Hindawi
Scientific Programming
Volume 2022, Article ID 4160821, 11 pages
https://doi.org/10.1155/2022/4160821

Research Article

Studying the Impact of Health Education on Student Knowledge and Behavior through Big Data and Cloud Computing

Zhilan liu[1] and Xiang Xu[2]

[1]*Department of Public Basic Education, Shanghai Civil Aviation Vocational and Technical College, Shanghai 200235, China*
[2]*School of Military Sports, Jiangxi Agricultural University, Jiangxi 330045, Nanchang, China*

Correspondence should be addressed to Xiang Xu; xuxiang@jxau.edu.cn

Received 25 April 2022; Revised 29 May 2022; Accepted 2 June 2022; Published 17 June 2022

Academic Editor: Lianhui Li

Artificial intelligence and big data, as emerging technologies that have attracted much attention in recent years, have broad application and development space in improving the development of intelligent and refined education in colleges and universities. The application of artificial intelligence and big data to the mental health education practice of college students has a very positive effect on accurately discovering and scientifically solving the mental health problems of college students. In order to combine big data and cloud computing platform organically, this paper introduces an intelligent algorithm based on multi-output support vector regression (MSVR) model and immune clone selection algorithm (ICSA). At the same time, we couple the two to obtain a new intelligent algorithm, namely, immune multiple output support vector regression (ICSA-MSVR) algorithm. Based on the prediction results of health education on students' knowledge and behavior by cloud computing platform, the necessary conditions for three intelligent algorithms to complete the task are summarized. Numerical experimental results show that ICSA-MSVR plays a role in both local search and global search, and is more effective in large-scale cloud computing task scheduling. In addition, in task scheduling, when the task completion time is short, ICSA-MSVR has a lower load imbalance than ICSA and MSVR, which can achieve better load balancing, and the load between virtual machines is closer. Finally, combined with the problems and the needs of students' health education, suggestions are put forward to deepen the application of technology in students' mental health education. This approach can provide corresponding ideas and reference methods for improving the scientificity, pertinence, and effectiveness of mental health education.

1. Introduction

Exercise and health education is an important part of students' knowledge and behavior, and it is an objective daily activity [1, 2]. With the development of technology, the ability to realize its value can be enhanced to a certain extent with the help of big data technology [3, 4]. How to dig, analyze, and use data in a scientific, efficient, and reasonable way, and explore the integration and development of traditional exercise and health education methods with modern information technology have become important means to release the value benefits of health education [5, 6]. In addition, the introduction of big data technology into the vision of health education can not only broaden the way of thinking of health education. This can also effectively integrate the universality of the use of big data and the specificity of health education, and improve the adaptability and realistic effectiveness of sports health education [7, 8].

Enhancing the applicability and effectiveness of sports health education is a realistic requirement under the new situation, new laws, and new scenarios. This situation is mainly manifested in two aspects.

(1) It is necessary to promote the renewal and reform of educational carriers here. These methods are aimed at profound changes taking place. Health education needs to strengthen the selection, update, and optimization channels of content resources and implementation methods on the basis of existing carriers, and conduct information dissemination

through normalized, personalized, and efficient carriers. This can effectively promote the achievement of the target effect.

(2) In addition, this method can optimize the predictive analysis capabilities of education. By combining artificial intelligence methods, we can maximize the advantages of health education.

Strengthening health education is the best path to implement the concept of health first and explore the integration of sports health education and big data, facing new situations, new laws, and new requirements. This is also a topic with outstanding strategic value in the post-epidemic era.

The post-epidemic era not only deepens the connotation of health education, but also puts forward more needs and requirements for it. The application of big data cannot be simply understood as a carrier innovation or technological innovation. The effective release of its series of value benefits requires the use of systematic thinking to make an overall plan for the two and to explore more measures to effectively integrate the two.

Big data is a new stage of informationization, digitalization, and intelligent development, which is both an objective external environment and an important technical carrier for health education. Relying on the new thinking and new concept of modern big data technology [9, 10], it is of distinctive significance and value to solve the methodological problems of health education under the new situation.

Although students entering university tend to be rational and mature in terms of knowledge and mind, due to the lack of deep understanding and experience of reality, they have certain vulnerability in terms of psychological and physical quality, which makes them easy to have mental and physical health problems due to some bad experiences. In the current situation where the pace of life and study pressure continue to increase, the issue of health education has obviously become a prominent problem among college students' groups. Therefore, more and more colleges and universities have incorporated health education into the education system of college students, hoping to enhance their knowledge behavior through professional and targeted health education.

However, the diversity of students' mental states and the lack of resources for health education teachers make it difficult to effectively meet students' individual physical and mental health education needs. This will result in health education often being reduced to a public knowledge curriculum [11, 12]. In such a situation, it is necessary to improve the ability of college teachers to screen, locate, and analyze the health education problems of college students with the help of technology. Artificial intelligence and big data, as the emerging technological content of computer science and technology innovation, can meet exactly this need.

In recent years, colleges and universities have gradually begun to explore the application of artificial intelligence and big data to the practice of college student health education, so as to enhance the relevance and effectiveness of mental health education [13, 14]. At present, the application of technology is still in the exploration stage, and no systematic and comprehensive application method has been formed.

In addition, in the field of school education, the value of focused data resources is extraordinary. However, due to the complexity of basic data resources, they do not directly serve educational activities, and their value can only be reproduced after certain processing, handling, and analysis. Only by focusing on the key points of exercise and health education and reproducing the value points can we synchronize and integrate the two to serve the educational activities. In other words, we can enhance the ability to achieve the goals of exercise and health education in a way that is driven by big data technology.

For a long time, the student population has shown a marked sensitivity to their psychological problems and is reluctant to talk about them too much. This makes it difficult to carry out mental health education for college students in an open and public way. Artificial intelligence and big data can rely on the functional advantages of the Internet platform, enabling universities and teachers to develop a personalized platform for college students' mental health education by using the Internet as a carrier. This allows students to log in to the system for content understanding and activity participation at any time and from any location on their own, truly eliminating the fear of health education for students.

Based on the above analysis, it is necessary to analyze the impact of health education on the knowledge and behavioral ability of students, especially contemporary college students, from the perspective of big data and cloud computing. This paper intends to introduce a novel big data combined with artificial intelligence approach [15, 16](based on immune multi-output support vector regression algorithm) and discuss the application of the novel approach in this field from the perspective of health education, in order to provide an idea for future applications of big data and cloud computing.

