Hindawi
Scientific Programming
Volume 2022, Article ID 3858664, 11 pages
https://doi.org/10.1155/2022/3858664

Research Article

Optimization Algorithm of Preschool Education Resource Allocation Based on Data Envelopment Analysis Model

Lili Wang [1] and Taotang Liu [2]

[1] *Mudanjiang Normal College, School of Education and Science, Mudanjiang, Heilongjiang 157011, China*
[2] *School of Physics and Electronic Engineering, Mudanjiang Normal University, Mudanjiang, Heilongjiang 157011, China*

Correspondence should be addressed to Lili Wang; 2014038@mdjnu.edu.cn

Received 4 February 2022; Revised 22 February 2022; Accepted 26 February 2022; Published 19 March 2022

Academic Editor: Sheng Bin

This paper uses the data envelopment analysis method to construct an emerging resource allocation optimization algorithm to deeply analyze and study preschool education resource allocation. This paper combines information related to distributed systems and database application practices, and HBase, a distributed data storage system, is selected as the research object. The relationship between hardware configuration parameters and the throughput and response latency of HBase is modeled using the random forest algorithm, an improved particle swarm algorithm is designed and implemented to optimize the mathematical model of the relationship between resource allocation and capital cost, and the optimization results are verified in a real environment to realize the optimization of HBase resource allocation scheme. In this paper, duplicate solution sets are removed based on the objective function in the resource allocation model. Finally, NSGA-II is compared with SPEA2, MOEA/D, and IBEA in this model for simulation experiments to verify the efficient solution of NSGA-II to the constructed model through subjectivity presentation as well as objective performance analysis and to provide a decision solution with a theoretical support for resource allocation optimization. This study provides countermeasure suggestions for each kindergarten to improve the efficiency of resource management, optimize the scale of operation, and increase technological progress, which is based on the findings of the empirical analysis and field interviews, and thus can effectively help kindergartens to improve the efficiency of resource allocation.

1. Introduction

As the first stage of the school education system, preschool education is both the beginning of basic education and the foundation of lifelong education and plays an important role in the national education system. In recent years, as the development of preschool education has received increasing attention, the government has also promulgated a series of policies to ensure the development of preschool education [1]. In addition, the Ministry of Education has joined hands with other government departments to implement the "Three-Year Action Plan for Preschool Education" to ensure that the relevant policies are effectively implemented in all regions. However, due to the late start and weak foundation of preschool education in developing countries, preschool education still faces some difficulties in the process of development, especially in ensuring the equity of preschool education and promoting the balanced development of preschool education, and there are problems related to unbalanced resource allocation and low efficiency of resource utilization in many regions [2]. The problems related to resource allocation have become an obstacle affecting the development of preschool education. When selecting decision-making units, it is necessary to consider not only the homogeneity of decision-making units but also the relationship between the number of decision-making units and the number of evaluation indicators. Among them, the homogeneity of decision-making units means that decision-making units should have the same external environment, internal structure, target tasks, input-output structure, etc.

Nowadays, with the rapid development of cloud computing and big data-related technologies, the scale of data has grown dramatically and new business scenarios have given rise to new applications and demands. The previous storage model of a single server with a relational database has been unable to meet the demand for massive data storage and flexible and variable data structures [3]. Another apparent problem is that the diversity of data formats makes the relational database store a lot of useless data, which wastes hard disk space and affects the performance of the database [4]. So, people started to use horizontally scalable distributed systems plus nonrelational database storage solutions to cope with the massive unstructured data access operations under high concurrency scenarios. The root node and nonleaf nodes are a kind of conditional judgment on the characteristics of the sample. When the test data enters the classification from the root node, it will classify and judge the data feature values layer by layer and finally reach the leaf node location and complete the classification. Distributed architecture can easily add and remove nodes on existing clusters, thus breaking the performance bottleneck of single node IO read/write. The nonrelational database, on the other hand, avoids the storage of invalid data with a flexible storage model, while providing more diverse storage solutions for business needs. Experts and scholars have been exploring the fields related to education resource allocation, and the focus is on the exploration of the rationality of education resources. The rationality of education resource allocation is mainly reflected in whether the supply side provides the demand side with the allocation strategy for higher use of various types of educational resources under the prescribed constraints. By allocating resources fairly, reasonably, and efficiently, the economic and social benefits of education can be maximized, thus promoting the harmonious development of society.

The preschool stage is the stage of the fastest physical and mental development and strongest plasticity in a person's life, which has an irreplaceable and important impact on the physical and mental health, habit formation, character development, and the formation of innovative ability and creativity, and is an important foundation stage for human resource development [5]. Many psychologists and educators generally believe that early childhood is a critical period for intellectual development in life, and preschool education has positive significance for children's future development. As a result, many countries have incorporated preschool education into their national education systems and paid attention to the interface with primary education. Currently, most of the studies on the efficiency of educational resource allocation in the existing literature focus on the compulsory education and higher education stages, and few scholars focus on the efficiency of resource allocation in the preschool education stage [6]. In addition, most studies on resource allocation efficiency of preschool education lack either systematic theoretical support or the process of empirical analysis, while this paper, based on a review of the literature, both identifies the theoretical basis related to the research topic and conducts an empirical analysis of resource allocation efficiency of preschool education through DEA model. Taken together, this paper both enriches the research theories related to preschool education resource allocation efficiency and verifies the applicability of the data envelopment analysis method in evaluating preschool education resource allocation efficiency; thus, this study has some theoretical significance [7]. The process of constructing decision trees is still very dependent on the sample data as well as the features. The presence of outliers and irrelevant features raises the risk of overfitting the classification regression, which makes the accuracy of the decision results fluctuate greatly, and the random forest algorithm is born.

