



## Digital Impact on the Development and Utilization of Preschool Education Curriculum Resources

Yu Yang<sup>1</sup> 

<sup>1</sup>City University of Zhengzhou, Zhengzhou, 452370, China  
[51221251@qq.com](mailto:51221251@qq.com)

Corresponding author: Yu Yang, [51221251@qq.com](mailto:51221251@qq.com)

**Abstract.** Young children's learning styles are constantly being updated and iterated, teachers' teaching modes and teaching methods are constantly advancing, and parents' educational requirements and goals are also evolving with the times. The advancement of technology and the development of the Internet make preschool education move forward with continuous reform, especially the development of big data brings new opportunities to preschool education, and the thinking of big data will change the whole preschool education model, and the selected topic of this paper is of great significance to the development of preschool education. In the analysis of the 10M data set, which is closer to the production data volume, the number of iterations 5 doubles to 10, the RMSE value decreases less than 0.05, and when the number of iterations doubles from 10 to 20, the RMSE decreases less than 0.03. That is, the RMSE convergence becomes smaller as the number of iterations increases, but the change is relatively weak, especially iterations 10 and 20 have little effect on the RMSE, so it takes into account Efficiency comprehensive consideration in the early operation of the system algorithm iteration parameter selection 10, later with the increase in data volume efficiency decreases can reduce the number of iterations to 5 can also meet the prediction accuracy requirements. Under the guidance of education theory, this paper discusses the changes brought by big data to preschool education from three aspects: first, big data technology brings changes to teachers' teaching mode, which not only improves teachers' efficiency and teaching level, but also makes children's learning from passive to active; second, big data technology brings changes to kindergarten managers' management style, which not only improves managers' work performance, but also enhances parents' satisfaction and reliance on kindergarten. Third, big data technology brings changes to the way of family education, which becomes more scientific and convenient. This paper also suggests that the construction of big data platform should be practical and reasonable to meet the users' habits, and should be more conducive to the interaction between teachers, parents and kindergarten managers.

**Keywords:** preschool education; Curriculum resources; Big data; resource development; Digital Impact.

**DOI:** <https://doi.org/10.14733/cadaps.2024.S16.114-128>

## 1 INTRODUCTION

With the continuous development of information technology and educational concepts, preschool education has changed a lot compared with before, paying more attention to the details of education. The application of big data technology in kindergarten teaching activities also enables teachers to conduct a comprehensive analysis of children more pertinently, which is of great help to continuously improve the level of preschool education and the pertinence of teachers' teaching, and is also conducive to improving the teaching and management of kindergartens. Therefore, it is an important job to carefully analyze the problems and deficiencies existing in the process of implementing preschool education in kindergartens at present [1] Pre-school teachers' exploration and application of information technology in pre-school education practice teaching is mainly reflected in the use of information technology to manage kindergartens in order to improve their own management level; Make use of information technology to assist the daily teaching content and improve the teaching effect; Use information technology to strengthen communication with parents, carry out joint parenting between home and school, and promote the healthy growth of children.

The big media environment has strong complexity, special immediacy and openness, and the mass determines the broad characteristics of the big data environment. Compatibility and equality enable the information of the big data environment to be shared. At the same time, interactivity, influence, permeability and concealment are also one of the important characteristics of the big data environment [8]. These characteristics not only affect how the ideological values of modern preschool education are shaped, but also affect the education and teaching of ideological and political work in preschool education. At the same time, its content and dissemination effectiveness are closely related. That is to say, in the big data environment, it is an inevitable requirement for the preschool education work to optimize and conduct in-depth research on the ideological and political education resources of preschool education. It is closely related to the ideological work of preschool education and has far-reaching significance and value [15]. Digital cultural repositories provide educators with easy access to a vast array of cultural materials from around the world. This content can be integrated into preschool curricula to expose young learners to different cultures, traditions, and historical contexts.

Among these studies, scholars have paid more attention to the impact of specific resources on preschool education. Taking human resources as an example, scholars have keenly noticed the prominent role of human resources in preschool education work, so they have conducted in-depth studies on the combination and collocation of human resources in preschool education work, and they have tried to expand the effect of preschool education work to the optimal effect through the optimal allocation of human resources [18]. Obviously, these studies are very meaningful and have a very high value for our ideological and political education resource allocation. Through the study of big data, it is possible to understand what children really lack and need, their preferred learning style of preschool education, and through technical analysis it is possible to find out what kind of curriculum is most attractive to children and what teaching mode is most acceptable to them and makes them happy and love learning. The innovations are:

1. For parents, big data can let them know their children's learning situation, health situation and teaching progress in the garden one day, so as to achieve the goal of better home education.
2. It provides a strategic idea for the construction and development of big data sharing platform for preschool education.

3. Integrate "blockchain" technology with the construction of educational big data sharing platform to promote the application of blockchain technology in the field of educational big data;

This paper studies the development and utilization of preschool education curriculum resources based on the background of big data. The structure is as follows:

The first chapter is the introduction part. This part mainly expounds the research background and research significance of the development and utilization of preschool education curriculum resources based on the background of big data, and proposes the research purpose, method and innovation of this paper. The second chapter mainly summarizes the relevant literature, summarizes the advantages and disadvantages, and puts forward the research ideas of this paper. The third chapter is the method part, which focuses on the development and utilization of combined big data and preschool education curriculum resources. The fourth chapter is the experimental analysis part. This part is experimentally verified on the dataset to analyze the performance of the model. Chapter five, conclusion and outlook. This part mainly reviews the main content and results of this research, summarizes the research conclusions and points out the direction of further research.