2. Immune-Based Multi-Output Support Vector Regression Algorithm

With the rapid development of science and technology, big data and cloud computing methods can be seen everywhere for different fields around the world. For example, the aerospace field often uses big data methods to detect tiny damage in equipment. For some nonlinear phenomena or some seemingly irregular situations, the use of big data analysis can often summarize the development trend of such problems and can have a better guiding role for future planning and development. This method is not only applied in the field of science and engineering, but also favored by researchers in the research process of social science. Taking the health education involved in this article as an example, the introduction of cloud computing and big data technology can to a certain extent liberate the traditional teaching mode that only relies on the teacher's name or urging mode. This method starts from the importance of each student, and reasonably analyzes and predicts the trend of each variable, so that a corresponding teaching mode can be formulated for each individual, and the purpose of teaching students in accordance with their aptitude is truly achieved.

In fact, the realization of big data and cloud computing is also realized on the basis of artificial intelligence or intelligent algorithms. Although big data technology and cloud computing platform seem to be simple in application, the mathematical principles contained in them are very complex. Taking the BP neural network with the simplest principle and the most convenient operation as an example, the calculation of each neuron is obtained by iteration and inversion of a series of nonlinear functions. Although this calculation is complex, with the blessing of computer engineering technology, researchers can obtain a predictable network engineering through code programming. However, it is a huge project for us to turn an initial black box into a practical network project. This requires repeated verification and trial and error to be completed.

Fortunately, with the emergence of artificial intelligence technology, more and more intelligent algorithms are applied in various fields. Compared with the early stage of research, there is only one prediction system, BP neural network, and various optimization algorithms have been introduced. Genetic algorithm is considered to be an effective method to find the optimal solution. The neural network optimized by genetic algorithm improves the solution method of weights and thresholds in its original algorithm, making the whole calculation process more reasonable.

In the process of finding the optimal solution, particle swarm optimization is also applied. Compared with the genetic algorithm, the particle swarm optimization process does not need to set too many parameters in advance. However, before the algorithm is calculated, it requires the researcher to determine the fitness function first, which requires the researcher to have certain prior knowledge.

Compared with the above two optimization algorithms, support vector machine is a new type of calculation method. It shows many unique advantages in solving small sample, nonlinear, and high-dimensional pattern recognition. At the same time, this algorithm can also be extended to other machine learning problems such as function fitting.

In deep learning, support vector machines are supervised learning models related to related learning algorithms. This computational model can be used to analyze data and identify patterns. In addition, it can also be used for classification and regression analysis.

As an upgraded version of the genetic algorithm, the immune algorithm can make up for the shortcomings of the calculation principle inherent in the genetic algorithm. Under the premise of retaining the excellent characteristics of the genetic algorithm, this intelligent algorithm tries to selectively and purposefully use some characteristic information or knowledge in the problem to be solved to suppress the degradation phenomenon in the optimization process.

The immune algorithm simulates the immune process of the human body resisting external antigens through antibodies. It is a swarm intelligence search algorithm with an iterative process of generate and test. In the complex trial calculation process, this algorithm can maintain global convergence on the premise of retaining the best individuals of the previous generation. Therefore, this artificial intelligence algorithm has strong adaptability.

In the process of studying the influence of health education on students' knowledge behavior, this paper attempts to form a new type of prediction system by coupling support vector machine and immune cloning. In particular, we should focus on combining this coupling algorithm with big data technology and cloud computing platform, in order to make this prediction system more adaptable.

2.1. Multidimensional Output Support Vector Regression. The support vector machine [17, 18] learning method was proposed by Vapnik et al. based on the theory of statistical learning. For each regression computational problem, they can be represented as establishing a functional mapping relationship between the input and output quantities. Here, we assume that the function is $y(x) = \omega \cdot x + b$, $\{x_i, y_i\}$ $(i = 1, 2, \cdots, k)$, $\{x_i, y_i\} \in R^d \times R^1$, and allow a certain amount of fitting error to exist by introducing a relaxation factor $\xi_i, \xi_i^* \geq 0$ in accordance with the structural risk minimization principle of statistical learning. In this way, the optimization problem mentioned in the text can be reduced to a minimization problem, and the optimization objective is established with the expression.

$$R(\omega, \xi_i, \xi_i^*) = \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{k}(\xi_i + \xi_i^*), \quad (1)$$

where ω is the fit coefficient; $\xi_i, \xi_i{}^*$ is the relaxation factor; C is the penalty factor; and k is the sample size.

The Lagrange multiplier method [19, 20] is often used by researchers to solve optimization problems for this convex quadratic optimization problem. The Lagrange function can represent the expression as follows.

$$\begin{aligned} L(\omega, b, \xi_i, \xi_i{}^*, \alpha_i, \alpha_i{}^*, \gamma_i, \gamma_i{}^*) = {} & \frac{1}{2}\omega \times \omega + C\sum_{i=1}^{1}(\xi_i + \xi_i{}^*) \\ & -\sum_{i=1}^{1}\alpha_i[\xi_i + \varepsilon - y_i + f(x_i)] - \\ & \sum_{i=1}^{1}\alpha_i^*[\xi_i{}^* + \varepsilon + y_i - f(x_i)] \\ & -\sum_{i=1}^{1}(\xi_i\gamma_i + \xi_i{}^*\gamma_i{}^*), \end{aligned} \quad (2)$$

where $\alpha_i, \alpha_i{}^*, \gamma_i, \gamma_i^*$ is the Lagrange coefficient; $\alpha_i, \alpha_i{}^* \geq 0; \gamma_i, \gamma_i^* \geq 0; i = 1, 2, \cdots, k$.

According to the KKT condition, the expression of its pairwise form maximization function can be expressed as follows.

$$W\left(\alpha_i,\alpha_i^*\right)=-\frac{1}{2}\sum_{i,j=1}^{k}\left(\alpha_i-\alpha_i^*\right)\left(\alpha_j-\alpha_j^*\right)\left(x_i\cdot x_j\right)+\sum_{i=1}^{k}\left(\alpha_i-\alpha_i^*\right)y_i-\sum_{i=1}^{k}\left(\alpha_i+\alpha_i^*\right)\varepsilon, \tag{3}$$

where x is the i-th sample input value and x is the j-th sample input value.

Therefore, solving the optimization problem yields the support vector regression model can be expressed as follows.

$$f(x)=\sum_{i=1}^{k}\left(\alpha_i-\alpha_i^*\right)\left(x,x_i\right)+b, \tag{4}$$

where x is the input value of the sample to be predicted.