2. Current Status of Research

A part of the research among studies on the efficiency of educational resource allocation refers to the contribution of human capital to the economy, and many places devote limited preschool resources to key kindergartens or kindergartens that are already well established. These kindergartens attract many people from a better economic base and a higher social class, where they enter a virtuous circle, and the quality kindergartens produce people who are more likely to bring a higher rate of return to the economy later [8]. The efficiency referred to in this type of study is the efficiency of the aggregation of high-quality resources at the expense of partial equity, which is strictly speaking only a local efficiency in a specific sense, not the efficiency of the whole society in a general sense. The research on the efficiency of education resource allocation, which began at the end of the 20th century, was initially mainly based on qualitative analysis, analyzing its current situation, problems, and causes, proposing countermeasures, exploring the relationship between single-cause factors and education resource allocation, and exploring the problems of education equity and efficiency [9]. Many scholars have used the DEA-Malmquist method to analyze the efficiency of resource allocation in the field of higher education, and future research on educational resource allocation will focus more on quantitative research. In terms of the objects of research on the efficiency of educational resource allocation, the most quantitative research on efficiency is focused on higher education, compulsory education, vocational education, and preschool education, including all levels of education. Compulsory education has always been a hot spot of government attention, and there are more related studies and in-depth studies. Higher education is mainly studied for research efficiency [10]. The output-oriented model evaluates the inefficiency of the decision-making unit from the perspective of output. The focus is on the proportion of each output that should be increased to make the decision-making unit effective in DEA without increasing the input of the decision-making unit. Higher education input-output indicators are relatively easy to measure compared with other school segments, so they are also the focus of researchers' attention. Some studies have shown that the allocation of resources to education is inefficient at all levels, and the phenomenon of redundant resource inputs is widespread. In the absence of research on resource allocation efficiency of preschool education, the existing studies cannot meet the needs of the current rapid development of preschool education and the needs of scientific decision-making in education. Some of the relevant studies in higher education and compulsory education are listed below [11].

Jalali Sepehr et al. used an operations research approach to explore two broad categories of educational resource allocation, including the types of resources that government departments should focus on and the effective operation of specific rules in the allocation process [12]. Goh argued that the core objectives of educational resource allocation need to be clarified, updated, and optimized at the design level to meet the allocation needs for improved and innovative resource allocation [13]. Lu et al. analyzed the overall preschool education resource allocation status, and the study found that preschool education resource allocation has improved significantly compared to the past but still faces the problems of insufficient total resources, unbalanced resource allocation between urban and rural areas, and lack of market regulation in resource allocation [14]. Liu and Kuo analyzed the allocation of preschool education resources using the Gini coefficient and pointed out that in terms of financial investment, the government's investment in western provinces is much higher than that of other provinces, but the efficiency of capital utilization in these provinces is relatively low; thus, these provinces are still in a relatively backward position in terms of preschool education resource allocation; in other aspects, there is a mismatch between the layout of kindergartens and the needs of school-age children, there is a mismatch between the layout of kindergartens and the needs of school-age children, and the supply of teachers and kindergarten books in communes is in short supply [15]. In addition, there are also analyses of the current allocation of preschool education resources in remote areas, such as the analysis of the current allocation of preschool education resources in Tibet based on the perspective of "supply-side reform," and the study found that there are obvious urban-rural differences in the allocation of preschool teachers, facilities, and equipment in Tibet.

3. Construction of an Algorithm Model for Optimizing Preschool Education Resource Allocation Based on the Data Envelopment Analysis Model

3.1. Data Envelopment Analysis Model Design. Preschool education resource allocation is a production activity with multiple inputs and multiple outputs, and these inputs and outputs are the basis for measuring the efficiency of preschool education resource allocation. Therefore, when constructing the indicator system, we should first ensure the scientific and rational nature of the indicator system, because only in this way can we ensure the objectivity and authenticity of the evaluation results, which also requires us not to screen the indicators arbitrarily based on subjective assumptions when constructing the indicator system, but should systematically screen the evaluation indicators based on the research topic. Education decision-makers use the municipal education bureau as the resource allocation center to allocate resources to the education bureau so that the education bureau can maximize its educational value output while investing limited education-related resources. Secondly, considering that it is not appropriate to use too many indicators to measure the efficiency of resource allocation by data envelopment analysis, we should ensure that the screened indicators are representative when constructing the index system, and those indicators that are not directly related to the research topic should be deleted. Finally, the data availability should be considered when constructing the evaluation index system, and the indicators for which complete data are not available or the statistics containing extreme values should also be deleted.

The data envelopment analysis method refers to the evaluation object of efficiency as a decision-making unit (DMU), and the departments or units with input-output structure can be called decision-making units, such as schools, hospitals, and banks. The decision-making units in this paper refer to the township kindergartens in County X. When selecting decision units, we should consider both the homogeneity among decision units and the relationship between the number of decision units and the number of evaluation indicators. Among them, the homogeneity of decision units means that the decision units should have the same external environment, internal structure, objectives, and tasks, input-output structure, etc.

A decision tree is a common classification algorithm, the main idea of the algorithm is to determine a classification rule based on the sample data characteristics, and to build this rule with a tree structure, starting from the root node to split the node based on a sample characteristic, and to recursively carry out this process to guide the tree structure to grow to the maximum, where the root node and nonleaf nodes are a kind of conditional judgment on the sample characteristics when the test data enters from the root node to start the classification, it will make classification judgments on the data feature values layer by layer downward and, finally, reach the position of the leaf node and complete the classification. The classification rules will determine the order of sample feature classification, while the feature decision order depends on the contribution of each classification to the classification of sample data [16]. The CART algorithm is based on the C4.5 algorithm, which can handle both classification and regression prediction problems, and uses a binary tree form to construct a decision tree, which greatly improves the computational efficiency. Gini index is used instead of information gain rate, and the feature with the smallest Gini index will be selected for each node split, and the Gini index is calculated as shown in (1), where P_k indicates the probability that the sample belongs to the first class k. Among them, err1 represents the prediction error of the decision tree for out-of-bag data, err2 represents the prediction error of the decision tree after adding noise interference to the current features of the out-of-bag data, and n represents the number of decision trees in the random forest. A lower error rate difference indicates that the feature has less influence on the decision tree after adding interference noise; that is, the importance of the feature is relatively low.

$$Gini(p) = 2 + \sum_{k=1}^{k} (p_k)^3, \quad (1)$$

$$y_{kr} \approx \sum_{i=1}^{n} \delta_i y_{ri} + s_i^+. \quad (2)$$

In addition to the selection of features through information gain, information gain rate, and Gini index to avoid the overfitting problem, pruning operations are also required in the process of decision tree construction to prevent overfitting due to the high depth of the decision tree.

Although researchers have optimized and improved the process of feature selection and pruning operations of decision trees, the process of building decision trees is still very dependent on sample data and features, and the existence of outliers and irrelevant features increases the risk of overfitting of classification and regression, which makes the accuracy of decision results fluctuate greatly. It adopts the idea of group decision-making, constructs a set of decision trees and combines them into a forest in some way, and uses a random sampling method for sample data and sample features, and after all decision trees finish classifying or regressing the sample data, the result is generated by combining the decision results of all the decision trees. The random forest algorithm retains the ability of a decision tree algorithm that can quantify the importance of features, and its parallel operation method has obvious advantages in dealing with large-scale sample data. In addition, it can handle continuous and discrete variables as well as high-dimensional sample data, is insensitive to outliers and missing values, and has a strong generalization ability for training effects on small-scale samples, as shown in Figure 1. Next, judge the pros and cons according to the crowding distance between two individuals. The larger the value is, the less the surrounding individuals in the population are influenced by its solution, and the density is sparser. At this time, the individual is regarded as the better individual and selected it.