## 2 RELATED WORK

Data sharing refers to individual researchers sharing their raw (or pre-processed) data with others in a formal or informal manner [5]. C. Tenopir argues that with the advent of the data-intensive era, data sharing has become important to provide researchers with a theoretical basis for reanalysis, reducing their duplication of effort and thus accelerating scientific innovation [17]. The concept of SmartPlanet was first proposed by Jin Y extended to all areas of society, including education, where the core of smart education is to provide learners with more personalized learning experiences and experiences [8]. John E. Huber, Taylor M. Moore, compared the differences in efficiency between private and public schools due to different organizational structures and systems, and found that the management approach and the market approach to resource allocation were the efficient institutional choices [9]. Jeff Huddy, Sally Power, David Halpin made a scientific and reasonable analysis and comment on the problems related to the reconstruction of education and education in countries such as the United Kingdom, the United States, Australia, New Zealand and Sweden [6]. In recent years, with the rise of international activities such as open courseware, open classrooms, and open courses, there will be a very large-scale electronic education resource formation. Courseware sharing and the subsequent development of socialized open courses such as MOOCs will also facilitate access to more educational resources [7]. Pesloch put forward the viewpoint of educational psychology, which made the educational theory change from experience summary to science. The main idea is that education should be suitable for children's psychological development, and the teaching contents, principles and methods must be combined with children's psychological reality, so that parents and teachers can stimulate children's interest in learning, enhance their initiative and consciousness in learning, and let children educate themselves [14]. Froebel first put forward the educational theory of kindergarten systematically and completely. Through the sublimation of general education to get preschool education, it has become a separate subject [3]. Mauro et al. studied related papers by analyzing four aspects of information, technology, methods and impacts involved in big data [12].

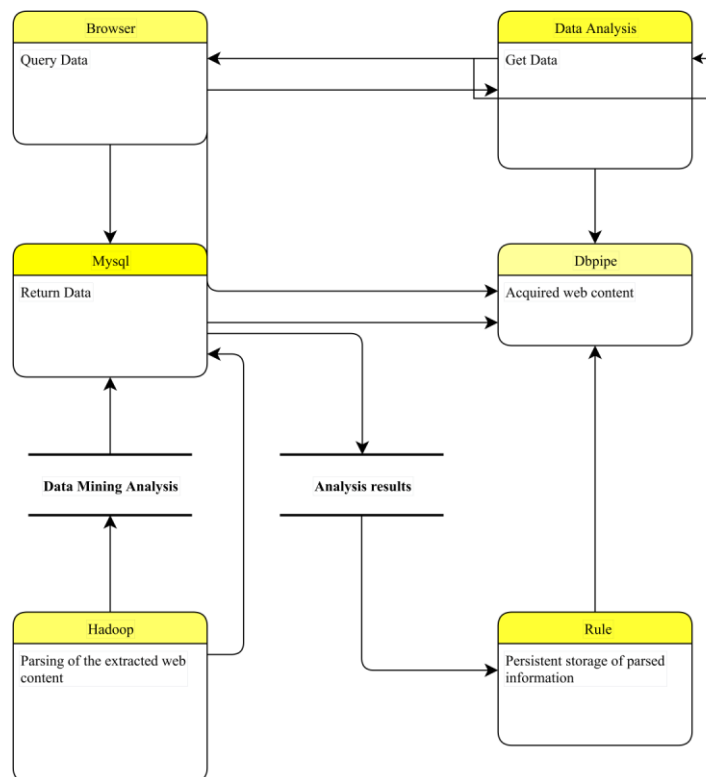
With the continuous development and popularity of big data technology and the continuous maturity of big data technology, it has a wide range of prospects and development space in various fields, which can better and more quickly extract the value and meaning of the data from the collected and integrated data, grasp the law of change of things, let the data speak, more accurately predict the change of things, grasp the law of change of things, and make more scientific decisions so as to better improve the living and learning environment [16]. In this paper, we propose a design scheme for the development and utilization of preschool curriculum resources based on the

background of big data, and carry out the process of matching big data and preschool education resources in the development and utilization of preschool curriculum resources based on the background of big data in Jing, so as to optimize education resources in terms of educational talents and students, and make the distribution of curriculum resources more reasonable.

### 3 METHODOLOGY

#### 3.1 MapReduce

MapReduce is a distributed computing framework, which was first proposed by Google. It is well known because it can process massive data in a distributed cluster consisting of thousands of cheap servers. It is widely used in big data processing. MapReduce can simultaneously complete the reading of large data sets. For writing and processing work, the user processes the key-value data set by defining the Map function to generate an intermediate key-value pair set, and then processes the intermediate set through the Reduce function, and finally merges all the values of the same key in it [19]. The detailed implementation process of MapReduce is shown in Figure 1.