The above analysis is only applicable to linear regression problems. However, for nonlinear regression, many researchers need to introduce a feature space and use a nonlinear mapping to map the data to a high-dimensional feature space for linear regression. In addition, the implementation of this algorithm requires replacing the inner product operation in linear regression with a kernel function in the high-dimensional feature space. The kernel function is defined as follows.

$$K\left(x_i,x_j\right)=\phi\left(x_i\right)\cdot\phi\left(x_j\right), \tag{5}$$

where ϕ represents the nonlinear mapping function.

After the same derivation process as linear regression, we can finally obtain the support vector model fitting function as follows.

$$f(x)=\sum_{i=1}^{k}\left(\alpha_i-\alpha_i^*\right)K\left(x,x_i\right)+b. \tag{6}$$

The traditional support vector regression has a one-dimensional variable (SVR) as the output variable. This artificial intelligence algorithm makes its application scenarios limited. In some complex systems, we need to build multi-input-multi-output mapping system to solve the problem. The main reason for this difference is that a 1D SVR does not perform such tasks. Therefore, some research exists to extend the one-dimensional SVR to make it applicable to multidimensional output systems to solve more complex problems in practical engineering.

We extend the one-dimensional insensitive loss function to multiple dimensions. In addition, the loss function can be defined, and its expression is expressed as follows.

$$L\left(u_i\right)=\begin{cases}0, & u_i<\varepsilon\\ \left(u_i-\varepsilon\right)^2, & u_i\geq\varepsilon,\end{cases} \tag{7}$$

where $u_i=\|e_i\|=\sqrt{e_i^Te_i}$; $e_i^T=y_i^T-\phi^T(x_i)W-b^t$; $W=[w^1,\cdots,w^Q]$; $b=[b^1,\cdots,b^Q]^T$, where ϕ is the nonlinear mapping kernel function; x is the sample input row vector; y_i is the sample output row vector $i=1,\cdots,n$; n is the number of samples; and Q is the dimensionality of the output variable.

Based on the loss function shown in the above equation, we can construct the optimization objective function, whose expression can be expressed as follows.

$$L_P(W,b)=\frac{1}{2}\sum_{j=1}^{Q}\left\|\omega^j\right\|^2+C\sum_{i=1}^{n}L\left(u_i\right). \tag{8}$$

To solve the mathematical optimization problem of multidimensional output support vector regression (MSVR) models, a large number of research results have proposed the use of iterative reweighted least squares (IRSL) to solve the problem.

In the optimization objective function of equation (8), we can approximate the loss function by replacing it with a first-order Taylor expansion.

$$L_P'(W,b)=\frac{1}{2}\sum_{j=1}^{Q}\left\|\omega^j\right\|^2+C\left(\sum_{i=1}^{n}L\left(u_i^k\right)+\frac{dL\left(u_i\right)}{du_i}\bigg|_{u_i^k}\frac{\left(e_i^k\right)^T}{u_i^k}\left[e_i-e_i^k\right]\right). \tag{9}$$

Meanwhile, we can construct a quadratic approximation of equation (9) instead of the original equation form. There are studies that confirm the approximate formula that can be used to represent the relationship between the independent variable and the response variable.

$$\begin{aligned}L_P''(W,b)&=\frac{1}{2}\sum_{j=1}^{Q}\left\|\omega^j\right\|^2+C\left(\sum_{i=1}^{n}L\left(u_i^k\right)+\frac{dL\left(u_i\right)}{du_i}\bigg|_{u_i^k}\frac{u_i^2-\left(u_i^k\right)^2}{2u_i^k}\right)\\&=\frac{1}{2}\sum_{j=1}^{Q}\left\|\omega^j\right\|^2+\frac{1}{2}\sum_{i=1}^{n}a_iu_i^2+CT.\end{aligned} \tag{10}$$

The main reason for using this approximate formulation is that W and b are decoupled in this formulation. This intelligent optimization solution does not require iteration, W and b; the approximate solutions of W and b can be calculated directly by taking the partial derivatives of W and b equal to 0. After the optimization objective is solved, the objective is to minimize the overall loss of W and b of the sample set. By the above operation, the multi-output support vector regression model is built.

2.2. Selection Algorithm Based on Immune Cloning. The biological immune system is a complex adaptive system. The human immune system is capable of recognizing pathogens and responding to them [21, 22]. The researchers used this mechanism to give this learning system some ability to learn, remember, and pattern recognize. Immune cloning systems can be used to describe the principles and mechanisms of information processing using computer algorithms to solve scientific and engineering problems.

Castro was the first to propose a clonal selection algorithm (ICSA). The algorithm is an intelligent method for solving complex problems inspired by the human immune system and simulating the function and mechanism of action of the biological immune system. It retains several

