algorithms. In the context of the COVID-19 pandemic, the business data of enterprises have fluctuated significantly (Lahmiri and Bekiros 2021; Xu et al. 2022), thereby aggravating performance prediction errors. The concept of efficiency from an input–output perspective makes it possible to solve this problem (Jiang et al. 2019). Efficiency is a more stable measure of an enterprise's capabilities. Focusing on the role of multidimensional efficiency, this research proposes a more effective approach by combining the theory of machine learning and operations research. Meanwhile, it is necessary to optimize the performance of low-performance enterprises. A scientific optimization solution would enhance the resistance of enterprises to the COVID-19 pandemic. Therefore, two critical capabilities are involved in the proposed model: performance prediction and performance optimization.

Notably, the healthcare sector has played an irreplaceable role in the fight against COVID-19 (Hossain et al. 2020), including healthcare equipment and services enterprises, as well as pharmaceutical, biotechnology, and life sciences enterprises. For example, with the support of healthcare technology enterprises, China has effectively improved its response to the pandemic by using "health codes." Owing to the efforts of biotechnology and pharmaceutical enterprises, various COVID-19 vaccines have been developed to reduce the risk of human mortality. Therefore, the proposed model is applied to the healthcare industry, representing not only an empirical study but also an avenue to help healthcare enterprises enhance their resistance to the COVID-19 pandemic.

In summary, this research investigates a valuable issue at the intersection of AI algorithms, operations research, and management science, and the main contributions of this research are threefold. First, the theory of machine learning and operations research are combined to develop an intelligent DEA-SVM model with interdisciplinary attributes, which typifies a robust theoretical innovation in combinatorial optimization. Second, the proposed approach conforms to the COVID-19 pandemic environment, and the theoretical innovation thus meets the needs of the times. Third, the proposed approach is successfully applied to the healthcare industry. Through a large volume of real data, the effectiveness and reliability of the approach are verified, and this has a significant application value.

The remainder of this paper is organized as follows: Sect. 2 presents a review of the relevant literature, Sect. 3 proposes the intelligent DEA-SVM model, Sect. 4 explains the data from the healthcare industry, Sect. 5 presents the numerical experiments and analyses, and the conclusions are presented in Sect. 6.

2 Literature review

In the era of big data, enterprise performance management relies heavily on information-driven decision-making, and machine learning is thus a promising source of new techniques (Mashrur et al. 2020; Tsai and Hung 2021). To enhance the prediction accuracy of credit performance, Wang and Ku (2021) propose a parallel artificial neural network (PANNs) ensemble model that creates several independent artificial neural networks (ANNs), each of which deals with the financial performance of the firms for each year, and the final output of the PANNs is aggregated via ensemble learning. Tong and Tong (2022) established a financial early warning system to predict financial operations using a decision tree algorithm. These studies predicted enterprise performance based on different AI techniques. The primary purpose of performance prediction is to screen out enterprises with financial or operational problems. Thus, this issue is a classification problem involving both high- and low-performance enterprises. Therefore, a classification algorithm based on machine learning theory, such as the SVM algorithm, is an effective tool (Huang et al. 2021a, b). To forecast the performance of listed companies, Yan et al. (2020) developed a new framework for a financial early warning system by combining several techniques such as SVM. Meanwhile, Li and Liu (2019) presented a novel method based on the gray kernel AR-SVM model, and the experimental results showed that their method was effective in the analysis of market preference data. Similarly, focusing on the impact of corporate debt financing capabilities and debt financing costs, Liu (2021) established an effective financial crisis early warning model by employing SVM. Considering the excellent classification capability of the SVM algorithm in previous studies, it was also introduced as a classifier in this study.