**Figure 1:** MapReduce detailed execution process.

#### 3.2 Based on Content Recommendations

The use of big data enables the mastery of young children's learning behaviors and the identification and resolution of the difficulties they encounter in their learning in order to achieve higher test scores. We need to understand how often a question can be written, what knowledge they apply to

answer it, what questions are simply skipped and not done, whether the question is relevant or not, the preparation done to answer it, and the similarities between the question and similar questions can help in the understanding of young children's learning behaviors. In addition, how the teacher has to advise the student is the most appropriate, and the student's work and problem solving can be detected immediately, so the teacher can pass on the information to the student in a very timely manner. Based on the knowledge and understanding of young children's learning behaviors, and thus the establishment of a responsive learning system, there is a great deal of science, and its strong advancement of students' learning efficiency. From the students' learning behavior, the test writers can then understand the students' abilities thus creating questions for them that suit their personalities. And create paths to enhance memory. From the results obtained from the big data analysis and then analyzed in terms of educational effectiveness, the researchers found that for a series of connected and increasingly difficult topics, students do better than when they do a series of randomly selected different topics that assess the same knowledge [4].

Make use of the professional direction selection and interest topics when users register or log in for the first time to label users, and the data of the skimming comes from the classification tree: the teaching resources are classified according to disciplines in the resource storage stage. Here, the vector space model method in the field of information retrieval is mainly adopted, that is, all the teaching resources to be processed are represented as a set  $D = \{d_1, d_2, \dots, d_n\}$ , the word set appearing in the resources is defined as  $T = \{t_1, t_2, \dots, t_n\}$ , and the normalized frequency of the word  $t_k$  appearing in the resource  $d_j$  is expressed as  $TF(t_k, d_j)$ ; FF is used to express the importance of the word  $t_k$  in all resources  $d_j$ , which is obtained by calculating the total number of resources divided by the number of resources containing the word  $t_k$  and taking the logarithm of the quotient; Then multiply them to get the complete calculation formula 1:

$$TF - IDF(t_k, d_j) = \underbrace{TF(t_k, d_j)}_{TF} * \log \underbrace{\frac{N}{n_k}}_{IDF} \quad (1)$$

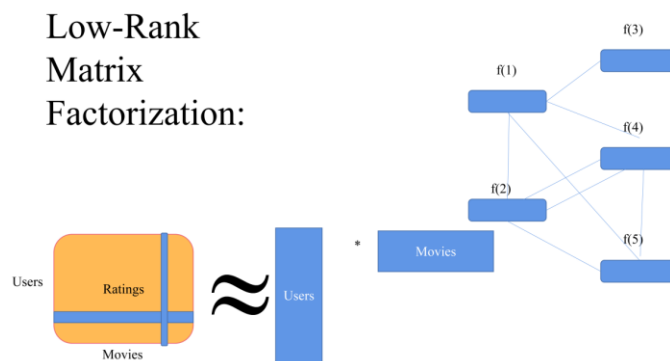
A higher TF-IDF value is calculated by formula 1 to filter out low-frequency words, evolve resource-true classification words, and combine with keyword matching in the classification tree to form resource classification results and store them; The TOP10 resources in the field; the TOP10 resources most concerned by users are synthesized in the order of tags and recommended to users [2].

### 3.3 Matrix Decomposition Model Based Recommendation Algorithm

The development of technology has led us into the era of information technology. Now we can see that in the historical process of kindergarten, whether it is kindergarten enrollment or educational research, whether it is reforming kindergarten education management or improving infrastructure construction, this series of work is appropriate can not leave the support of the element of information, learn to collect, analyze and master information is the kindergarten director in management should do, which is the key to the growth of kindergarten So in the actual In the actual management process, information is the market, information is the business opportunity, accurate information is the key to success, information does not flow will easily lead to the failure of plans and decisions. The value of information first should be valued and created by the director, and the ability to get the right information on time, and to accurately handle and apply this information

determines the success or failure of management to a great extent. The process of management is not a constant, due to the development of things change, the relevant information should be constantly updated, in order to effectively command, coordinate, monitor and carry out the daily management.

The recommendation algorithm based on matrix decomposition model is also a kind of collaborative filtering recommendation algorithm. By decomposing the user's preference value matrix between items into the matrix of potential factors that affect users and resources, and multiplying the decomposed matrices to obtain all users' preference value models for items, the prediction can be made. In this paper, the matrix decomposition algorithm, that is, the cross least squares ALS algorithm implemented in MLHB, is adopted. Firstly, the matrix of potential factors is learned, and the learned matrix of potential factors is divided into smaller matrix blocks for product. Finally, the calculation results of the blocks are connected into a preference matrix. Figure 2 shows the matrix decomposition algorithm model of ML1 movie recommendation demonstration:



**Figure 2:** MLlibALS model algorithm.

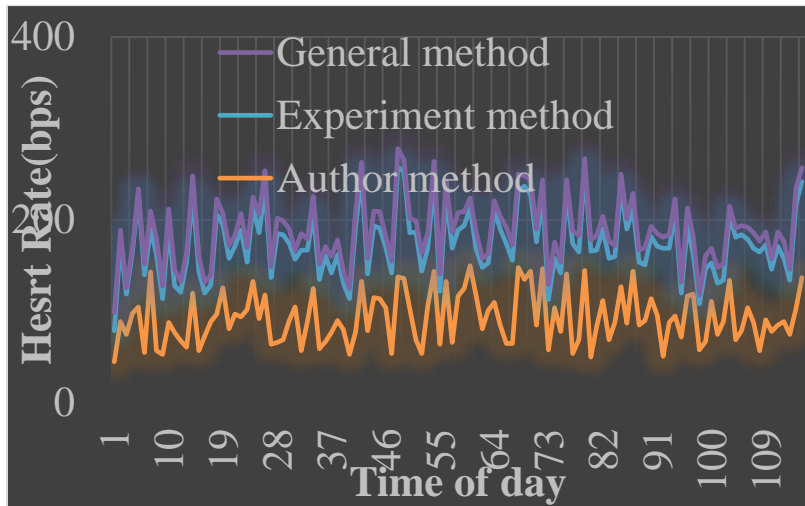
$$f[i] = \arg \min_{\omega \in R^d} \sum_{j \in Nbrs(i)} (r_{ij} - \omega^T f[j])^2 + \lambda \|\omega\|_2^2 \quad (2)$$