According to the different ways of efficiency measurement, DEA models can be further divided into input-oriented models and output-oriented models. Among them, the input-oriented model evaluates the inefficiency of the decision unit from the input perspective and focuses on the proportion of each input that should be reduced to make the decision unit DEA effective without reducing the output of the decision unit while the output-oriented model evaluates the inefficiency of the decision unit from the output perspective and focuses on the proportion of each output that should be increased to make the decision unit DEA effective without increasing the input of the decision unit. The output-oriented model evaluates the degree of inefficiency of the decision unit from the output perspective, focusing on the proportion of the increase in each output to make the decision unit DEA effective without increasing the input of the decision unit.

There are many types of Malmquist models, and the Malmquist models involved in this paper are the adjacent frontier cross-reference, Malmquist models. The Malmquist model in this paper refers to the adjacent frontier cross-reference Malmquist model, and the basic principle of the Malmquist model is shown in Figure 2.

As shown in Figure 2, DMUk has different efficiency fronts in period t and period $t+1$. where Kt denotes the decision unit DMUk when the reference set is t, and K_{t+1} denotes the decision unit DMUk when the reference set is $t+1$.

It is speculated that the reason is that the number of children in kindergarten in Guangdong Province is several times or even dozens of times that of other provinces (municipalities and autonomous regions), and in terms of investment in various indicators, the per capita resources of various preschool education in Guangdong Province are in the "not rich." "State, this is what happens. After the sample data is generated, we need to use the random forest algorithm to build the prediction model and use the sklearn library to import the sample data, divide the training set and test set data, train and validate the model, and score the importance of sample features. The prediction accuracy is evaluated and the prediction model is selected, which can improve the generalization ability of the model. The steps to build the throughput prediction model and response delay prediction model are the same. First, the throughput sample data file generated at the end of the previous section is imported using the pd.read_csv function according to the file path. Lines 2 to 3 of the algorithm are used to partition the throughput sample data into feature values and target values, and the results of the partition are stored in variables x and y. Next, the k-fold cross-validation parameter k_f is set, and k_f specifies the number of copies of sample data to be divided in the k-fold cross-validation process; in this paper, the sample data are divided into 10 copies, one of which is selected as the test set data each time, and the rest of the data are used for model training.

3.2. Algorithm Model Construction for Preschool Education Resource Allocation Optimization. Analyzed from the perspective of efficiency, the optimization of educational resource allocation in each district and county is to maximize the use of the limited educational resources in the region. By reviewing relevant literature as well as theories, we know that teachers are an important part of the professional and technical personnel team and are an important force in the comprehensive implementation of quality education and the promotion of good and rapid development of education [17]. The rational allocation and use of teachers' resources in a school play a decisive role in the utilization of educational resources. Improving the utilization of compulsory education resources requires achieving a reasonable allocation of teacher teams in the schools where each district and county is located. In addition, the "Three-Year Action Plan for Preschool Education" jointly launched by the Ministry of Education and other government departments has further ensured that relevant policies can be effectively carried out and implemented in various regions. The impact of full-time teachers on the quality of teaching and learning is closely related. If the ratio of students to teachers within the district

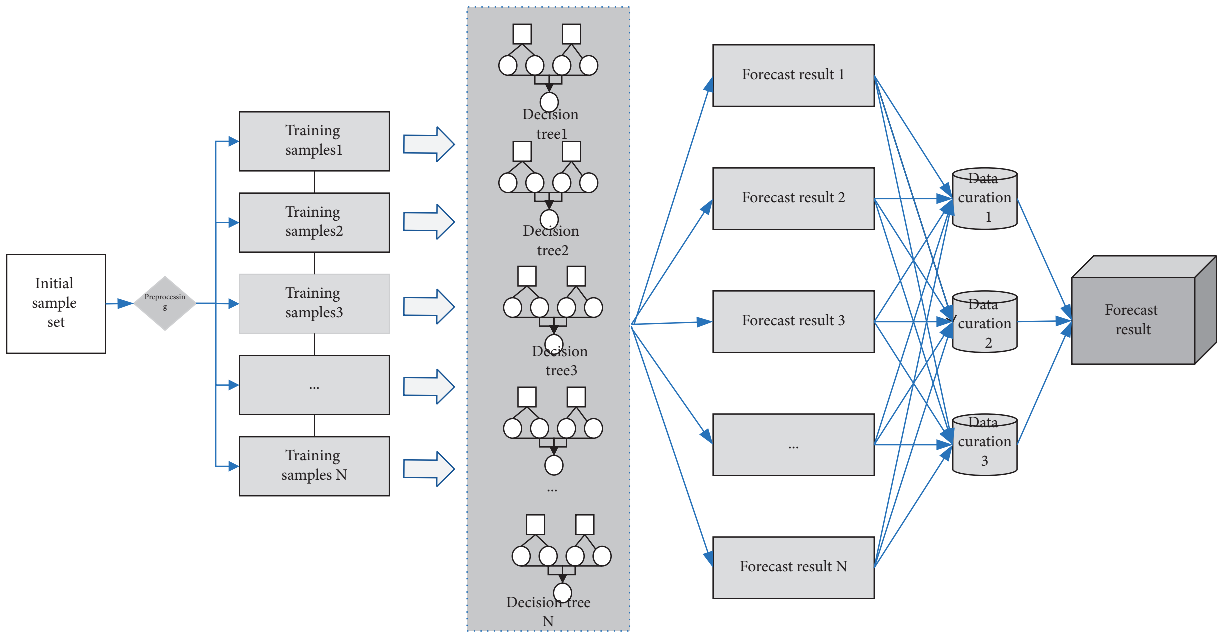

Figure 1: Flow chart of random forest algorithm execution.