However, most studies ignore the impact of significant changes in an enterprise's external environment on prediction results. Under the COVID-19 pandemic, there were huge fluctuations in financial and operational data. Thus, this irregular change is a great challenge for traditional prediction techniques. Therefore, a more reliable approach, which can convert original business data into more stable data, is required in the context of the COVID-19 pandemic. Moreover, to enhance resilience to the great changes in the external environment, it is essential to investigate scientific optimization solutions for low-performance enterprises. Consequently, the traditional SVM algorithm needs to be improved. Efficiency is a relatively stable measure of an enterprise's capabilities, such as marketing, research and development (R&D), and operations capabilities (Li et al. 2010). Based on linear programming, DEA is a quantitative analysis approach for evaluating the relative effectiveness of comparable units of the same type (Huang et al. 2019; Soltanifar and Sharafi 2022). With DEA-based models, enterprise organizational capability was studied by Lin et al. (2021), sustainable development capability was measured by Jin et al. (2022), and technological innovation capability was evaluated by Gu et al. (2018). In terms of performance optimization, DEA-based approaches can also generate the optimization solutions for input/output resources. Concerning the financial field, Wu et al. (2022) proposed a novel DEA model, and they found that some samples had insufficient variables that needed to be increased, while others had redundant resources that required to be reduced. The corresponding optimization solutions are discussed further. Zhao and Ge (2020) applied DEA to academic entrepreneurship issues, and Wu et al. (2018) designed a DEA-based performance evaluation system for construction enterprises. From an input–output perspective, their approaches successfully generated solutions for performance optimization. Previous studies have demonstrated that enterprise performance can be optimized by increasing (reducing) insufficient (redundant) resources.

3 Proposed intelligent model

In this section, an intelligent DEA-SVM model is developed for two purposes. One purpose is to accurately predict enterprise performance using fluctuation data during the COVID-19 pandemic, while the other is to generate scientific performance optimization solutions for low-performance enterprises from the perspective of input/output resources. According to the SVM (Du et al. 2021) and DEA (Raayatpanah et al. 2021) theories, the proposed model is developed as follows, and the corresponding schematic is depicted in Fig. 1.

In the context of the COVID-19 pandemic, it is supposed that there are n healthcare enterprises and n decision-making units (DMUs), where DMU_j ($j = 1,2,\ldots,n$). From an input–output perspective, each enterprise has m input resources, x_{ij} ($i = 1,2,\ldots,m$), and s output resources, y_{rj} ($r = 1,2,\ldots,s$). Thus, the relative efficiencies of these n

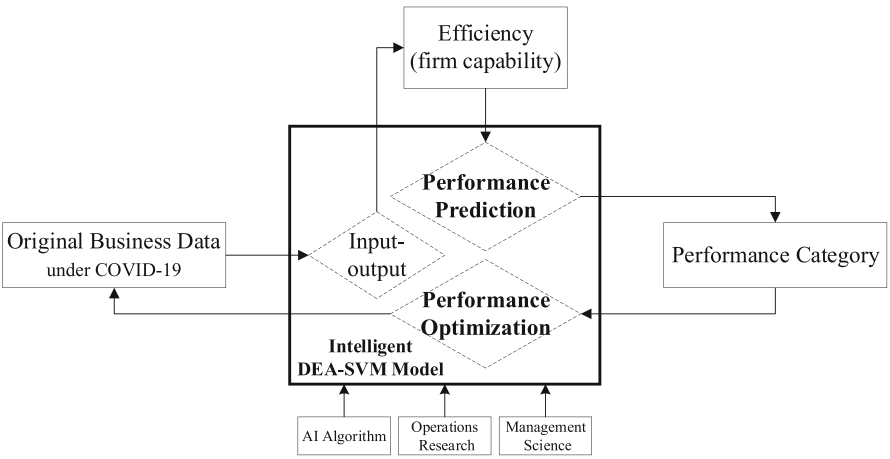

Fig. 1 Schematic of the intelligent DEA-SVM model

DMUs are generated by Eq. 1. Here, ε is the non-Archimedean infinitesimal. The s_i^- and s_r^+ indicate the redundancy/shortage of the input/output variables, respectively. This model is based on the assumption of constant returns to scale, and the efficiency value is thus represented as technical efficiency (TE).