### 3.4 Unary Gaussian Distribution

In order to solve the cold start problem in the recommendation system, that is, how to recommend personalized resources for users without any interest preferences and associated friends, it is usually necessary to assist the content-based recommendation algorithm. This paper also adopts the content-based recommendation algorithm. Recommended algorithms to solve this problem. In this paper, the design of the content-based recommendation algorithm is implemented based on the subject classification tree. We start with the simplest and most common univariate Gaussian distribution, whose probability density function is:

$$p(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (3)$$

where  $\mu$  and  $\sigma$  denote the mean and variance, respectively, and this probability density function curve is drawn as the familiar bell-shaped curve, with the mean and variance uniquely determining the shape of the curve [13]. If we take measurements at an infinite number of time points each day, it becomes the case of Figure 3. Note that by using the measurement time as the horizontal axis in the figure below, a line in each color represents a (infinite number of time points of measurement) infinite dimensional sample. When sampling an infinite dimension to get an infinite number of points, it can actually be understood as sampling a function.



**Figure 3:** Function sampling diagram.

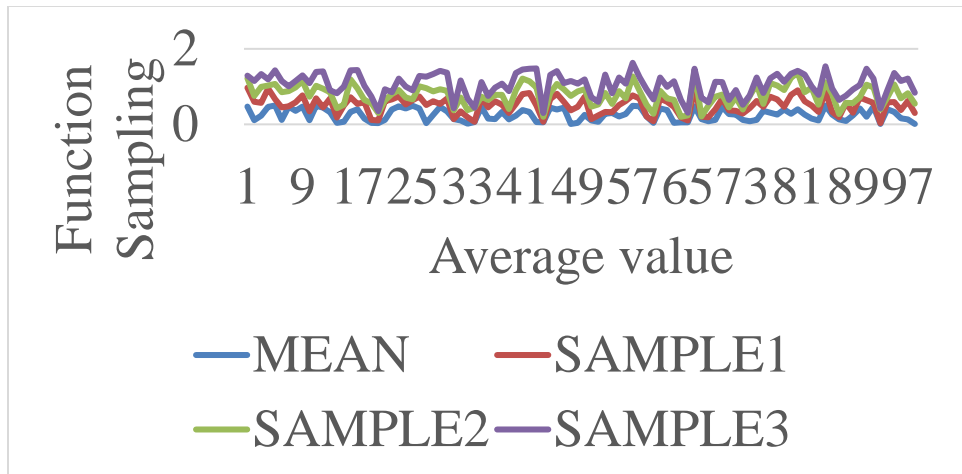
Kernel function is the core of a Gaussian process, which determines the nature of a Gaussian process. The function of kernel in Gaussian process is to generate a covariance matrix (correlation coefficient matrix) to measure the "distance" between any two points. The most commonly used kernel function is Gaussian kernel function, which also becomes radial basis function RBF. Its basic form is as follows.

Where  $\sigma$  and  $l$  are superparameters of Gaussian kernel.

$$K(x_i, x_j) = \sigma^2 \exp\left(-\frac{\|x_i - x_j\|_2^2}{2l^2}\right) \quad (4)$$

### 3.5 Gaussian Process Visualization

The following figure visualizes the Gaussian process, where the blue line is the mean of the Gaussian process, the light blue region 95% confidence interval (obtained from the diagonal of the covariance matrix), and each dashed line represents a function sample (here 100-dimensional simulated continuous infinite dimension is used). Figure 4 shows the prior of the Gaussian process (here the zero mean is used as a prior), showing how the Gaussian process updates its own mean function and covariance function [11].



**Figure 4:** How a Gaussian process updates its own mean function and covariance function.

### 3.6 Multivariate Normal Distribution

The definition of the probability density function of a multivariate normal distribution is:

$$\rho(x) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)} \quad (5)$$

Obviously, when  $d=1$ , multivariate Gaussian is consistent with univariate Gaussian. Sometimes the symbol  $N(\mu, \Sigma)$  is used to represent the Gaussian probability density function with the mean value of  $\mu$  and the covariance of  $\Sigma$  [10].

In order to better understand what multivariate Gaussian is, we consider some cases in two-dimensional space, which is visible. In this case, there are:

$$\Sigma[\sigma_1^2] \quad (6)$$

### 3.7 Preschool Education Mining Technology

The preschool education mining technology is mainly used to process the preschool education big data, and to present the relationship between the relevant variables through the model, and then to judge the learners' future learning situation. The following are the four research goals endowed by the data mining technology:

1. Organize learners' knowledge and behavioral motivation, organize specific learning motivations such as knowledge cognition and learning attitude, establish a model, and then judge the learner's future learning situation.
2. Explore and improve the domain model with the best teaching content and sequence.