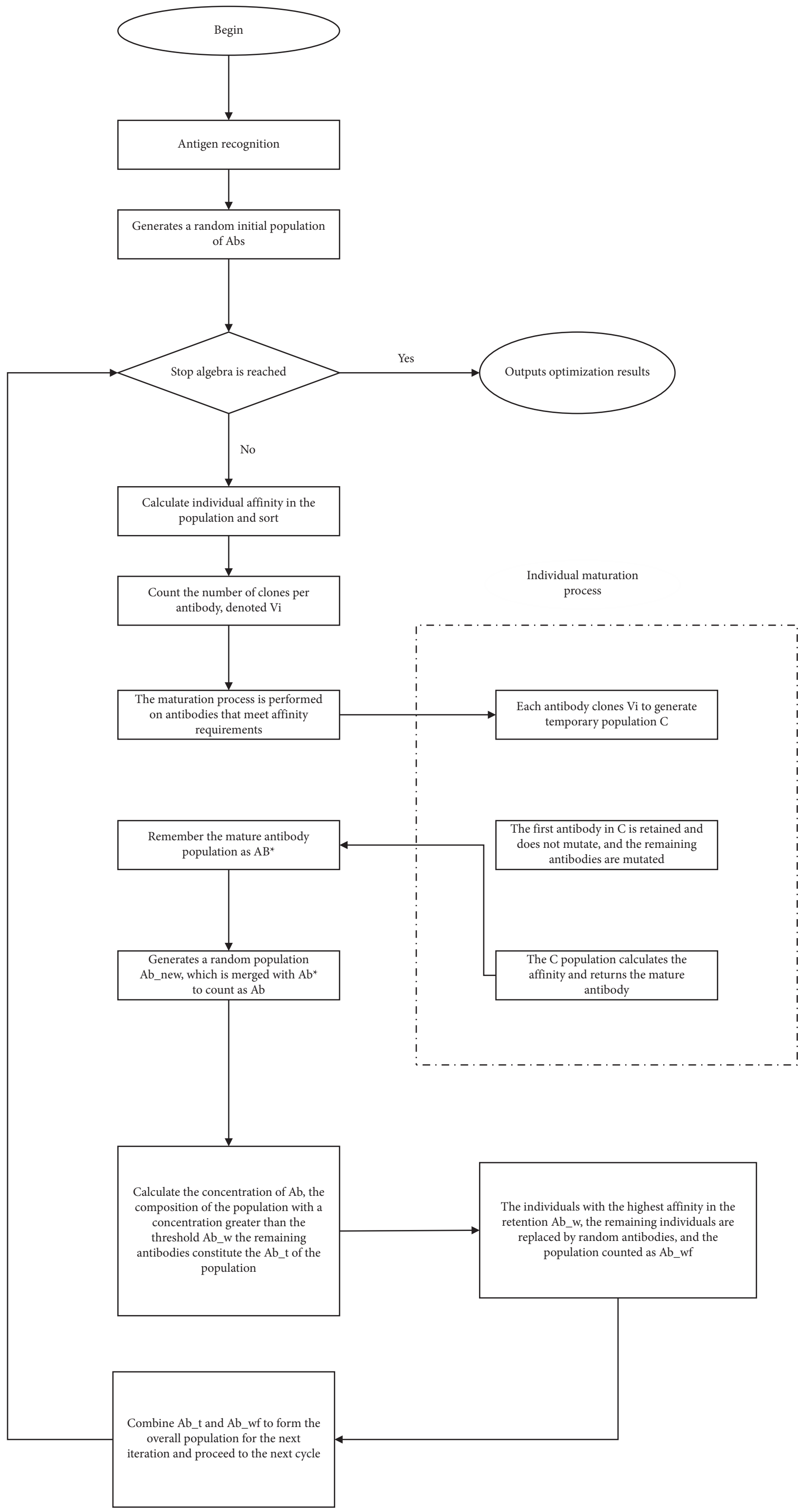

Figure 1: Flowchart of ICSA.

characteristics that are characteristic of the biological immune system. The advantages of this algorithm mainly include global search capability, diversity maintainer, extreme robustness, and parallel solving search process. Researchers can introduce this intelligent algorithm idea into the process of solving optimization problems.

The coupled algorithm introduced in this paper adds a population suppression process to the ICSA to control the average concentration of the population and avoid premature convergence of the algorithm to a local optimal solution. This operation increases the global optimization capability of the artificial intelligence. The detailed process of the ICSA is shown in Figure 1.

The typical multi-peaked function Rosenbrock function (banana function) [24] is used to test the optimization ability of the improved ICSA. Such function expressions can be expressed as follows.

$$f(X) = \sum_{i=1}^{n-1}\left(100\left(x_{i+1} - x_i^2\right)^2 + \left(1 - x_i\right)^2\right), \tag{11}$$

where $X = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^N$.

The global minimum point of the Rosenbrock function is obtained when all independent variables take the value of 1, and the minimum value of the function value is 0. We use the 10-element Rosenbrock function to test the optimization effect of the ICSA for multivariate functions. The search interval of the independent variable is $(-10, 10)$, and the algorithm parameters are set in Figure 2. As shown in the figure, NP denotes the number of antibody population sizes, G denotes the maximum number of cycles, and NC represents the number of clones.

We run the optimization algorithm 10 times and find the minimum value point of the function about 4 times. The results of these 10 times of optimization are shown in Figure 3. As shown in Figure 3, the ICSA has good optimization-seeking capability for multidimensional multi-peak functions. This optimization algorithm can be applied to solve the optimization problem of MSVR model and the prediction performance of research health education based on the coupled ICSA-MSVR algorithm.

The main performance is that when the number of numerical experiments exceeds 5 times, the three algorithms show different prediction performances. When there are less than 5 experiments here, the prediction effects of the three algorithms are basically the same.

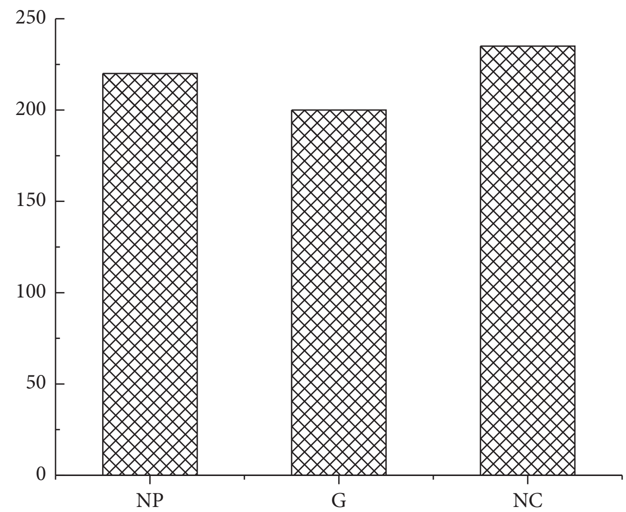

Figure 2: Specific application parameters of the ICSA.

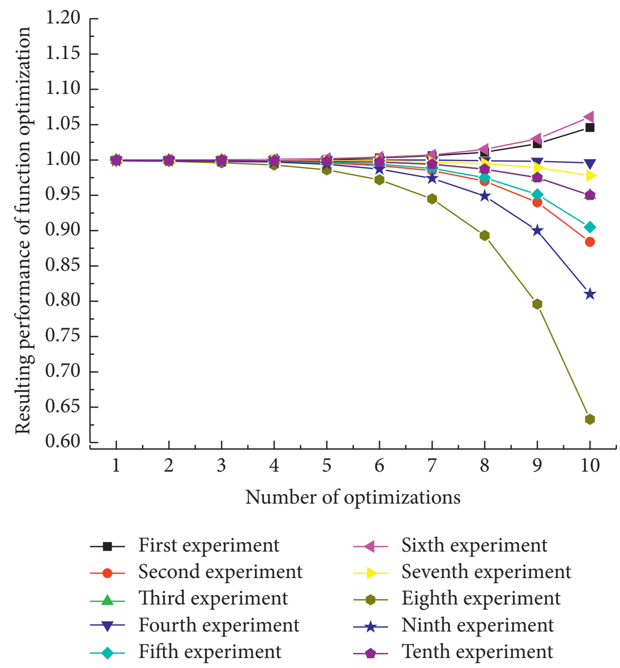

Figure 3: Rosenbrock function optimization results.