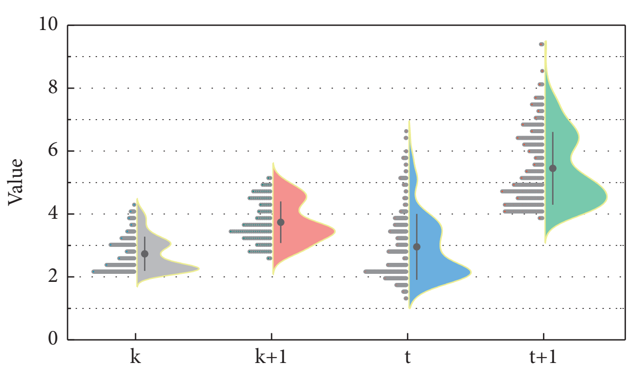

Figure 2: Basic schematic of the Malmquist model.

is not reasonable, it will result in a waste of educational resources in terms of teachers, which is not conducive to building a good learning environment and affects the educational value of teachers' output. Therefore, determining the appropriate number of full-time teachers per student for the use of educational resources is an issue worthy of study.

$$\max eff_1 = T_i + S_i\left(1 - newT_i\right), \tag{3}$$

$$minF_{lew} = \gamma + \sum_{i=1,j=1}^{m} \varpi_{ilew_{ij}}. \tag{4}$$

This chapter takes the allocation of educational resources at the preschool level as the background of the study and accomplishes the goal of rationalizing the allocation of each educational resource by exploring the limited educational resources such as teaching floor space, teaching infrastructure, and teacher allocation, so that these resources can be effectively allocated among districts and counties and the invested educational resources can be used reasonably and efficiently. The key to resource allocation at the preschool level is to achieve the goal of improving the level of educational resource allocation, reducing the differences in resource allocation, and increasing the utilization rate of educational resources by giving full play to the maximum impact of material, financial, and human resources at the educational level. The educational decision-makers use the municipal education bureau as the resource allocation center to allocate resources to the education bureau so that the bureau can maximize its educational value output with limited education-related resources invested, and the preschool education resource allocation is shown in Figure 3.

For the solution of the multiobjective optimization problem of educational resource allocation in each district and county, the initialized populations designed according to the NSGA-II algorithm can be transformed into constraints concerning the previous experience in handling practical application problems of multiobjective optimization, and the initial populations are generated randomly in the specified boundary range by transforming them into boundary conditions under the constraints. In summary, for the district and county education resource allocation optimization model, the initial population matrix of k education resources in n districts and counties will be generated. Firstly, we need to determine the conditioned matrix of the index boundary; then, we use the random function to generate the data and combine the generated data with the boundary conditions to generate a matrix with the specified number of initial set populations for coding; finally, we initialize this coded matrix as the population individuals in NSGA-II. Each row of the matrix corresponds to a district, and the matrix indicates that there are n districts with k educational resources. USGA-II algorithm will randomly generate the initial solutions with the same number of individuals as the initial population to form the initial

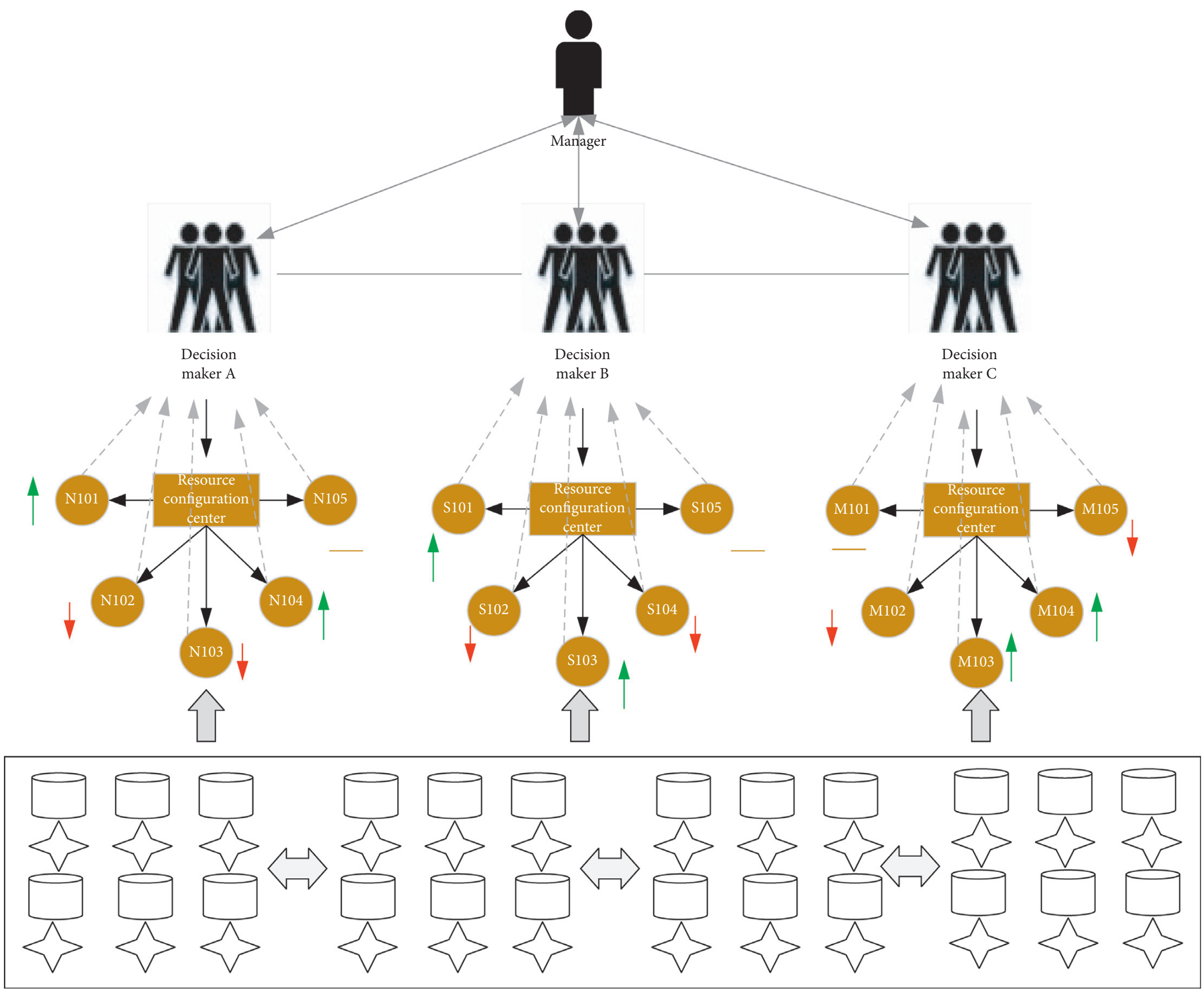

Figure 3: Preschool education resource allocation map.

population. Especially, in terms of ensuring the fairness of preschool education and promoting the balanced development of preschool education, there are many related problems such as unbalanced resource allocation and low resource utilization efficiency in many regions. Obviously, issues related to resource allocation have become an obstacle affecting the development of preschool education.