3. Study various learning software to see if it can really provide effective support for teaching.

4. Establish a model including the data calculation of the learner field and the teaching Strategy of the preschool education software to improve the learning efficiency of the learner.

Researchers use the following five primary techniques to achieve the above four research objectives:

1. Prediction. Multiple to single integration of predictable variables to achieve a unified model construct that provides technical support for the speculation of a single variable in a passive stance, as if the researcher were able to predict the risks that the learner might encounter by analyzing the learner's online learning environment and the communication with others and the feedback from the test results.

2. Clustering. For example, the researcher groups learners according to their learning forms and learning habits in the online learning environment, and then analyzes the learning patterns and characteristics of each group of learners to provide them with different teaching methods and learning contents.

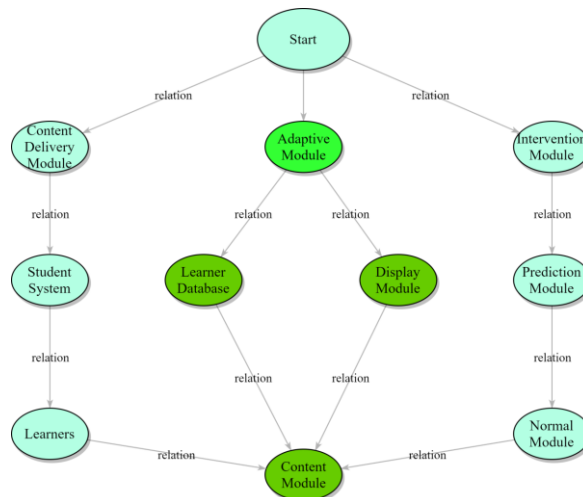
3. Relationship mining. Study the interrelationships hidden in the data and select a result as a rule, for example

For example, using relationship mining, we study the relationship between learners' learning activities and their performance when learning through the Internet, and apply the results to learning.

and apply its findings to the presentation of learning content.

4. Simplification of the human judgment process. This approach is the most important technique in data analysis in order to enable humans to quickly judge and distinguish data characteristics, and therefore describe the data in a way that everyone can understand.

5. Model building. Collecting relevant data and finding correlations between them to construct models that can explain the phenomenon and serve as a model for future analysis. As shown in Figure 5



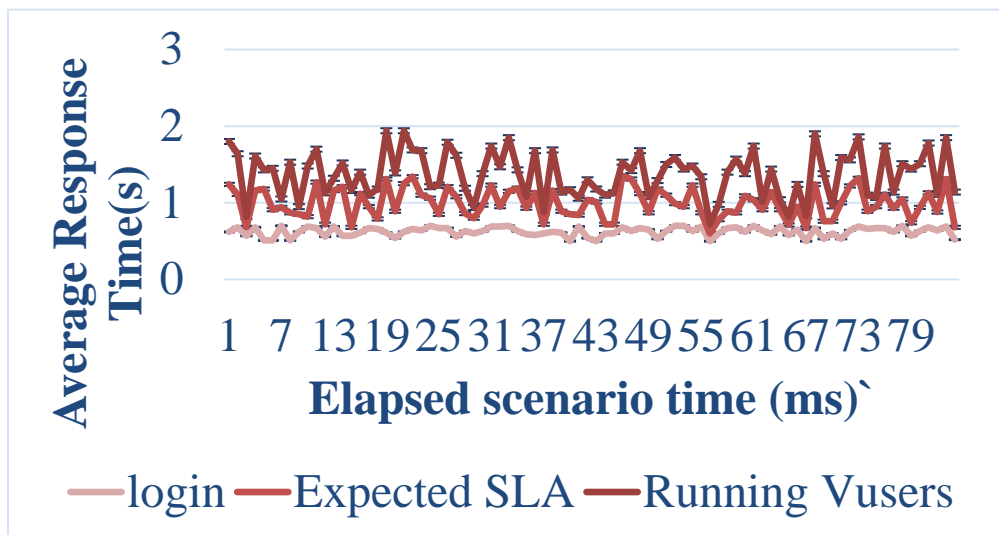
**Figure 5:** Big data-based adaptive learning system composition and operation process.

## 4 RESULT ANALYSIS AND DISCUSSION

### 4.1 Subject Tree Maintenance Subsystem Performance Index Test

For educational big data, openness and access are the prerequisites for sharing. Only openness can be shared. Only by access can there be data that can be opened and shared, and the circulation and exchange of data can be promoted. Therefore, the most important point is that the data resources of educational big data must be opened first, and then obtained, in order to achieve the ultimate purpose of sharing.

The maximum number of online users is tested, and it is estimated that the system has registered 100 maintenance personnel for the subject classification tree. According to the principle of 28, only 20% of the users use the system 80% of the time, and the maximum number of simultaneous online users will not exceed 30% in the peak period. 50 people, the performance of the test system is shown in Figure 6. It can be seen from the figure that the system response time is 0.7s even at the slowest time, which is in line with the design goal of less than 1s.



**Figure 6:** System response time test.

System response time test: 500 concurrent users and 800 concurrent users are taken for stress test, and the performance of the test system is shown in Figure 6, respectively. According to the analysis of two pressure curves in the figure, when the number of concurrent users is 500, the response time of the system is below 0.1s, but when the number of concurrent users increases to 800, the average response time of the system is less than 0.15s, which meets the requirement of less than 1s in performance requirements. Moreover, it can be seen that the increase of the number of users in the system has a linear relationship with the response time of the system, which means that the load capacity of the system to applications is linearly expanded within the limitation of the hardware platform. Now, the concurrent operation of 800 has already met the needs of the construction unit. Even if the number of users increases after N, it can meet the needs of large users only by adding server data or expanding hardware support capacity.