$$\min\theta = \theta_0 - \varepsilon\left(\sum_{r=1}^{s} s_r^+ + \sum_{i=1}^{m} s_i^-\right) \tag{1}$$

$$\text{s.t.}\begin{cases} \sum\limits_{j=1}^{n} \lambda_j x_{ij} + s_i^- = \theta_0 x_{i0} \\ \sum\limits_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{r0} \\ \lambda_j, s_i^-, s_r^+ \geq 0 \\ i = 1, 2, \ldots, m \\ r = 1, 2, \ldots, s \\ j = 1, 2, \ldots, n \end{cases}$$

Under the assumption of variable returns to scale, Eq. 1 is transformed into Eq. 2, and the corresponding efficiency value is pure technical efficiency (PTE). Based on the values of TE and PTE, scale efficiency (SE) is further calculated as SE = TE/PTE.

$$\min\theta = \theta_0 - \varepsilon\left(\sum_{r=1}^{s} s_r^+ + \sum_{i=1}^{m} s_i^-\right) \tag{2}$$

$$\text{s.t.}\begin{cases} \sum\limits_{j=1}^{n} \lambda_j x_{ij} + s_i^- = \theta_0 x_{i0} \\ \sum\limits_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{r0} \\ \sum\limits_{j=1}^{n} \lambda_j = 1 \\ \lambda_j, s_i^-, s_r^+ \geq 0 \\ i = 1, 2, \ldots, m \\ r = 1, 2, \ldots, s \\ j = 1, 2, \ldots, n \end{cases}$$

Consequently, for a non-efficient DMU, the target values of the input/output resources are generated based on s_i^- and s_r^+. The proportion of resources that should be reduced is then defined (see Eq. 3). Here, *pro* is the proportion, D_O is the original data, and D_T is the target data.

$$pro = D_O - D_T / D_O \tag{3}$$

The efficiency values obtained are used as the inputs for the performance prediction. It is well known that an SVM constructs a hyperplane or set of hyperplanes that can be used for classification. In this research, enterprises are divided into high- and low-performance categories after the COVID-19 outbreak. Hence, the first task of the SVM algorithm is to determine the hyperplane that maximizes the margin (w) between the two classes. Here, $w * x + b = 0$ denotes the hyperplane. Thus, the problem is to determine w and b by solving the following objective function using quadratic programming (see Eq. 4).

$$\begin{aligned} &\min_{w,b} \frac{1}{2}||w^2|| \\ &\text{s.t. } y^{(i)}\left(w^T x^{(i)} + b\right) \geq 1 \\ &i = 1, 2, \ldots, n \end{aligned} \tag{4}$$

An ideal SVM analysis should produce a hyperplane that completely separates the vectors (cases) into two nonoverlapping classes. However, perfect separation may not be possible, or it may result in a model with so many cases that it does not classify correctly. Therefore, the slack parameter ξ is introduced to allow some instances to fall into the margin, but they are penalized. This means that parameter C is also introduced to penalize the misclassified instances and those within the margin (see Eq. 5).

$$\begin{aligned} &\min_{w,b,\xi \geq 0} \frac{1}{2} W^T W + C \sum_i \xi_i \\ &\text{s.t. } y_i\left(W^T x_i + b\right) \geq 1 - \xi_i \\ &i = 1, 2, \ldots, n \\ &\xi_i \geq 0 \end{aligned} \tag{5}$$

Using the Lagrange operator, this problem is transformed into Eqs. 6 and 7:

$$L(w, b, \xi, \alpha, \lambda) = \frac{1}{2} W^T W + C \sum_i \xi_i + \sum_i \alpha_i \left(1 - \xi_i - y_i\left(W^T x_i + b\right)\right) - \sum_i \lambda_i \xi_i \tag{6}$$

$$L(\xi, \alpha, \lambda) = \sum\nolimits_i \alpha_i - \frac{1}{2} \sum\nolimits_{i,j} \alpha_i \alpha_j y_i y_j X_i^T X_j \tag{7}$$

Furthermore, the dual form of this problem (Eq. 8) and the Karush–Kuhn–Tucker conditions (Eq. 9) to be satisfied are given.