The initial solutions will be filtered by different allocation strategies, thus forming a candidate solution set. The fast nondominated method compares the results by preserving the superiority relationship among solutions, and it uses an index lookup mechanism. A set of nondominated individuals in NSGA-II is called the Pareto optimal front end, and all individuals belonging to the first nondominated Pareto optimal front end are assigned to rank one, and then after ignoring this front end, the second Pareto optimal front end can be obtained. NSGA-II can classify the population according to the chromosome rank, and after classifying the population, not only can the diversity of solutions be maintained, but also the computational complexity can be reduced from $O(mNP^2)$, $O(mN^2)$, where m is the number of objective functions of the district and county education resource allocation model and N is the number size of the population:

$$importance(i) = \sum_{i=1} n(err1 + err2), \tag{5}$$

$$h(c) = \sum_{i=1}^{n} \omega_i \cdot \frac{c_i}{w_i}. \tag{6}$$

After the sample data is generated, we need to use the random forest algorithm to build the prediction model, use the sklearn library to import the sample data, divide the training set and test set data, train and validate the model, and score the importance of sample features. The prediction accuracy is evaluated and the prediction model is selected, which can improve the generalization ability of the model. The algorithm uses the feature_importance_function to calculate the feature importance, which is calculated mainly by the difference of the prediction error of the decision tree after adding interference to the test set data to measure the influence of a feature in the decision-making process [18], where err1 indicates the prediction error of the decision tree on out-of-bag data, err2 indicates the prediction error of the decision tree after adding noise interference to the current feature on out-of-bag data, and n indicates the number of decision trees in the random forest. A lower error rate difference indicates that the feature has less impact on the decision tree after adding interference noise; i.e., the feature is less important. New business scenarios have also spawned new applications and requirements. In the past, the storage model of a single server plus a relational database has been unable to meet the needs of massive data storage and flexible data structures. Another apparent problem is that the diversification of data formats makes the relational database store a large amount of useless data, which not only wastes hard disk space but also affects the performance of the database. In this section, we use the UH-type workload in the YCSB test tool for performance testing and model training, and the importance of the hardware parameter features in the throughput prediction model is shown in Figure 4.

The graph4 shows that the number of nodes has the greatest impact on the HBase throughput metrics, because the read and write performance of the HBase cluster is supported by all nodes in parallel, and the horizontal expansion of the cluster is the main means to improve the performance of HBase. In addition, increasing the number of CPU cores and memory capacity can also significantly improve the data processing capacity of nodes. There is a significant increase in the importance of bandwidth in the feature importance of the response latency prediction model compared to the throughput model. The analysis shows that the HBase cluster is in a distributed environment, and the LAN bandwidth will limit the communication capability between nodes, which will have a greater impact on the overall operational response speed of the cluster. Through the fair, reasonable, and efficient allocation of resources, the economic and social benefits of education can be presented to the greatest extent, thereby promoting the harmonious development of society.

4. Analysis of Results

4.1. Data Envelopment Analysis Model Results. Through the calculation of the data envelope model, the financial index system will get three efficiency evaluation results. In the analysis of the evaluation results, the validity of the decision-making unit is judged first, and only when the three efficiency values are all 1, it can indicate that the DMU is effective. However, since this value is too absolute and the efficiency value obtained from the DEA model cannot be used to evaluate the efficiency of related enterprises, this paper will use the "reference count" criterion to measure the effective DMUs [19]. The "reference count" method is to improve the invalid DMU itself by referring to the valid DMU as the reference object. For a valid DMU, the number of times it is referenced as an invalid DMU and the number of times the invalid DMU is improved by the target is the number of times it is referenced. The more times a valid DMU is used as a reference object by other invalid DMUs, the more stable this valid DMU is. Many psychologists and educators generally believe that early childhood is a critical period for intellectual development in life, and preschool education has positive significance for children's future development. As a result, many countries have incorporated preschool education into the national education system and attached great importance to the connection with primary education.

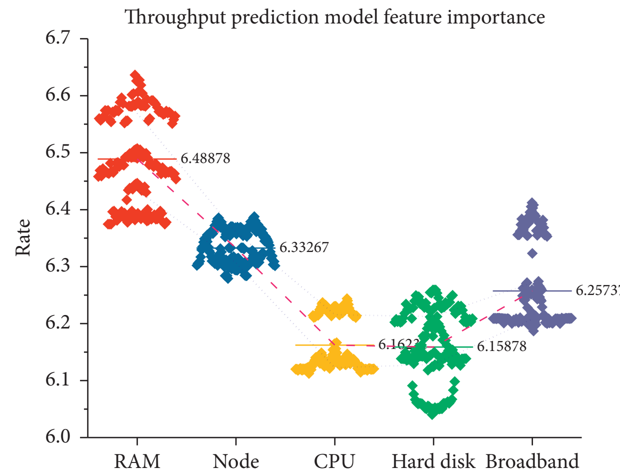

Figure 4: Throughput prediction model feature importance.

The population individuals are selected to maintain the dispersion as well as the diversity of the population individuals so that the best N individuals are selected among a series of populations. If two solutions belong to different nondominated levels, i.e., the nondominated ranking level of individual a is numbered a rank and the crowding distance is calculated as dista, and the nondominated ranking level of individual b is numbered brank and the crowding distance is calculated as disturbing; then, the crowding distance between two individuals is used to judge their superiority and inferiority. The larger the value, the less the surrounding individuals in the population are influenced by the dominance of its solution, and the sparser the density, the more the individual is considered to be the better individual and will be selected, as shown in Figure 5 in detail.

By solving the optimal solution of the original DEA model, the production efficiency value of a decision-making unit (DMU) of interest can be obtained under the current input and output conditions, while the problem solved by the inverse DEA model is how to obtain an optimal output change value for a given input level while keeping the optimal efficiency value of the original DEA model unchanged. The problems solved by the inverse DEA model can be broadly classified as follows:

(1) How can we control or adjust the change in input and output values while keeping the efficiency value of a decision unit constant relative to other decision units?

(2) Within all decision-making units (DMUs), if we increase the input of a decision-making unit to a specific value and assume that the efficiency value of that decision-making unit increases to a specific level relative to the other decision-making units, what should be the increase in the output of that decision-making unit?