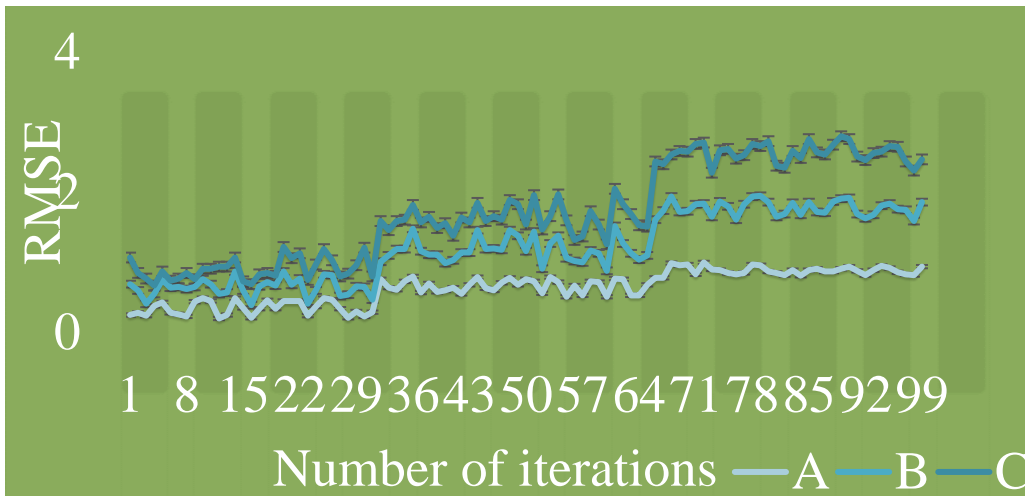
To test the recommendation accuracy of the system, offline experiments were conducted to verify the prediction accuracy of the recommendation algorithm using the MovieLens dataset, and the rootmeansquare error (RMSE) was used to calculate the prediction accuracy. The experiments are as follows: 1. Experimental design, to verify the impact of different data volumes on the convergence of the algorithm, 100K, 1M and 10MH datasets of MovieLens are selected for testing, in which the data volume of the 10-volume dataset is at the 100,000 level, the data volume of the 1M dataset is at the million level, and the data volume of the 10M dataset is at the 10 million level; the dataset is divided into two parts using the two-eight principle, i.e., 80% training data, 20% test data. Experimental data, under the premise of taking into account the prediction accuracy and operational efficiency, experiments were conducted on 100K, 1M, 10M data for 5, 10, 20, 3 iterations of tests, where the daily test RMSE data are shown in Table 1.

| <i>Data</i> | <i>1</i>    | <i>2</i>    | <i>3</i>    | <i>4</i>    | <i>5</i>    | <i>Number of iterations</i> |
|-------------|-------------|-------------|-------------|-------------|-------------|-----------------------------|
| <i>100k</i> | <i>0.93</i> | <i>0.93</i> | <i>0.93</i> | <i>0.93</i> | <i>0.93</i> | <i>5</i>                    |
| <i>100k</i> | <i>0.92</i> | <i>0.93</i> | <i>0.93</i> | <i>0.92</i> | <i>0.92</i> | <i>10</i>                   |
| <i>100k</i> | <i>0.92</i> | <i>0.92</i> | <i>0.92</i> | <i>0.92</i> | <i>0.92</i> | <i>20</i>                   |
| <i>1M</i>   | <i>0.89</i> | <i>0.89</i> | <i>0.88</i> | <i>0.89</i> | <i>0.88</i> | <i>5</i>                    |
| <i>1M</i>   | <i>0.86</i> | <i>0.87</i> | <i>0.86</i> | <i>0.86</i> | <i>0.87</i> | <i>10</i>                   |
| <i>1M</i>   | <i>0.86</i> | <i>0.86</i> | <i>0.86</i> | <i>0.86</i> | <i>0.86</i> | <i>20</i>                   |
| <i>10M</i>  | <i>0.84</i> | <i>0.84</i> | <i>0.84</i> | <i>0.84</i> | <i>0.84</i> | <i>5</i>                    |
| <i>10M</i>  | <i>0.81</i> | <i>0.81</i> | <i>0.81</i> | <i>0.81</i> | <i>0.81</i> | <i>10</i>                   |
| <i>10M</i>  | <i>0.81</i> | <i>0.81</i> | <i>0.81</i> | <i>0.81</i> | <i>0.81</i> | <i>20</i>                   |

**Table 1:** Experimental measurement RMSE data sheet.

The key point of analyzing the influence of data volume on RMSE is that the increase of data volume will lead to the decline of prediction accuracy. On the premise of comprehensive efficiency, the RMSE data of each test data in the above table is iterated for 10 times for comparison. As shown in fig. 7, when the amount of data increases by 10 times and the RMSE appointment decreases by 0.1, that is, the increase of the amount of data will not have a significant impact on the recommendation algorithm, and the increase of the amount of data is more beneficial to the convergence effect of the algorithm.