$$\max_{\alpha} W(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} y^{(i)} y^{(j)} \alpha_i \alpha_j (x^{(i)}, x^{(j)}) \tag{8}$$

$$\begin{aligned} &\text{s.t. } 0 \leq \alpha_i \leq C \\ &i = 1, 2, \ldots, n \end{aligned}$$

$$\sum_{i=1}^{n} \alpha_i y^{(i)} = 0$$

$$\begin{cases} \alpha_i = 0 \ y^{(i)} \left(w^T x^{(i)} + b\right) \geq 1 \\ \alpha_i = C \ y^{(i)}\left(w^T x^{(i)} + b\right) \leq 1 \\ 0 < \alpha_i < C \ y^{(i)}\left(w^T x^{(i)} + b\right) = 1 \end{cases} \tag{9}$$

Moreover, there are situations whereby a nonlinear region can separate groups more efficiently. The kernel function is then employed to map the data into a different space, where a hyperplane (linear) cannot be used to perform the separation. This problem is then transformed into Eq. 10.

$$\begin{aligned} &\max_{\alpha \geq 0} \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j K\left(X_i, X_j\right) \\ &\text{s.t.} \sum_i \alpha_i y_i = 0 \\ &\alpha_i \leq C \\ &i = 1, 2, \ldots, n \end{aligned} \tag{10}$$

The data are mapped into a new space and the inner product of the new vectors is obtained. The image of the inner product of the data is the inner product of the data images. The two kernel functions are shown below.

$$K\left(X_i, X_j\right) = X_i^T X_j \quad \text{(Linear Kernal)}$$

$$K\left(X_i, X_j\right) = \exp\left(-\frac{||X_i - X_j^2||}{2\sigma^2}\right) \quad \text{(RBF Kernal)}$$

4 Experimental data

In this study, the data of the healthcare industry were rigorously collected for numerical experiments, and the variables and datasets in line with the proposed DEA-SVM model were selected scientifically.

4.1 Data acquisition

First, the data of healthcare enterprises were collected through the Wharton Research Data Services (WRDS) database, and the Global Industry Classification Standard (GICS) was employed to screen out healthcare enterprises among various types of enterprises. Healthcare equipment and services enterprises and pharmaceutical, biotechnology, and life sciences enterprises were both considered. Second, to effectively validate the capability of the DEA-SVM model under the COVID-19 pandemic, the enterprises' business data before the outbreak of the COVID-19 (2018), and the enterprises' performance data after the outbreak of the COVID-19 (2020) were collected.

4.2 Variable explanation

Based on the analysis of relevant variables in previous studies, 10 vital variables were selected for this study.

1. Total revenue (million). This variable represents the gross income received from all the divisions of the company. High total revenue is usually desirable for companies. It is the primary output of the firms from the input–output perspective.
2. Cost of goods sold (COGS, million). This variable represents all the costs directly allocated by the company to production, such as material, labor, and overhead.
3. Capital expenditure (million). This variable represents the cash outflow or the funds used for additions to the company's property, plant, and equipment (PPE).
4. Research and development expense (R&D expense, million). This variable represents all the costs related to the development of new products or services, and it is indispensable for analyzing healthcare enterprises.
5. Employees (thousand). This variable represents the number of company workers.
6. Selling, general, and administrative expense (SGA expense, million). This variable represents all the commercial expenses incurred in the regular course of business pertaining to the securing of operating income.
7. Working capital (million). This variable represents the difference between total current assets minus total current liabilities.
8. Total current asset (million). This variable represents the cash and other assets that are expected to be realized in cash or used in revenue generation.
9. Total property, plant, and equipment (Total PPE, million). This variable represents the cost, less accumulated depreciation, of tangible fixed properties used in the production of revenue.
10. Performance category. In the WRDS database, quality ranking is an item of appraisal of firm performance. Based on this item, the performance category variable was defined as two categories. Numbers 1 and 0 represent high and low performances, respectively.

Importantly, four efficiency variables were generated based on these variables. The four efficiency variables comprehensively reflect a firm's capability from different perspectives.

1. Operation Efficiency. Inputs are cost of goods sold and capital expenditures; output is total revenue.
2. Research & development Efficiency. Inputs are R&D expense and employees; output is total revenue.
3. Marketing Efficiency. Inputs are SGA expense and working capital; output is total revenue.