2.3. Inverse Analysis Method Based on ICSA-MSVR Coupling Algorithm. The values of the control parameters (penalty coefficients C, sensitivity coefficients ε, kernel function parameters σ) need to be specified artificially in the process of MSVR model building. In order to control the parameter values to achieve the minimum sample training error and the best generalization accuracy of the MSVR model, we propose to solve the parameters optimally by using the ICSA.

In the model training phase, the overall error function of the training sample set is defined as the optimization objective, and the error of individual samples is also adopted as the insensitive loss function. The training samples are divided into learning samples and testing samples, and we use the K-fold cross-validation method to calculate the overall sample error. This way the optimization objective function expression is expressed as follows.

$$\begin{aligned}\left(C^*, \varepsilon^*, \sigma^*\right) &= \underset{C,\varepsilon,\sigma}{\arg\min}\, L_{\text{all}}(C, \varepsilon, \sigma),\\ L_{\text{all}}(C, \varepsilon, \sigma) &= \sum_{m=1}^{k}\sum_{i=1}^{k_m} L\left(u_i\right),\end{aligned} \tag{12}$$

where $L_{\text{all}}(C, \varepsilon, \sigma)$ denotes the overall training loss function; k denotes the number of sample aliquots; k_m denotes the number of training sample aliquots; and the superscript asterisk indicates the optimal parameter.

After the MSVR model is trained, i.e., the optimal MSVR model parameters are optimized by the ICSA. This completes the process of building the prediction model for the positive method inverse analysis process.

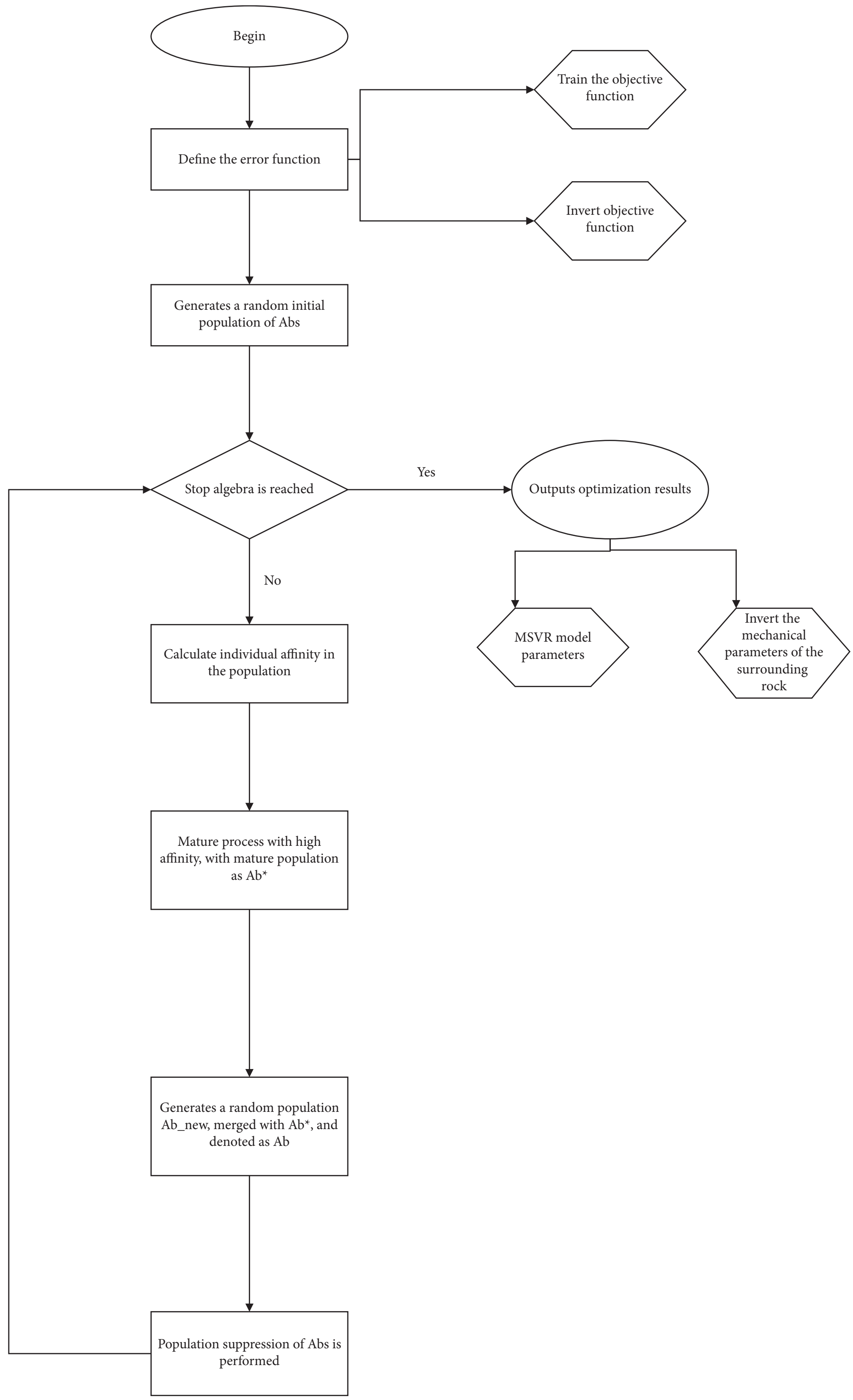

Figure 4: The calculation flow of the ICSA-MSVR coupling algorithm.

In the optimization process of the ICSA, in order to expand the search range of the parameters and the search efficiency, we mapped the parameters to be optimized exponentially; i.e., the range of values of the parameters in the population is the natural logarithm of the actual range of values. In addition, in the actual affinity calculation, the antibody individuals are mapped exponentially, and the calculation formula can be expressed as follows [23, 24].

$$P' = \exp(P). \tag{13}$$

The flow of the coupled ICSA-MSVR algorithm is shown in Figure 4. This newly introduced AI algorithm differs in the definition of the error function and the resultant output part of the model training process and the parameter identification process.

3. Simulation Results and Analysis

We take the results of a health education prediction for students' knowledge behaviors as an example and compare their task completion through the cloud computing platform. CloudSim is a cloud computing simulation platform jointly developed by the GridLab and Gridbus project at the University of Melbourne, Australia. It focuses on simulating cloud environments and testing the scheduling policies of different service models. To test the effectiveness of this paper's algorithm in cloud computing task scheduling, the CloudSim platform is used under Intel i5 processor, 12 GB RAM, and WINDOS10 operating system. In this subsection, we compare and analyze the improved algorithm (ICSA-MSVR), immune cloning algorithm (ICSA), and multidimensional output support vector machine (MSVR) introduced in this paper in three aspects: convergence speed, task completion time, and load imbalance of cloud computing task scheduling.