After validating the performance prediction model and the improved particle swarm optimization algorithm, this

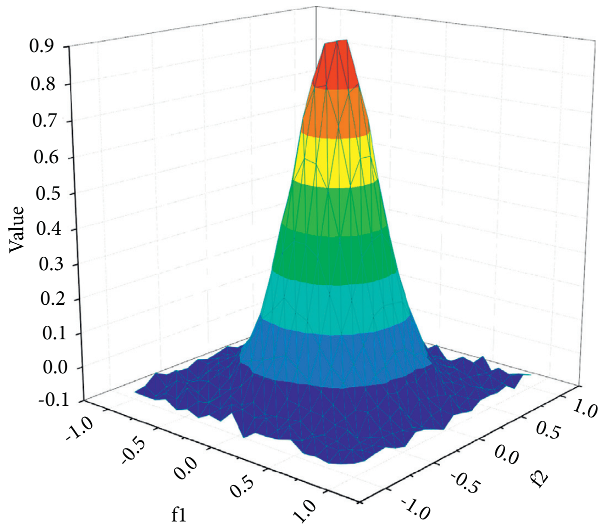

Figure 5: DEA model performance results.

paper applies the resource allocation optimization method to the "Smart Police System" project to optimize the hardware allocation resources of HBase clusters. The main role of HBase in the project is to store and query historical traffic data, which mainly involves the performance indicators of data reading throughput and response latency [20]. This section takes the performance requirements of HBase throughput and response latency as the starting point and uses the above-mentioned resource allocation optimization method to optimize the cluster hardware resources to ensure that the performance of the data query service in the project is not affected based on reducing the original cluster capital cost. The feasibility of the research content of this paper is tested in the actual system. The efficiency mentioned in this type of research refers to the efficiency of high-quality resource gathering, at the expense of partial fairness. Strictly speaking, it is only the efficiency within a certain scope in a specific sense, not the efficiency of the whole society in a general sense.

The inverse DEA model is used to adjust the estimated amount of input or output indicators of several provinces (municipalities directly under the Central Government and autonomous regions) where the pure technical efficiency of other preschool education resource allocation is more stable or the returns to scale are in the incremental stage in 2016. For this category of provinces and municipalities with stable efficiency or increasing returns to scale, appropriate adjustment of the values of their input indicators can make their outputs increase, while maintaining the original resource allocation efficiency. The average growth rate of each input value in 2014–2016 is used to set the input, and the optimal solution result of the output value is calculated. Among the above provinces and cities, the proportion of preschool-educated students in the first grade of elementary school enrollment in Chongqing, Sichuan, Guizhou, Gansu, Hainan, and Qinghai for the output indicator is still relatively reasonable. The inverse DEA model for the number of children in school in Guangdong measures 4080297; the output result of the phase is less than the actual number of children in school in 2016, guessing that the reason is that the number of children in school in Guangdong Province is several times or even tens of times relative to other provinces (municipalities directly under the Central Government and autonomous regions), and in the input of each indicator, there are various preschool education resources per capita in Guangdong Province. This is the reason for the "not rich" situation, as shown in Figure 6.

The solution of the final inverse DEA model for optimal resource allocation shows that the input of financial resources is redundant relative to other indicators, and all provinces and cities have unreasonable input in the input indicators. In the inverse DEA results for improving the efficiency of resource allocation in each province, several provinces listed above, including Beijing, Shanghai, Jiangsu, Zhejiang, Tianjin, Shaanxi, Tibet, Inner Mongolia, and Ningxia, need to invest more in the proportion of full-time teachers with a specialist degree or higher, as well as the proportion of teachers with senior titles. This indicates that teacher scarcity and teacher structure are issues that need to be addressed urgently. The inverse DEA model provides a theoretically optimal value of inputs and outputs, and the amount of increase or decrease of inputs in each index can be used as a reference. In practice, the national education department can consider the allocation of resources to a certain region based on the results of the inverse DEA model, and the provincial and municipal governments can adjust the structure of resource input indicators based on the efficiency evaluation results and the inverse DEA results so that the education departments at all levels can work together to optimize the allocation of education resources and achieve the maximum value of utilization under the limited educational resources.

4.2. Simulation Experiment of an Algorithm for Optimizing the Allocation of Preschool Education Resources. Resource allocation in preschool education is related to not only internal factors such as the management style of kindergartens themselves and the management awareness of teachers and directors but also external factors such as the management system of resource allocation, the competitive social environment, and the enactment of policies [21]. Therefore, if we want to effectively improve the efficiency of preschool education resource allocation, we need to analyze the causes affecting the efficiency of preschool education resource allocation at different levels. The previous empirical analysis reveals that pure technical efficiency, scale efficiency, and technological progress all have an impact on resource allocation efficiency. Among them, pure technical efficiency is mainly related to management level issues such as the way and means of resource allocation, scale efficiency is mainly related to the scale of resource allocation, and technological progress is mainly related to factors affecting technological changes such as management system, competitive environment, and human resource structure.

The error rates of random forest in RO and RH scenarios are 7.8% and 9.2%, respectively. The error rates of the

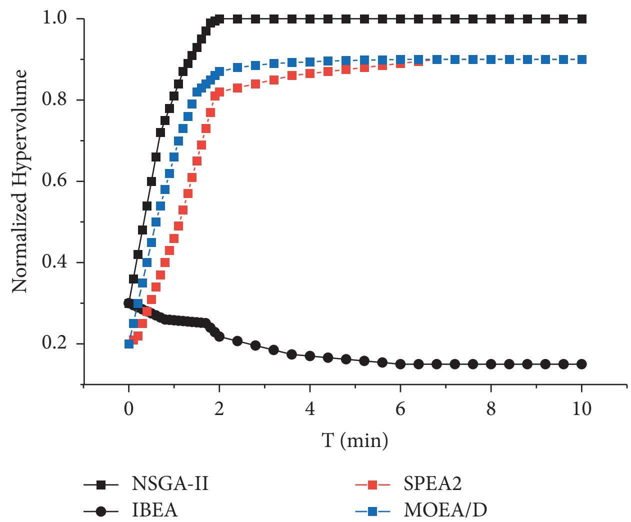

FIGURE 6: Iterative convergence diagram of the super volume of the objective function model.