4. Finance Efficiency. Inputs are total current assets and total PPE; output is total revenue.

4.3 Dataset explanation

To improve the reliability and effectiveness of the numerical experiments, an equal number of high- and low-performance enterprises were included in the data samples. In addition, each enterprise has business data for 2018 and performance data for 2020. After removing invalid data, the final sample comprised 120 healthcare enterprises. Among them, there were 60 enterprises each in the high- and low-performance categories. These 120 enterprises were randomly divided into two data groups: training and testing. There were 90 enterprises in the training data group, with 45 high- and 45 low-performance enterprises. For the testing data group, 15 high- and 15 low-performance enterprises were involved.

The descriptive statistical analyses for the business data variables are given in Table 1, and these data are normalized before the numerical experiments using Eq. 11. Here, V_{new} is the normalized data, V is the original data, V_{min} is the minimum value of the corresponding variable, and $Vmax$ is the maximum value of the corresponding variable.

$$V_{new} = \frac{V - V_{min}}{V_{max} - V_{min}} \tag{11}$$

5 Numerical experiments

Numerical experiments are conducted for two purposes: one is to examine whether the proposed model can achieve satisfactory prediction results in the context of the

Table 1 Descriptive statistical analysis for business data variables

Variable	Unit	Maximum	Minimum	Mean	Median
Total revenue	Million	214,319.000	0.430	5860.826	172.232
COGS	Million	202,446.000	0.302	2793.382	68.402
Capital expenditures	Million	3670.000	0.005	192.404	6.295
R&D expense	Million	11,901.000	0.200	614.311	25.585
Employees	Thousand	135.100	0.016	9.890	0.541
SGA expense	Million	33,315.000	1.830	1717.064	104.609
Working capital	Million	25,231.000	0.482	1429.617	136.183
Total current assets	Million	49,926.000	2.415	3349.103	171.142
Total PPE	Million	17,035.000	0.049	1010.600	26.802

COVID-19 pandemic, and the other is to examine whether the proposed model can provide scientific performance optimization solutions for enterprises in advance.

5.1 Performance prediction and analysis

In this section, the proposed intelligent DEA-SVM model is compared with the traditional SVM algorithm. Both approaches utilized the experimental data of 120 healthcare enterprises: 90 training samples and 30 testing samples.

Because the business data fluctuated significantly in the context of the COVID-19 pandemic, the DEA-SVM model intelligently transformed the original business data into efficiency data, which improved the stability of the inputs. Consequently, the DEA-SVM model took the efficiency data of four variables (operations, R&D, marketing, and finance efficiencies) in 2018 as inputs and the performance category data in 2020 as the output. However, the traditional SVM algorithm predicted the performance by utilizing the original business data. Therefore, the inputs were the eight business variables (cost of goods sold, capital expenditure, R&D expense, employees, SGA expense, working capital, total current asset, and total PPE) corresponding to the four efficiency variables, and the output was also the performance category.

To ensure the reliability of the results, experiments were conducted with different cost values ($c = 0.5$, $c = 1$, $c = 5$, $c = 10$) and kernel types (linear and radial basis function (RBF)), which are two key parameters in the SVM algorithm. The SVM type was C-SVC. The prediction accuracy results are shown in Figs. 2 and 3.

From Figs. 2 and 3, it is clear that the intelligent DEA-SVM model outperforms the traditional SVM algorithm under different cost values and kernel types; however, there is a slight drop when $c = 0.5$, and the kernel type is RBF. Second, the results of the linear kernel type are generally better than those of the RBF kernel type. In particular, when the kernel type is linear and $c = 1.00$, the prediction accuracy of the DEA-SVM model reaches 90%, which is 20% higher than that of the traditional SVM algorithm.

