

Applied Mathematics and Nonlinear Sciences

<https://www.sciendo.com>

Visual learning analysis of physical virtual simulation experiments based on heterogeneous data features

Guanqi Tao^{1,†}, Yinshu Wang², Yina Fan¹

1. Experimental and Practical Education Innovation Center, Beijing Normal University at Zhuhai, Zhuhai, Guangdong, 519087, China.
2. Faculty of Arts and Sciences, Beijing Normal University at Zhuhai, Zhuhai, Guangdong, 519087, China.

Submission Info

Communicated by Z. Sabir
 Received December 3, 2022
 Accepted April 19, 2023
 Available online October 9, 2023

Abstract

In order to provide a way to develop the teaching effectiveness of physics experiments, this paper optimizes the platform search engine by combining heterogeneous data features and representing document information as feature vectors based on visual learning analysis methods. The algorithm is dynamically adjusted according to the authority to build a network database. And the virtual physics experiments have interacted with virtual experimental equipment to build a physics virtual simulation experiment platform. The results show that the overall level of visual student portrait analysis is above 40%, and the average completion efficiency of visual evaluation tasks 1-9 is 87.9%, which helps the digital transformation and upgrading of experimental physics teaching and promotes the construction of high-quality virtual simulation experimental teaching system.

Keywords: Visual learning; Heterogeneous data; dynamically adjusted; Search engine; Virtual experiment.
AMS 2010 codes: 68T05

†Corresponding author.
 Email address: tgqbnu@163.com

1 Introduction

Compared to other disciplines, physics virtual simulation experiments focus more on inquisitiveness and the design of experimental methods [1]. In the experimental process, the initial parameters of the experimental components are randomly generated at the beginning of the virtual simulation experiment, resulting in diverse experimental results [2-3]. Also, the experimental measurement results can be affected by differences in the accuracy of different models of experimental instruments, resulting in different readings. The physics experiments are step-by-step, and the correlation between the answers is extremely strong, so the complex physics virtual simulation experimental data relationship sorting needs to be carried out with the aid of visual learning analysis for diagnosis, feedback, and other teaching sessions [4-6]. The visualization and analysis process includes steps such as data collection and pre-processing, model building and hypothesis verification, visual representation, and knowledge acquisition, which is a non-linear process [7]. Driven by the concept of precision teaching, visual learning analysis focuses on the intrinsic learning conditions, learning process, and personalized development of learners, which places higher demands on the comprehensiveness and sampling rate of collected experimental data, and the accuracy and execution frequency of learning analysis models.

Learning analytics has become one of the important contemporary supports for providing instructional feedback by processing and analyzing data, using known models and methods to explain significant issues affecting learners' learning, and assessing learners' learning behavior. Valle, N et al. investigated the question of whether the target outcomes of LAD are consistent with domain measures through educational data techniques, raising the application of learning analytics interest in providing a way to visualize information in a meaningful way [8]. Rojas R et al. illustrated that learning basic visual concepts plays an important role in applications such as retrieval and annotation. A neighborhood concept distribution method was also proposed, and the learnability of the proposed method was analyzed [9]. Cheng L et al. constructed a virtual physics experiment teaching platform based on a B/S three-tier architecture system and an MVC model using interactive techniques. The platform started by building a teaching platform for virtual experiment scenarios, analyzing students' mastery of university physics experiments before and after using it, and conducting a systematic survey on students' satisfaction to obtain the strength of the teaching platform's impact on university physics course learning [10]. Wang, Y et al. developed a multi-fusion algorithm to achieve robust whole-body tracking in virtual reality-assisted assembly simulations. The accuracy and robustness of the data were examined by assessing the tracking confidence of the effect of correlation and self-closing errors on the tracking skeleton [11]. Scott M et al. demonstrated that analytical data science could be combined with statistical data modeling to provide easy access to modern analytical tools and visual access to trusted analytical decision points on modern devices [12]. In summary, the above technologies are still in their infancy in the application of virtual simulation experimental teaching, and there is an urgent need to promote relevant theoretical and empirical research and establish scientific and systematic guidance programs and evaluation systems for virtual simulation experimental teaching.

In this paper, based on the visual learning analysis method driven by precision teaching in the design process, the document information is firstly represented as feature vectors by combining the heterogeneous data features generated during the virtual simulation experiments. And according to the different relevance levels, the document information is grouped to optimize the platform search engine. Secondly, the algorithm is dynamically adjusted according to the authority to give experiment authority to different learners, and the learning effect of students' virtual simulation experiments is precisely diagnosed so as to build a network database. In the network database, the learning and forgetting probability of learning behavior big data visualization are introduced into it, the learning difficulty is used as its feature vector, and the skill difficulty is substituted into it to improve the

learning efficiency. Finally, using deep data mining, the set of key features of the data is extracted to visualize the learning behavior pattern for computer verification. From the perspective of user credibility level, the physical, virtual experiment interacts with virtual experimental equipment to build a physical, virtual simulation experiment platform. And the simulation results analysis proves that the platform can provide a reference for the innovation of physics experimental teaching mode.