The strengths and weaknesses of the tested algorithms in terms of convergence speed are mainly reflected in the minimum number of steps to compute the iterations. At the scheduling scale of 200 cloud tasks and 10 VMs, we can compare the convergence speed of ICSA and ICSA by the relationship between the number of algorithm iterations and task completion time. In the analysis, we set the number of antibody population size to 220 and the number of clones to 235.

As shown in Figure 5, ICSA-MSVR converges better than ICSA, both ICSA-MSVR and ICSA converge quickly in the first 100 iterations, and ICSA-MSVR converges faster than ICSA. In addition, the ability of ICSA-MSVR algorithm to develop near the optimal solution is improved and the convergence speed of the algorithm is accelerated. Meanwhile, ICSA-MSVR gradually leveled off after 250 iterations, and the task completion time was less than that of ICSA.

The strengths and weaknesses of the tested algorithms in terms of cloud task completion time are directly reflected in the magnitude of task completion time. We set the number of virtual machines to 10 and the number of cloud tasks to 40, 80, 120, 160, and 200 [25], and then we can compare the task completion time of the three algorithms, ICSA-MSVR, ICSA, and MSVR, in cloud task scheduling and analyze them. The specific numerical experimental results are shown in Figure 6.

As can be seen in Figure 6, ICSA-MSVR takes less time for task completion and is better optimized than the other two algorithms. As the number of cloud tasks increases, the task completion time also increases. When the number of tasks is 40, the task completion time of ICSA-MSVR is 8s and 20s less than that of ICSA and MSVR, respectively. The number of tasks gradually increases, and the task completion time difference of each algorithm increases, and when the number of tasks reaches 200, the task completion time of ICSA-MSVR is 24s and 35s less than that of ICSA and MSVR, respectively, which decreases by 3.7% and 5.3%. The above analysis results prove that ICSA-MSVR works in both local search and global search, which is more effective on larger scale cloud computing task scheduling.

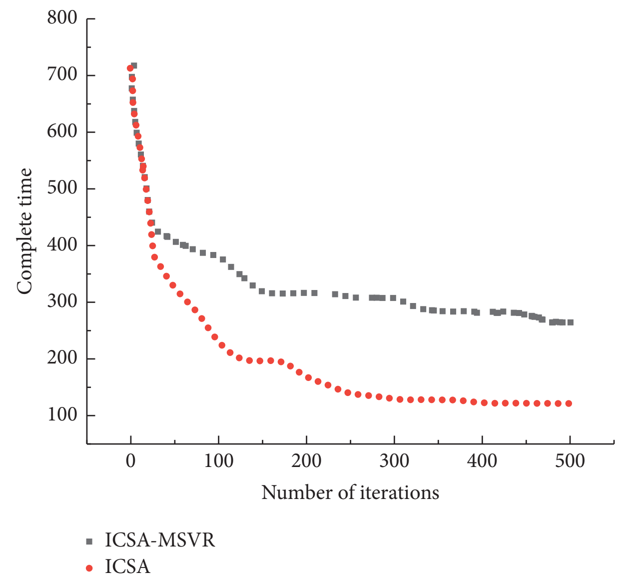

Figure 5: Convergence comparison of algorithms.

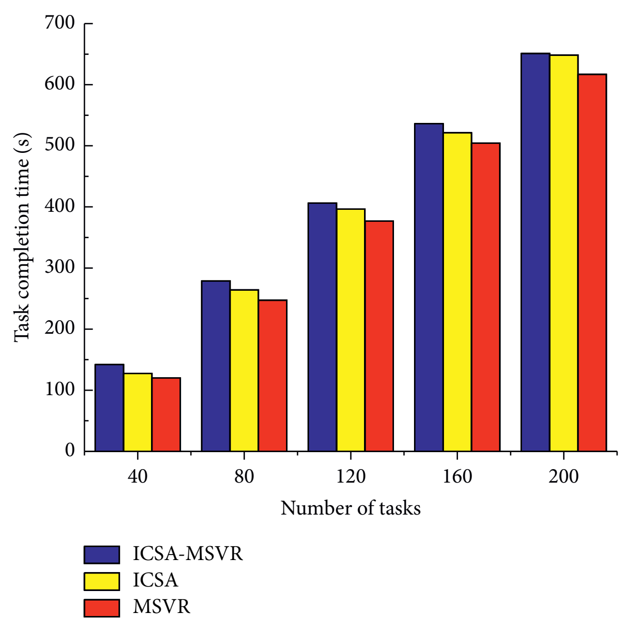

Figure 6: Comparison of task completion time.

What we know is that the degree of load imbalance (DI) of the test algorithm is an important concept, and DI measures the degree of imbalance between virtual machines.

In this paper, we use the standard deviation to represent the imbalance DI. The smaller the DI value, the closer the amount of load among the virtual machines. The better the

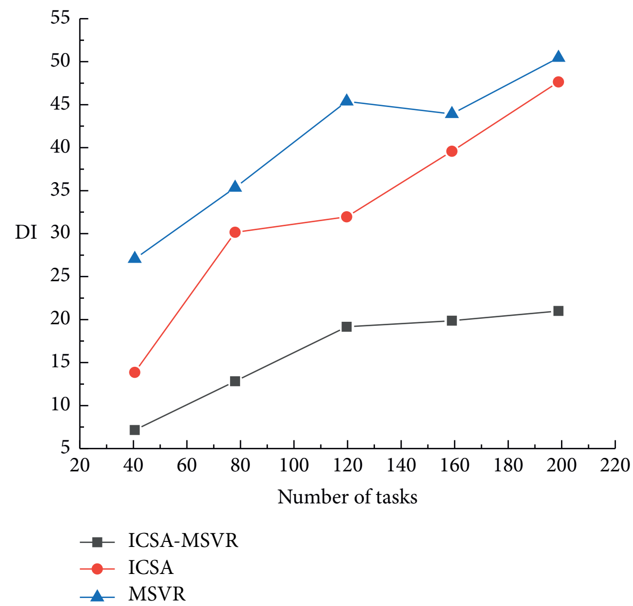

Figure 7: Comparison of DI values of the three algorithms.