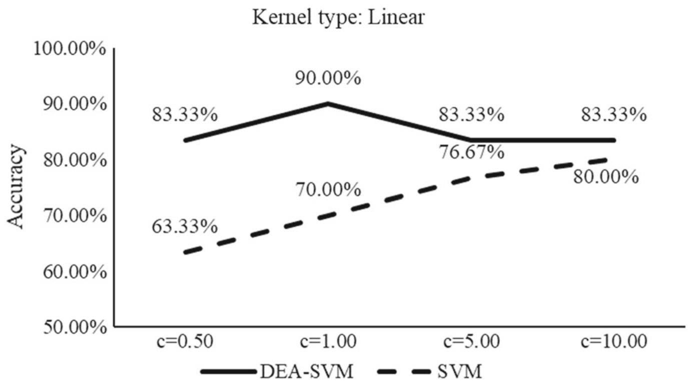

Fig. 2 Comparison of prediction results (kernel type of SVM is linear)

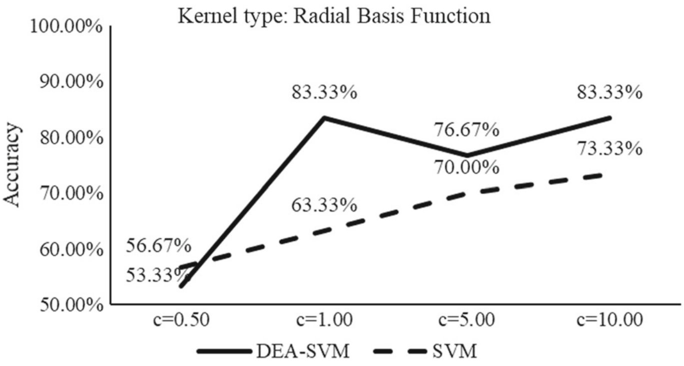

Fig. 3 Comparison of prediction results (kernel type of SVM is RBF)

Owing to the pronounced impact of the COVID-19 pandemic on healthcare enterprises, business data fluctuated irregularly. There would be considerable errors if the original business data were directly used for the performance prediction. The efficiency variable generated from the input–output perspective can steadily reflect an enterprise's capabilities. Therefore, the proposed intelligent DEA-SVM model achieved a better prediction capability by combining the theory of AI algorithms and operations research.

5.2 Performance optimization and analysis

For enterprises with low-performance, it is necessary to optimize their performance in advance to combat the negative impact of the COVID-19 pandemic. Thus, the purpose of this section is to verify how to conceptualize optimization solutions for enterprises with low performance using the DEA-SVM model.

By carefully observing the efficiency data of 120 sample enterprises, it is found that the average efficiency value in the high-performance category is significantly higher than that in the low-performance category, and this result is consistent for three efficiency types: technical, pure technical, and scale. In particular, for the technical efficiency used in the prediction experiment, the average value of the high-performance category is 18% higher than that of the low-performance category.

As a result, the proposed DEA-SVM model further computes the improvement proportion of eight business variables to realize performance optimization, that is, the proportion by which each variable should be reduced. Figure 4 shows the average improvement proportion among the 120 enterprises.

From Fig. 4, it is demonstrated that the proposed DEA-SVM model can generate the proportion of improvement for each business variable. To enhance the healthcare industry's resistance to the COVID-19 pandemic, working capital and R&D expenses, which are the most redundant variables, should be reduced by 94.51% and 90.22%, respectively, while capital expenditure only need to be reduced by 47.01%. The redundancy of employees and total PPE are medium, and they should be reduced by 73.57% and 82.60%, respectively.

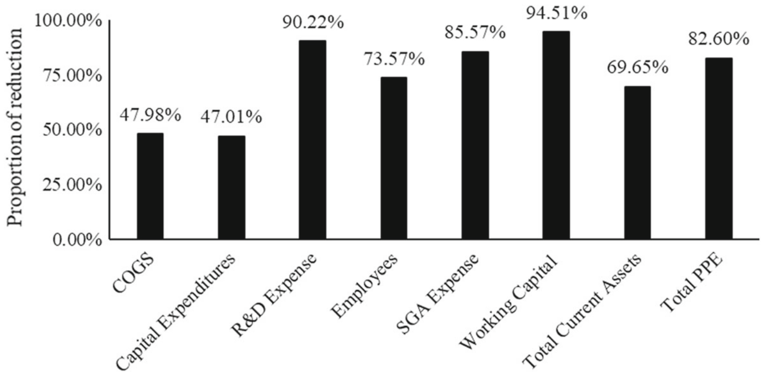

Fig. 4 Performance optimization for all enterprises (average value)