The influence of the number of iterations of the loss function on the prediction accuracy under the same amount of data is analyzed to determine the selection of system iteration parameters at different stages. Figure 6 takes the 10M data set, which is closer to the production data volume, as an example. The number of iterations from 5 doubles to 10, and the RMSE value decreases by less than 0.03. When the number of iterations doubles from 10 to 20, the RMSE decreases. less than 0.01. That is to say, with the increase of the number of generations, the RMSE convergence will become smaller but the change is relatively weak, especially the 10 and 20 iterations have little effect on the RMSE. Therefore, considering the efficiency, the algorithm iteration parameter selection 10 in the early stage of the system operation, and the later with the increase of the data volume. The efficiency drop can reduce the number of iterations to 5 and also meet the prediction accuracy requirement.



**Figure 7:** RMSE curve of 10m data set with three iterative values.

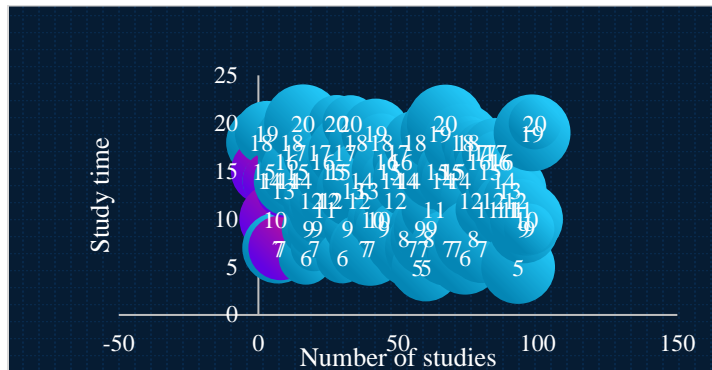
## 4.2 Data Visualization

Data visualization is a means and method that allows users to visualize the characteristics of data. Through this way of visualizing the characteristics of data, educational researchers or users of education-related data can discover relevant laws and gain inspiration through graphical presentation. By displaying the results of data analysis graphically, the characteristics and rules can be more intuitively and clearly expressed, and it is easier to summarize the data laws. In daily teaching research and management, information on students' learning status, corresponding behavioral preferences, and behavioral patterns can be quite abstract and difficult to understand clearly and thoroughly when presented in a conventional tabular or text-based way, but displaying it in a graphical and visual way is especially suitable for this kind of information, so that teachers can make in-depth and intuitive interpretations of the data characteristics shown in these data, and even make a thorough understanding of students' status without much effort. Teachers can interpret the data in depth, and without even thinking about it, they can make accurate judgments about students' status and assign timely guidance or advice.

In all aspects of campus life, as well as in teachers' daily teaching and research and students' learning lives, teachers are often faced with the choice between information and material resources. Especially in today's relatively abundant information resources, it will take more time and effort to select the vast amount of information resources if they are actively searched. The personalized recommendation function in the data analysis platform uses collaborative filtering recommendations to recommend education-related resources. As shown in Figure 8.

As a recommendation algorithm and technology, collaborative filtering is the most commonly used recommendation algorithm at present, and there are many researches on it, which is a very mature recommendation method. The collaborative recommendation algorithm analyzes and calculates the similarity between users based on the calculated score list of system users (teachers or students), finds the nearest neighbors according to the similarity, and then predicts and judges the user's preference for a certain item based on the data of the nearest neighbors. According to the preference degree, the item score list is formed according to the preference degree, and finally the item with the highest score is recommended to the user. The personalized recommendation

process based on collaborative filtering recommendation is to first obtain user rating data, then perform similarity calculation, and finally perform personalized recommendation.



**Figure 8:** Similarity between users.

## 5 CONCLUSIONS

This paper firstly studies the background of personalized recommendation system based on big data and the significance of its application in education field, and points out the urgent need of using personalized recommendation system based on big data to solve the information overload problem in education system through the investigation of existing personalized recommendation technology and recommendation technology based on big data. We also make a comprehensive analysis of the key problems faced by personalized recommendation systems at this stage, such as how to support real-time, reduce technical complexity, and the uniqueness of educational resources, to clarify the difficulties and construction objectives of the system, and give a clear description of the overall architecture design and detailed implementation through this paper. The number of iterations of the loss function on the prediction accuracy under the same data volume is analyzed to determine the selection of iterative parameters for different stages of the system. Taking the 10M data set, which is closer to the production data volume, as an example, the RMSE value decreases by less than 0.05 when the number of iterations doubles from 5 to 10, and by less than 0.03 when the number of iterations doubles from 10 to 20. That is, the RMSE convergence will become smaller but the change will be relatively weak with the increase of the number of iterations, especially iteration 10 and iteration 20 have little effect on RMSE. Therefore, considering the efficiency comprehensively, the algorithm iteration parameter selection of 10 in the early stage of system operation, and the efficiency decrease with the increase of data volume in the later stage can reduce the iteration number to 5, which can also meet the prediction accuracy requirements. My suggestions are as follows: First, the compilation of kindergarten textbooks and teaching outlines should be scientific and reasonable, and the content of the textbooks should be beneficial to the all-round and healthy growth of children. It is necessary to integrate and unify the contents of textbooks, so as to avoid a big gap in children's education level caused by the problems of textbooks. Secondly, we should take various measures to cultivate more talents, offer corresponding majors in colleges and universities, and strengthen the education of applied preschool education talents. Cultivate higher-level talents and strengthen the strength of preschool teachers, so as to promote the development of preschool education. Third, we should continue to strengthen the research on preschool education, improve the existing education system and make the teaching content and methods more scientific. Fourthly, increase the financial input of preschool education and promote the fair distribution of preschool education resources. Fifth, care for the vulnerable groups and provide them with equal opportunities.