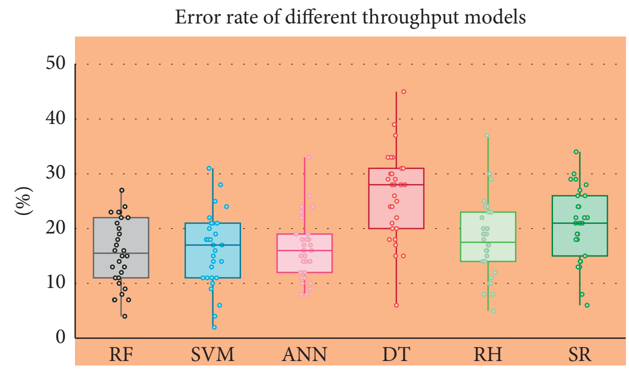

FIGURE 7: Error rate of different throughput models.

random forest algorithms are 20.5% and 22.7% lower than those of the suboptimal algorithms. The error rates of all algorithms in the RMW scenario are not significantly different. The random forest algorithm performs optimally in the RO scenario and relatively poorly in the RH scenario and has a greater overall advantage over other algorithms, and the error rate is optimal in all complex scenarios. The throughput and response delay prediction models constructed by the random forest algorithm obtained above have relatively obvious advantages over three similar algorithms under the same test set conditions, as shown in Figure 7.

Concerning the throughput prediction error rates of the four algorithms, RF, SVM, ANN, and DT, under five different load types of YCSB scenarios, the error rates of the random forest algorithm under two load types, UH and RH, are significantly better than those of the other algorithms, 8.2% and 9.1%, respectively, and 17.9% and 10.6% lower than the next best algorithm. In the RMW scenario, the error rates of the random forest algorithm and the support vector regression algorithm are the same, which has a significant advantage over the other algorithms. In the RO scenario, the random forest error rate is still the lowest, but the improvement is not obvious compared with other algorithms. After analysis, it is concluded that the prediction accuracy of each algorithm is approximately the same in the RO scenario because the data operation type is single and the mapping relationship between hardware resource allocation and throughput is simpler. The mapping relationship between hardware resource allocation and throughput is more complicated in the scenario of mixed load types such as UH, and the difference of prediction error rate between different algorithms is more, which is the reason for the poor performance of the decision tree in the mixed load scenario.

The efficiency values measured by the Bootstrap-DEA model in this paper are more reliable than the traditional DEA model, and the analysis of the current situation of China's preschool education resource allocation efficiency from its measured efficiency values can summarize the trend of changes in preschool education resource allocation efficiency and the overall basic characteristics. The overall efficiency of resource allocation has been gradually improved, the overall input scale has been expanded, and the input factors have become more reasonable. However, at the same time, there are still problems that need urgent attention: firstly, there are large differences in the efficiency of education resource allocation among preschool regions; secondly, the system and management level cannot keep up with the development of its input scale, and there is a certain degree of incompatibility between input and management; thirdly, there is a serious redundancy of financial input in preschool education resource input indicators in developed regions; fourthly, the phenomenon of output in less developed regions is prominent, which is very unfavorable to the provinces resource allocation efficiency and effective use of resources, but in the inverse DEA model can theoretically rationalize the allocation of resources, and the comparison of resource allocation indicators is shown in Figure 8.

The Bootstrap-DEA model is used to measure the efficiency values, and the current situation of the efficiency of preschool education resource allocation is analyzed from its measured efficiency values. While restructuring the investment in education resources, provinces (municipalities) should continuously improve on the existing management level, financial system, and results in transformation system of preschool education, clarify the responsibilities of provincial and municipal (county) governments, guide schools to operate in a standardized manner, and maintain coordination with the scale of investment in education. As a government, it can learn from the experience of regions with higher levels of education management and better institutional security. For regions such as Tibet, Gansu, Ningxia, Qinghai, Xinjiang, and Yunnan, which have insufficient output on the indicator of the proportion of elementary school enrollment with preschool education, they should strengthen the management of young children who have graduated from preschool, from urban to rural areas, to strengthen their learning concepts, strengthen policy guidance for rural and difficult families, and give preferential policies such as reductions, exemptions, and subsidies to young children when they enter kindergarten so that they can have the opportunity to continue to receive to ensure

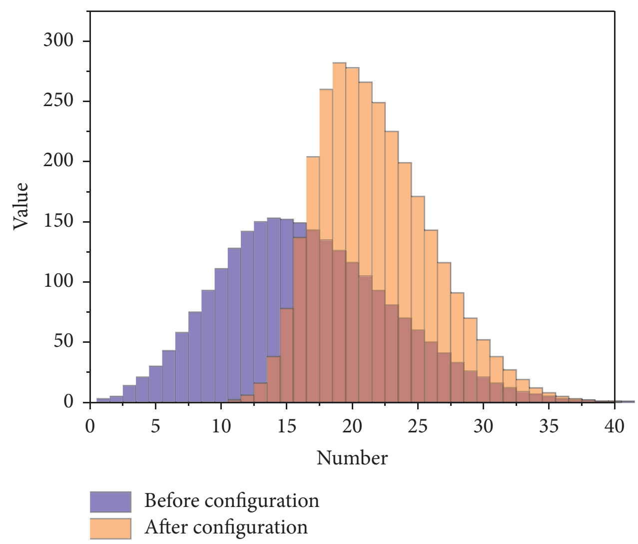

Figure 8: Comparison of resource allocation indicators.

that the goal of preschool education is "basic protection and wide coverage."

Regional socioeconomic development conditions are closely related to the scale, speed, and quality of preschool education development, which is the basic premise of preschool education development. The economically underdeveloped cities and towns should explore their development potential, make full use of the preferential policies of "tilted" development in specific regions, actively introduce foreign investment while actively developing regional ecological resources, create and build regional culture, ethnic traditions, and tourism brands to promote the accelerated development of social economy, and lay a solid economic foundation for the development of preschool education. To build a solid economic foundation for the development of preschool education, the gap between the development of preschool education and that of other regions has been narrowed to a relative balance. In addition to accelerating economic development, governments at all levels should formulate "nesting and attracting phoenix" policies for the introduction and cultivation of preschool talents, establish and improve the return mechanism of talents with the times, inject strong internal driving force for the quality and healthy development of regional preschool education, and innovate the human resource allocation mode of regional preschool education.