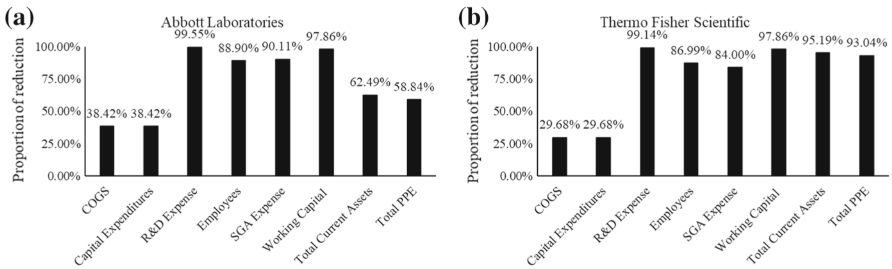

Fig. 5 Performance optimization for two case-study enterprises

However, these results are the average values for 120 sample enterprises, whereas the optimization solutions for each enterprise are different. Some enterprises need to reduce the number of input variables, while others do not. In fact, the DEA-SVM model can simultaneously generate the performance optimization solution for each enterprise, and Fig. 5 shows the optimization solutions for two randomly selected enterprises. Abbott Laboratories is a healthcare equipment and services enterprise in the USA, and Thermo Fisher Scientific is a pharmaceutical, biotechnology, and life sciences enterprise also in the USA. Evidently, the results for the two enterprises are completely different. Thus, each enterprise should optimize its performance according to its unique situation.

6 Conclusion

In the context of the COVID-19 pandemic, this research presented an intelligent DEA-SVM model to predict and optimize the performance of healthcare enterprises. The WRDS database was used to collect business (performance) data before (after) the COVID-19 outbreak. Through a numerical experiment on performance prediction, four efficiency variables that reflect an enterprise's capabilities are generated: operation, R&D, marketing, and finance efficiencies. The prediction accuracy of the DEA-SVM model was then compared with that of the traditional SVM algorithm. By combining various parameters, it was comprehensively verified that the DEA-SVM model had an excellent prediction capability. The prediction accuracy was optimal when the kernel type was linear and the cost parameter was 1.00. In terms of the numerical experiment for performance optimization, the optimization solutions for low-performance enterprises were obtained. Working capital and R&D expense were the most redundant resources for all sample enterprises. For each specific enterprise, the corresponding optimization solution was simultaneously generated. These experiments not only proved the theoretical innovation of the DEA-SVM model but also reflected its application value. In the future, efficiency-related parameters could be improved through additional experiments. It is also important to explore the feasibility of applying this model to other fields.

Acknowledgements This work was supported by National Natural Science Foundation of China [72211530040; 71931008]; Shanghai Shenkang Hospital Development Center (SHDC22021204); Shanghai Municipal Health Commission (202140160); National Key R&D Program of China (2021YFF0900400).

Funding Fundings are disclosed in Acknowledgements.

Data Availability Enquiries about data availability should be directed to the authors.

Declarations

Competing Interests The authors have not disclosed any competing interests.

References

Alraja MN, Imran R, Khashab BM, Shah M (2022) Technological innovation, sustainable green practices and SMEs sustainable performance in times of crisis (COVID-19 pandemic). Inf Syst Front. https://doi.org/10.1007/s10796-022-10250-z

Du Y, Yu H, Li ZJ (2021) Research of SVM ensembles in medical examination scheduling. J Comb Optim 42(4):1042–1052

Gu W, Saaty TL, Wei LR (2018) Evaluating and optimizing technological innovation efficiency of industrial enterprises based on both data and judgments. Int J Inf Technol Decis Mak 17(1):9–43

Hossain MS, Muhammad G, Guizani N (2020) Explainable AI and mass surveillance system-based healthcare framework to combat COVID-I9 like pandemics. IEEE Netw 34(4):126–132

Huang H, Gao W, Ye CM (2021a) An intelligent data-driven model for disease diagnosis based on machine learning theory. J Comb Optim 42(4):884–895