2 Big data visualization of learning behavior

2.1 Visual Categorization of learning behaviors

Visualized large database learning mainly uses the information resources and communication tools in the network platform, and the temporal and spatial separation and media courseware as the main learning method, reflecting the learners' behavior of their own orientation, motivation, and supervision [13-16]. The visualized learning path is a linear process, i.e., active learning. In the context of big data, learners can choose learning sessions and learning resources according to their learning purposes and learning styles and change from the traditional instructor, environment, and resources to an interactive learning environment mediated by the Internet, and learners' learning behaviors change from instructor-led to autonomous control in the computer environment.

The Bayesian theory-based classification calculation method is easy to use and has a low error rate [17]. The learning behavior is viewed as a set of feature vectors consisting of conditional attribute values in the form of $X(X_1, X_2, \dots, C)$, X_i representing the conditional attribute feature values and C being the classification target attribute. In the network model for big data visualization, the target attribute is the final learning method [18]. If the sample set contains N attributes, A_1, A_2, \dots, A_n constitutes a sample feature vector containing M categories for $\{C_1, C_2, \dots, C_m\}$. If the feature vector for X is $\{X_1, X_2, \dots, X_n\}$, the category for X is calculated as $P(C_i|X)$. The maximum result obtained is the main learning behavior of the visualized network database, then:

$$P(C_i|X) = P(X|C_i)P(C_i)/P(X) \tag{1}$$

The independent assumptions made on the basis of equation (1) above are:

$$P(X|C_i) = \prod_{j=1}^n P(X_j|C_i) \tag{2}$$

On the basis of the above, learning can be divided into three parts by resources, and learning resources are divided as shown in Table 1.

Table 1. Classification of visual learning resources

	Category	Traits
Courses	Course Resource Collection	
Behavior	Big data countability of learning behavior	Learning resource hits, learning outcome scores, surrounding evaluations
Advantages	Degree of data visualization	

2.2 Visual learning supervision

Visual learning supervision requires the computerization of process management, in which the school administrator and the physics teacher have the highest authority of daily management as managers and are responsible for the daily management and operation maintenance [19-21]. According to visual learning supervision, the sorting problem of management is transformed into a classification problem, and Figure 1 shows the results obtained by training the classifier. The document information is represented as a feature vector and grouped according to the relevance level. If the relevance rank of the returned documents f_1, f_2 under query x_1 is $f_1 > f_2$, the feature vector can be redefined: $y_2 - y_1, y_3 - y_1, y_3 - y_2$ as positive samples and $y_1 - y_2, y_1 - y_3$ as negative samples, while their labels will be reassigned at the same time.

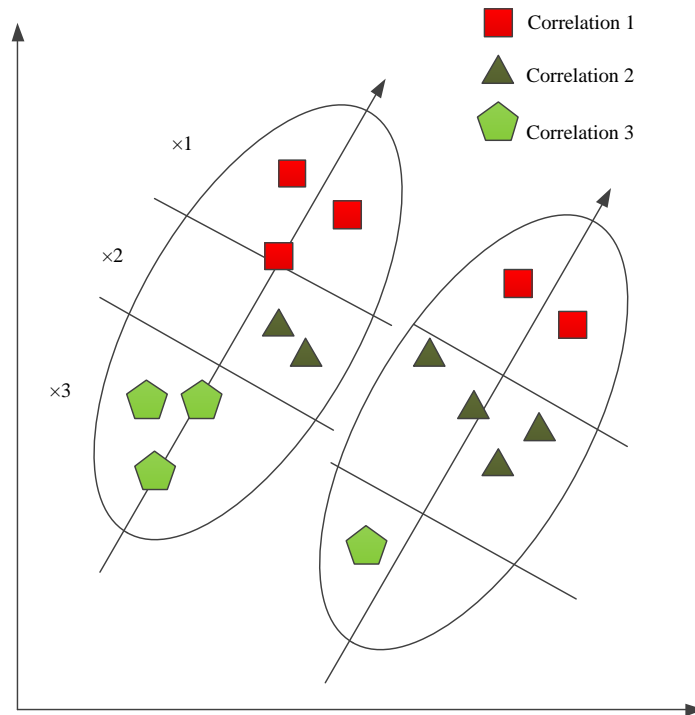


Figure 1. Document Information Grouping

The sorting intensity of the next iteration of the sorted documents is expressed as:

Among the same group, documents with different relevance levels are composed into a new vector and classified using the grouping model to obtain:

$$\begin{cases} \min_{\alpha, \varepsilon} \frac{1}{2} \|\alpha\|^2 + \sum_{e=1}^m \varepsilon_e \\ \text{s.t. } b_e < \alpha, a_e^{(1)} - a_e^{(2)} \geq 1 - \varepsilon_e \end{cases} \quad (3)$$

In equation (3), a denotes the document feature vector, b_e denotes the document relevance label, α denotes the parameter vector, ε denotes the slack variables, and $\varepsilon_e \geq 0, e = 1, \dots, m$.

The sorting intensity of the next iteration of the sorted document is expressed as:

$$\delta_e = \sum_{i:\{e,i\} \in E} \delta_{ei} - \sum_{e:\{i,e\} \in E} \delta_{eiz} \quad (4)$$

In equation (4), δ_{ei} represents the intensity of document e sorted up, δ_{eiz} represents the intensity of document e sorted down, and the absolute values of δ_{ei} and δ_{eiz} are equal. To obtain the specific value of δ_e , it is necessary to calculate δ_{ei} :

$$\delta_{ei} = \frac{-\sigma |\Delta Z_{ei}|}{1 + k^{\sigma(s_e - s_i)}} \quad (5)$$

Based on the exchange of document e with the document i , the sorted position is obtained to generate the difference ΔZ_{ei} , and the query DCG is calculated as:

$$dcg(query) = \sum_e \frac{2^{\text{label}(e)} - 1}{\log_2(e+1)} \quad (6)$$

Based on the above calculation, the labels are arranged in descending order, which is given by the following equation:

$$ndcg(query) = \frac{dcg(query)}{idcg(query)} \quad (7)$$

From equation (7), the document position swaps during the query are calculated one by one to obtain the $lambda$ value. If $lambda$ is a positive value, the larger the value, the greater the rise in the ranking of the documents. If $lambda$ is a negative value, the larger the value, the larger the document's ranking decreases.

The value of $lambda$ is calculated for each document, and the regression tree is trained, with the output of the tree as a fixed value for each leaf node:

$$\lambda_{gh} = \frac{-\sum_{a_e \in R_{gh}} \sum_{\{e,i\} \in E} |\Delta Z_{ei}| \rho_{ei}}{\sum_{a_e \in R_{gh}} \sum_{\{e,i\} \in E} |\Delta Z_{ei}| \sigma \rho_{ei} (1 - \rho_{ei})} \quad (8)$$

where g denotes the number of leaves and h denotes the number of trees.

After obtaining λ_{gh} by equation (6), the product of the obtained result and the learning rate is calculated, which is added to the regression tree to update the score, and the documents under the query are reordered, and the $lambda$ values of the documents are calculated as well as the different values generated by the document sorting position, and the above steps are repeated until the number of trees reaches the set value. On this basis, the platform search engine is optimized so as to realize the construction of a physical, virtual simulation experiment platform based on big data visualization of learning behavior.

The learner's ability to learn behaviors includes the correct answer rate and error rate, which are not the same for each behavior; the more learned behaviors a skill requires, the fewer correct guesses are made. Using $S_j^{(n)}$ as a node observation variable to identify the child nodes in the learning node, when

there are j child nodes in the n th knowledge point, the value of $S_j^{(n)}$ results in 1 or 0. K_n is a behavioral ability hidden variable representing the learning status of the learner, with 1 being a learned skill and 0 being an unlearned skill. C_n is a state variable representing the learner's state at the n th learning behavior, with 1 being correct and 0 being incorrect. $P(L_0)$ is the initial value of the learned behavior, which is used to represent the probability that the learner has mastered the skill. $P(S)$ is the error rate, which represents the probability that the learner has mastered the skill at the n th session and will still make an error, which can be expressed as:

$$P(C_n = false | K_n = true) = \frac{\exp\left(-\sum_{j=1}^n \delta_j S_j^n + \gamma\right)}{1 + \exp\left(-\sum_{j=1}^n \delta_j S_j^n + \gamma\right)} \quad (9)$$

On the basis of equation (9), $P(G)$ represents the learning outcome correct rate, which is the probability that the learner has not mastered the skill in question n , but still gets the correct answer. Since the learning state function C_n is probabilistically dependent on the behavioral ability K_n and the sub-node $S_j^{(n)}$. Therefore, in this paper, using the principle of big data visualization, the correct rate equation is obtained using logistic regression expansion as:

$$P(G_n) = P(C_n = true | K_n = false) \quad (10)$$

$$P(G_n) = \frac{1}{1 + \exp\left(-\sum_{j=1}^n \alpha_j S_j^n + \beta\right)} \quad (11)$$

In the above equation, α_j and δ_j are the degree of influence of the j rd child node on the correctness and error rate. β and γ are the amount of deviation from the correct and error rates.