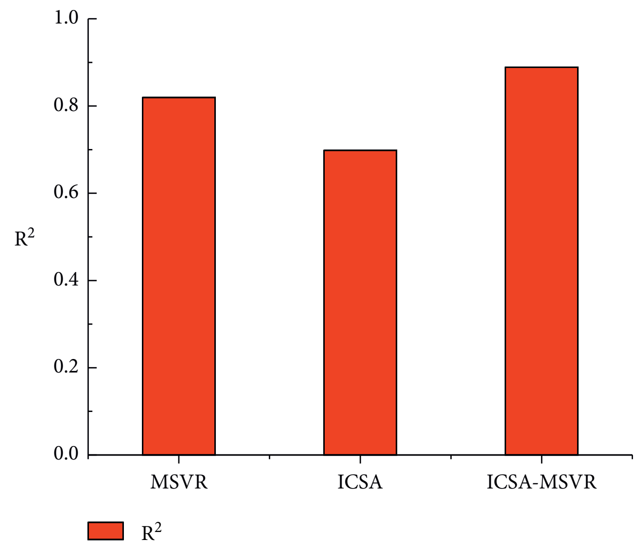

Figure 8: Comparison of the coefficients of determination of the three algorithms.

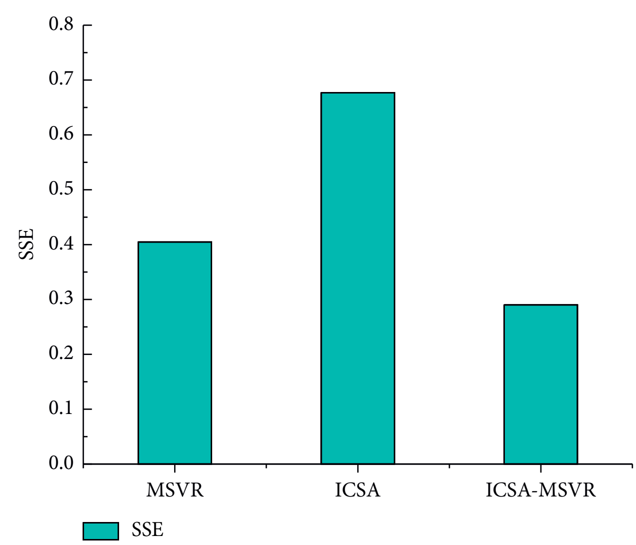

Figure 9: Comparison of the residual sum of squares of the three algorithms.

load balance degree is, the more reasonable the scheduling policy is. The DI is expressed as follows.

$$DI = \sqrt{\frac{\sum_{j=1}^{n} (\text{Time} - AL)^2}{n}}, \quad (14)$$

where AL is the average load of the virtual machine and is the average task completion time of the virtual machine; Time is the load of the virtual machine; and n is the number of virtual machines.

When the number of cloud tasks is 40, 80, 120, 160, and 200, the load imbalance DI of the three algorithms, ICSA-MSVR, ICSA, and MSVR, is compared and analyzed as shown in Figure 7.

It can be seen from Figure 7 that the DI values of the three algorithms increase as the number of tasks increases, and the DI values of ICSA-MSVR are smaller than those of ICSA and MSVR. This situation indicates that in task scheduling with short task completion time, ICSA-MSVR has lower load imbalance than ICSA and MSVR, which can achieve better load balancing and closer load amount among virtual machines.

To further investigate the effectiveness of the three intelligent algorithms in the application of health education prediction of students' knowledge behavior, the next step is proposed by comparing the coefficient of determination (R^2) and the sum of squared residuals (SSE) of the three algorithms. It is well known that the closer the square of the correlation coefficient (R^2) is to 1, the smaller the sum of squared residuals (SSE) is, and the better the fit is. The R2 and SSE corresponding to the three algorithms are plotted in Figures 8 and 9.

As can be seen from Figures 8 to 9, compared with the three algorithms, ICSA-MSVR obtained the largest coefficient of determination and the smallest sum of squared residuals, which can indicate that this method has the relatively best prediction performance.

In addition, we interpolated the predicted performance parameters of ICSA-MSVR for this case using the trajectory interpolation method. The processing results are shown in Figure 10. From Figure 10, we can see that the predicted performance indexes obtained by interpolation have good continuity after the optimization process of the coupled ICSA-MSVR algorithm. This method can provide some theoretical references for the subsequent analytical studies.

We can get from the above simulation case of health education analysis based on cloud computing that artificial intelligence and big data are technical contents that are emphasized and exploited in various fields in recent years, and applying them to mental health education of college

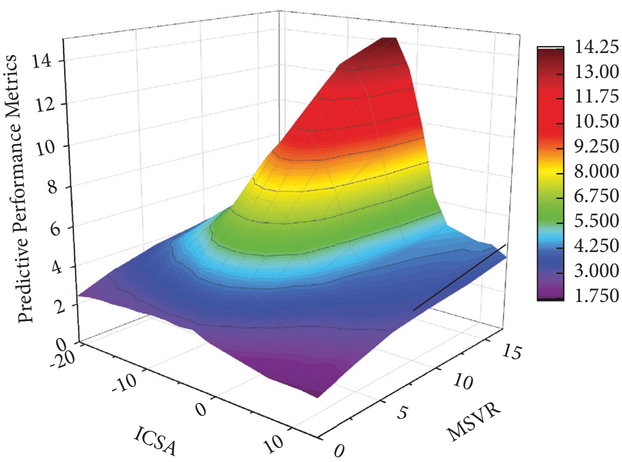

Figure 10: Trajectory interpolation results of ICSA-MSVR prediction performance indicators.

students can not only improve the technicality and accuracy of mental health education. At the same time, this method can also solve the current problem of effective teachers' time and energy in mental health education, which makes it difficult to meet students' personalized psychological education needs. The methods introduced above can make health education carried out in a more scientific and effective way.

4. Conclusion

(1) In this paper, we propose an improved immune clone selection algorithm by adding a population suppression process to improve the convergence to local extrema and the prematureness of the algorithm. After the arithmetic test, the algorithm has good solving ability for multidimensional optimization problems and converges faster. This coupled optimization algorithm can accompany the cloud computing platform to perform certain prediction work on student health education problems.

(2) Taking a certain health education prediction result for students' knowledge behavior as an example, we test the prediction performance of three algorithms, ICSA-MSVR, ICSA, and MSVR, through the cloud computing platform. The results show that ICSA-MSVR works in both local search and global search, and is more effective in scheduling larger scale cloud computing tasks. In addition, ICSA-MSVR has less load imbalance than ICSA and MSVR in task scheduling with short task completion time, allowing better load balancing. The amount of load is closer between virtual machines. In the meantime, compared with the three algorithms, ICSA-MSVR obtained the largest coefficient of determination and the smallest sum of squared residuals, which can indicate that this method has the relatively best prediction performance.