Huang H, Li SL, Yu Y (2019) Evaluation of the allocation performance in a fashion retail chain using data envelopment analysis. J Text Inst 110(6):901–910

Huang H, Shih PC, Zhu YL, Gao W (2021b) An integrated model for medical expense system optimization during diagnosis process based on artificial intelligence algorithm. J Comb Optim. https://doi.org/10.1007/s10878-021-00761-x

Jiang B, Lio WC, Li X (2019) An uncertain DEA model for scale efficiency evaluation. IEEE Trans Fuzzy Syst 27(8):1616–1624

Jin FF, Zhang Y, Garg H, Liu JP, Chen J (2022) Evaluation of small and medium-sized enterprises' sustainable development with hesitant fuzzy linguistic group decision-making method. Appl Intell 52(5):4940–4960

Lahmiri S, Bekiros S (2021) The effect of COVID-19 on long memory in returns and volatility of cryptocurrency and stock markets. Chaos Solitons Fractals 151:111221

Li FT, Liu L (2019) Decision making based on grey model and support vector machine. Clust Comput- J Netw Softw Tools Appl 22(2):S4603–S4609
Li SL, Shang J, Slaughter SA (2010) Why do software firms fail? Capabilities, competitive actions, and firm survival in the software industry from 1995 to 2007. Inf Syst Res 21(3):631–654
Lin C, Ting SF, Lee L, Lin ST (2021) Firm capability assessment via the BSC and DEA. Ind Manag Data Syst 121(3):567–593
Liu BY (2021) Influence of enterprise cloud computing database application and enterprise debt financing based on SVM algorithm: internal control quality analysis. J Ambient Intell Human Comput. https://doi.org/10.1007/s12652-021-03103-1
Mashrur A, Luo W, Zaidi NA, Robles-Kelly A (2020) Machine learning for financial risk management: a survey. IEEE Access 8:203203–203223
Raayatpanah MA, Khodayifar S, Weise T, Pardalos P (2021) A novel approach to subgraph selection with multiple weights on arcs. J Comb Optim 44:242–268
Shorfuzzaman M (2021) IoT-enabled stacked ensemble of deep neural networks for the diagnosis of COVID-19 using chest CT scans. Computing. https://doi.org/10.1007/s00607-021-00971-5
Soltanifar M, Sharafi H (2022) A modified DEA cross efficiency method with negative data and its application in supplier selection. J Comb Optim 43(1):265–296
Tong LL, Tong GL (2022) A novel financial risk early warning strategy based on decision tree algorithm. Sci Program. https://doi.org/10.1155/2022/4648427
Tsai JK, Hung CH (2021) Improving AdaBoost classifier to predict enterprise performance after COVID-19. Mathematics 9(18):2215
Wang MF, Ku HJ (2021) Utilizing historical data for corporate credit rating assessment. Expert Syst Appl 165:113925
Wu PY, Ma J, Guo XY (2022) Efficiency evaluation and influencing factors analysis of fiscal and taxation policies: a method combining DEA-AHP and CD function. Ann Oper Res 309(1):325–345
Wu WD, Ren CX, Wang YH, Liu T, Li L (2018) DEA-based performance evaluation system for construction enterprises based on BIM technology. J Comput Civ Eng 32(2):04017081
Xu WJ, Fu ZN, Li H, Huang JL, Xu W, Luo YY (2022) A study of the impact of COVID-19 on the Chinese stock market based on a new textual multiple ARMA model. Stat Anal Data Min. https://doi.org/10.1002/sam.11582
Yan DW, Chi GT, Lai KK (2020) Financial distress prediction and feature selection in multiple periods by lassoing unconstrained distributed lag non-linear models. Mathematics 8(8):1275
Zamani ED, Griva A, Conboy K (2022) Using business analytics for SME business model transformation under pandemic time pressure. Inf Syst Front. https://doi.org/10.1007/s10796-022-10255-8
Zhao JF, Ge ZF (2020) Analysis of influencing factors of academic entrepreneurship based on blockchain. Wirel Commun Mobile Comput 2020. https://doi.org/10.1155/2020/8825318