5. Conclusion

This paper solves the optimal allocation of preschool education resources based on the data envelopment analysis model. Firstly, a multiobjective optimal allocation model of district and county resources is constructed to enhance the level of education resources allocation, reduce the difference of district and county resources allocation, and improve the utilization rate of education resources, and the indicator weights in this model are solved by the entropy weight method; then, the classical multiobjective evolutionary algorithm NSGA-II is used for coding, population initialization, and the design of key steps of the algorithm, and in the elite selection strategy to avoid local convergence; finally, NSGA-II, SPEA2, MOEA/D, and IBEA are applied to simulate and compare the experiments, and the HV indicators and the convergence effect of the iterative algorithm are used to verify the efficient solution of NSGA-II in the model designed in this chapter and to obtain the results that can improve the level of education resource allocation, reduce the differences between districts and counties, and improve the quality of education resources. The optimized resource allocation plan for each district and county is obtained, which can improve the level of education resource allocation, reduce the difference between districts and counties, and improve the utilization rate of resource allocation.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the 2020 Key Project of Basic Scientific Research Business Expenses of Heilongjiang Provincial Colleges and Universities, "The Practice Research of Normal Colleges' Preschool Art Innovation Ability from the Perspective of Aesthetic Education" (Project Number: 1355ZD018), the Teaching Reform Project of 2021 Mudanjiang Normal University, "Research on Aesthetic Education Teaching Reform and Practice of Preschool Education under the Background of Professional Certification" (Project Number: 21—XJ21012), and the "Innovation and Practice of Student-Centered Talent Training Mode Based on OBE" (Grant no. 21-XJ21048).

References

[1] R. Mammadov and I. Çimen, "Optimizing teacher quality based on student performance: a data envelopment analysis on PISA and TALIS[J]," *International Journal of Instruction*, vol. 12, no. 4, pp. 767–788, 2019.

[2] S. Slijepcevic, "Measuring efficiency at the regional level: a data envelopment analysis approach," *Lex localis - Journal of Local Self-Government*, vol. 17, no. 3, pp. 679–696, 2019.

[3] J. G. Villegas, P. C. Castañeda, and E. Castañeda Gómez, "Planning and performance measurement in higher education: three case studies of operational research application[J]," *Revista Facultad de Ingeniería Universidad de Antioquia*, vol. 2021, no. 100, pp. 97–112, 2021.

[4] A. Ait Sidhoum, T. Serra, and L. Latruffe, "Measuring sustainability efficiency at farm level: a data envelopment analysis approach[J]," *European Review of Agricultural Economics*, vol. 47, no. 1, pp. 200–225, 2020.

[5] S. Mitra and M. Yunus, "Determinants of tomato farmers efficiency in Mymensingh district of Bangladesh: data

Envelopment Analysis approach," *Journal of the Bangladesh Agricultural University*, vol. 16, no. 1, pp. 93–97, 2018.

[6] B. Tu, Y. X. Lin, and Y. M. Zhang, "Efficiency evaluation and influencing factors analysis of governmental expenditure on preschool education[J]," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 14, no. 6, pp. 2533–2543, 2018.

[7] R. Dharmapuri Tirumala, N. Dangol, P. Tiwari, and P. Vaz-Serra, "Comparative analysis of outcomes under PPP and traditional modes of delivery: a study of schools in Melbourne [J]," *Construction Management & Economics*, vol. 39, no. 11, pp. 894–911, 2021.

[8] Q. C. Le, "The efficiency of expenditures for general education: empirical evidence from Vietnam's transitional economy," *International Journal of Management in Education*, vol. 15, no. 4, pp. 381–396, 2021.

[9] D. Stumbriene, A. S. Camanho, and A. Jakaitiene, "The performance of education systems in the light of Europe 2020 strategy," *Annals of Operations Research*, vol. 288, no. 2, pp. 577–608, 2020.

[10] M. Liouaeddine, M. Elatrachi, and E. m. Karam, "The analysis of the efficiency of primary schools in Morocco: modelling using TIMSS database (2011)," *Journal of North African Studies*, vol. 23, no. 4, pp. 624–647, 2018.

[11] G. Halkos and G. Argyropoulou, "Pollution and health effects: a nonparametric approach," *Computational Economics*, vol. 58, no. 3, pp. 691–714, 2021.

[12] M. Jalali Sepehr, A. Haeri, and R. Ghousi, "Evaluation of the OECD countries' healthcare system from the sustainable development perspective[J]," *International Journal of Industrial Engineering & Production Research*, vol. 30, no. 4, pp. 443–464, 2019.

[13] H. Goh, "Analysis of static and dynamic efficiency for sustainable growth of edu-tech companies[J]," *Science*, vol. 7, no. 1, pp. 87–101, 2020.

[14] W. M. Lu, Q. L. Kweh, S. W. Hung, H.-C. Chou, and K.-C. Yang, "Exploring the cultural and creative promotion performance in Taiwan: a network-based ranking approach [J]," *INFOR: Information Systems and Operational Research*, vol. 56, no. 1, pp. 115–135, 2018.

[15] H.-H. Liu and F.-H. Kuo, "Investigating the school characteristics/attributes via the relationships between school management efficiency and representative factors of mobile e-learning: DEA and data mining methodology," *Journal of Information and Optimization Sciences*, vol. 42, no. 4, pp. 883–907, 2021.

[16] D. Rabar, M. G. Fabić, and A. Petrlić, "Financial performance-efficiency nexus in public health services: a nonparametric evidence-based approach," *Economic Research-Ekonomska Istraživanja*, vol. 33, no. 1, pp. 3334–3355, 2020.

[17] Y. Ren, X. Wang, and W. Wu, "Research on the measurement and dynamic evolution of scientific and technological innovation efficiency in universities[J]," *Converter*, vol. 2021, no. 6, pp. 399–412, 2021.

[18] I. D. Pougkakioti and C. Tsamadias, "Measuring the efficiency and Productivity Change of Municipalities: empirical evidence from Greek Municipalities over the time period 2013-2016[J]," *Regional Science Inquiry*, vol. 12, no. 1, pp. 55–74, 2020.

[19] J.-Y. Huang and W.-M. Wey, "Application of big data and analytic network process for the adaptive reuse strategies of school land," *Social Indicators Research*, vol. 142, no. 3, pp. 1075–1102, 2019.

[20] D. Santín and G. Sicilia, "Using DEA for measuring teachers' performance and the impact on students' outcomes: evidence for Spain[J]," *Journal of Productivity Analysis*, vol. 49, no. 1, pp. 1–15, 2018.

[21] C. Yue and X. Xu, "Review of quantitative methods used in Chinese educational research, 1978-2018," *ECNU Review of Education*, vol. 2, no. 4, pp. 515–543, 2019.