(3) Of course, the current research and practice on the application of artificial intelligence and big data in student health education is still in the exploratory stage, and there are more problems and shortcomings. This requires combining the needs of students' health education and deeply tapping and using the advantages of technologies such as artificial intelligence and big data to truly bring into play the positive functions of technology in education. The combination mechanism of intelligent algorithms introduced in this paper can provide some theoretical reference for such research.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] J. Taylor, G. Forsell, E. Perweiler, and M. Sienkiewicz, "Longitudinal evaluation practices of health workforce development programs: an incremental approach to evaluability assessment," *Evaluation and Program Planning*, vol. 69, pp. 68–74, 2018.

[2] B. Resnick, J. P. Leider, and R. Riegelman, "The landscape of US undergraduate public health education," *Public Health Reports*, vol. 133, no. 5, pp. 619–628, 2018.

[3] B. Knüsel, M. Zumwald, C. Baumberger et al., "Applying big data beyond small problems in climate research," *Nature Climate Change*, vol. 9, no. 3, pp. 196–202, 2019.

[4] S. B. Hao, H. L. Zhang, and M. Song, "Big data, big data analytics capability, and sustainable innovation performance [J]," *Sustainability*, vol. 11, no. 24, 2019.

[5] J. Amankwah-Amoah and S. Adomako, "Big data analytics and business failures in data-Rich environments: an organizing framework," *Computers in Industry*, vol. 105, pp. 204–212, 2019.

[6] X. Xu, D. Li, M. Sun et al., "Research on key technologies of smart campus teaching platform based on 5G network," *IEEE Access*, vol. 7, pp. 20664–20675, 2019.

[7] C. C. Ekin, K. Cagiltay, and N. Karasu, "Effectiveness of smart toy applications in teaching children with intellectual disability," *Journal of Systems Architecture*, vol. 89, pp. 41–48, 2018.

[8] M. Latif, I. Hussain, R. Saeed, and U. Maqsood, "Use of smart phones and social media in medical education: trends, advantages, challenges and barriers," *Acta Informatica Medica*, vol. 27, no. 2, pp. 133–138, 2019.

[9] E. Thomas, "Mobilizing the next generation of health advocates: building our collective capacity to advocate for health education and health equity through SOPHE advocacy summits," *Health Promotion Practice*, vol. 20, no. 1, pp. 12–14, 2019.

[10] J. C. Lawrence, L. L. Knol, J. Clem, R. de la O, C. S. Henson, and R. H. Streiffer, "Integration of interprofessional education (IPE) core competencies into health care education: IPE meets culinary medicine," *Journal of Nutrition Education and Behavior*, vol. 51, no. 4, pp. 510–512, 2019.

[11] A. J. Park, J. M. Ko, and R. A. Swerlick, "Crowdsourcing dermatology: DataDerm, big data analytics, and machine

learning technology," *Journal of the American Academy of Dermatology*, vol. 78, no. 3, pp. 643-644, 2018.

[12] D. J. Kim, J. Hebeler, V. Yoon, and F. Davis, "Exploring determinants of semantic web technology adoption from IT professionals' perspective: industry competition, organization innovativeness, and data management capability," *Computers in Human Behavior*, vol. 86, pp. 18–33, 2018.

[13] W. A. Bhat, "Bridging data-capacity gap in big data storage," *Future Generation Computer Systems*, vol. 87, pp. 538–548, 2018.

[14] M. J. Sanders, R. Bruin, and C. T. Tran, "How to make a standard inhaler device into a 'smart' inhaler that teaches technique[J]," *Journal of Aerosol Medicine and Pulmonary Drug Delivery*, vol. 31, no. 2, p. A14, 2018.

[15] J. M. Liang, W. C. Su, Y. L. Chen, S. L. Wu, and J. J. Chen, "Smart interactive education system based on wearable devices[J]," *Sensors*, vol. 19, no. 15, 2019.

[16] S. Streiter, S. D. Berry, and A. Schwartz, "Teach smarter, not harder: integrating adult learning theory (ALT) into a geriatric fellowship didactic curriculum[J]," *Journal of the American Geriatrics Society*, vol. 67, p. S238, 2019.

[17] S. Verma, S. S. Bhattacharyya, and S. Kumar, "An extension of the technology acceptance model in the big data analytics system implementation environment," *Information Processing & Management*, vol. 54, no. 5, pp. 791–806, 2018.

[18] Y. J. Guo, Z. L. Yang, S. Z. Feng, and J. Hu, "Complex power system status monitoring and evaluation using big data platform and machine learning algorithms: a review and a case study," *Complexity*, vol. 2018, Article ID 8496187, 2018.

[19] A. Eltom, W. Elballa, N. Sisworahardjo, R. Hay, and G. Kobet, "Smart distribution course for 21St century power sector workforce[J]," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5639–5647, 2018.

[20] M. Hasanipanah, R. Naderi, J. Kashir, S. A. Noorani, and A. Zeynali Aaq Qaleh, "Prediction of blast-produced ground vibration using particle swarm optimization," *Engineering with Computers*, vol. 33, no. 2, pp. 173–179, 2017.

[21] M. L. Wilson, S. Sayed, S. Horton, and K. A. Fleming, "Artificial intelligence can augment global pathology initiatives - a," *The Lancet*, vol. 392, no. 10162, p. 2352, 2018.

[22] M. C. Alvarez-Herault, A. Labonne, S. Toure et al., "Corrections to an original smart-grids test bed to teach feeder automation functions in a distribution Grid [jan 18 373-385]," *IEEE Transactions on Power Systems*, vol. 33, no. 4, p. 4647, 2018.

[23] X. Li, X. Liu, C. Z. Li, Z. Hu, G. Q. Shen, and Z. Huang, "Foundation pit displacement monitoring and prediction using least squares support vector machines based on multipoint measurement," *Structural Health Monitoring*, vol. 18, no. 3, pp. 715–724, 2019.

[24] Y. Gong and L. Jia, "Research on SVM environment performance of parallel computing based on large data set of machine learning," *The Journal of Supercomputing*, vol. 75, no. 9, pp. 5966–5983, 2019.

[25] S. Chen, X. Wu, and H. Yin, "A novel projection twin support vector machine for binary classification," *Soft Computing*, vol. 23, no. 2, pp. 655–668, 2019.