Yu Yang, <https://orcid.org/0009-0009-0498-4565>

## REFERENCES

- [1] Chen, Z.; Huang, W.; Ma, L.: Application and Development of Big Data in Sustainable Utilization of Soil and Land Resources, IEEE Access, (99), 2020, 1-1. <https://doi.org/10.1109/ACCESS.2020.3016318>
- [2] Demirba, R.; Sezgin, S.: Book Review: Guidelines on the Development of Open Educational Resources Policies, Athabasca University Press (AU Press), 12(2), 2021, 39-45. <https://doi.org/10.19173/irrodl.v22i2.5513>
- [3] Froebel.: Research on the Utilization and Development of Bamboo Resources through Problem Analysis and Assessment, IOP Conference Series: Earth and Environmental Science, 300(5), 2019, 28. <https://doi.org/10.1088/1755-1315/300/5/052028>
- [4] Girmallesh, H.N.: Utilization of Electronic Resources by the Researchers of University of Agricultural and Horticultural Sciences, Shivamogga: A study, International Journal of Library and Information Studies, 45(3), 2018, 26-29.
- [5] Harju, A.; Balldin, J.; Ladru, D.E.: Children's Education in 'Good' Nature: Perceptions of Activities in Nature Spaces in Mobile Preschools, SAGE PublicationsSage UK: London, England, (3), 2021, 22. <https://doi.org/10.1177/2043610619900519>
- [6] Jeff, H.; Sally, P.; David, H.: Children's Agentive Orientations in Play-Based and Academically Focused Preschools in Hong Kong, Early Child Development & Care, 185(11-12), 2015, 1828-1844. <https://doi.org/10.1080/03004430.2015.1028400>
- [7] Jenkins, J.M.; Whitaker, A.A.; Nguyen, T.: Distinctions Without a Difference? Preschool Curricula and Children's Development, Journal of Research on Educational Effectiveness, 2(1), 2019, 1-36. <https://doi.org/10.1080/19345747.2019.1631420>
- [8] Jin, Y.: Research on Framework of Preschool Education Curriculum Resource Management System Based on web, Revista de la Facultad de Ingenieria, 32(13), 2017, 293-299.
- [9] John, E.; Huber.; Taylor, M.; Moore.: Application and Development of Big Data in Sustainable Utilization of Soil and Land Resources, IEEE Access, (99), 2020, 6.
- [10] Ladigohon, M.C.; Nguyen-Orca, M.F.; Parco, J.: A Study on Media Utilization and Preference for Food and Nutrition Information of Selected College Students in the University of the Philippines Los Baos, 2020(1), 2020, 14-17.
- [11] Lebeninik, M.; Pitt, I.; Stari, A.I.: Use of Online Learning Resources in the Development of Learning Environments at the Intersection of Formal and Informal Learning: The Student as Autonomous Designer, University of Ljubljana, 2, 2015, 25-26. <https://doi.org/10.26529/cepsj.144>
- [12] Mauro.: Study on Resources of Tea Tree in Hainan Province and Its Development and Utilization, Francis Academic Press, 3, 2017, 88.
- [13] Misra, R.P.; Mothukur, A.: Utilization of e-Resources at Technical Information Centers (Library) of Defence Research and Development Laboratory (DRDL), Hyderabad, International Journal of Library and Information Studies, 2, 2019, 47-56.
- [14] Pesloch.: Sustainable Development and Utilization of Plant-Based Resources for the Future, Industrial Biotechnology, 11(6), 2015, 293-294. <https://doi.org/10.1089/ind.2015.29024.lpw>
- [15] Serrano, E.L.; Zuiga, L.H.P.; Valenzuela, P.L.C.: Alignment Between Teacher Performance Standards and Preschool Curriculum, College of Education, University of South Florida, (1), 2019, 65.
- [16] Tao, A.: Study on the Relationship Between Water Resources Utilization and Economic Development Based on Decoupling Theory in Beijing, IOP Conference Series Earth and Environmental Science, 687(1), 2021, 7. <https://doi.org/10.1088/1755-1315/687/1/012207>

- [17] Tenopir, C.; Siew, N.M.: The Development and Validation of a Figural Scientific Creativity Test for Preschool Pupils, *Creative Education*, 06(12), 2015, 1391-1402. <https://doi.org/10.4236/ce.2015.612139>
- [18] Wakabayashi, T.; Andrade-Adaniya, F.; Schweinhart, L.J.: The Impact of a Supplementary Preschool Mathematics Curriculum on Children's Early Mathematics Learning, *Early Childhood Research Quarterly*, 53(4), 2020, 329-342. <https://doi.org/10.1016/j.ecresq.2020.04.002>
- [19] Zhao, H.; Liu, Z.; Yao, X.; Yang, Q.: A Machine Learning-Based Sentiment Analysis of Online Product Reviews with a Novel Term Weighting and Feature Selection Approach, *Information Processing & Management*, 58(5), 2021, 102656. <https://doi.org/10.1016/j.ipm.2021.102656>