In order to achieve the relevance and effectiveness of physical and virtual simulation experiments, this paper assigns different learning and experimental privileges to different learners according to a privilege dynamic adjustment algorithm. The virtual role with the smallest permission priority value T calculated by integrating student performance and student classroom progress in the virtual scenario will be assigned permission first. The permission priority value for the student role a is calculated as follows:

$$T_a = \frac{P_a}{\bar{p}} w_p + \frac{g_a}{\bar{g}} w_g \quad (12)$$

For the visual analysis of big data on learning behavior, assuming that a learning skill does not contain any knowledge points, the learners' correct and error rates are both 0.5. Adding β and γ to this causes the learners' learning behavior ability to vary and have different correct and error rates. If $\alpha_j, \delta_j, \beta$ and γ are included as responses to the learned behavioral ability, their values do not change with the change in learning difficulty. It is clear that the correct and error rates are not fixed conditional functions in the construction of visualized web database learning but depend on the learning behavior coefficients.

Big data visualization of learning behavior extends the correct and error rates of learning by considering both learning and forgetting probabilities, as well as the effect of learning difficulty on behavior and correct and error rates when there is a difference in learning difficulty. For example, “1+5” has a higher correct rate compared with “16×26+4”. In the network database constructed in this paper, the learning and forgetting probabilities of the big data visualization of learning behavior are introduced into it, and the learning difficulty is taken as its feature vector.

The system behavior function is defined as:

$$P = \sum_{i=1}^n X_i Y_i \left(\frac{\alpha_i}{\bar{\alpha}_i} \right) \quad (13)$$

When the value of P is larger, it indicates that the learning forgetting exists more efficiently.

Using $P(T)$ to denote the probability of learning, for the part of the learner’s knowledge not mastered in the n rd, the probability of having mastered $n+1$ questions through the act of learning is:

$$P(K_n + 1) = true | K_n = false = \frac{1}{1 + \exp\left(-\sum_{j=1}^n \alpha_j S_n^{(j)} + \beta\right)} \quad (14)$$

Use $P(F)$ to denote the probability of forgetting the part of the n nd knowledge point that the learner has mastered. The probability that the learner has not mastered in question $n+1$ is denoted as:

$$true = \frac{\exp\left\{-\sum_{j=0}^m \delta_j S_n^{(j)} + \gamma\right\}}{1 + \exp\left\{-\sum_{j=0}^m \delta_j S_n^{(j)} + \gamma\right\}} \quad (15)$$

In the above equation, α_j and δ_j are the degree of influence of the j rd subnode on the probability of learning and the probability of forgetting, and β and γ are the amount of deviation from the probability of learning and the probability of forgetting. As with the initial behavioral ability, $\alpha_j, \delta_j, \beta$ and γ indicate the change in the learned behavioral ability with the change in difficulty.

For the obtained visualization results of $P(G)$ and $P(S)$, the learned behavioral ability is not only related to the mastery of the skill but also to the difficulty of the skill by substituting the skill difficulty into it. Therefore, in this paper, based on the visualization of big data of learning behavior, the construction of a network database learning can better improve learning efficiency.

2.3 Learning behavior analysis and visualization module

Database interfacing is accomplished by collecting information on learners’ daily learning behaviors using $xAPI$ data guidelines. $xAPI$ is a guideline to measure the learned content and behavior data. In computer networks, each $xAPI$ record has a grammatical structure of subject, predicate, and object, where the subject represents the initiator of the activity, the predicate represents the specific action instruction, and the object represents the object that bears the activity. The guidelines for the establishment of each database need to store learner interactions within the scope of the business, including the time when learning behaviors such as answering questions, watching multimedia courseware, and exploring knowledge points occur, as well as the content. The learning behavior data

collection module is established using the unified port specified $xAPI$. It regularly receives multiple types of data collected and reported by each data center and gathers and organizes these obtained data to send to the data storage module. The work content of the storage level is to collect the data information on the system and complete the work of filtering, aggregating, storing, and backing up the data, $xAPI$ and the guidelines can store the data for a long time.

The process of constructing the visualized web database is shown in Figure 2. The main purpose of building the learning behavior data analysis module is to read the collected and stored data and use deep data mining to bring up the set of key features of the data and visualize the learning behavior pattern for computer verification. The universal patterns of behavioral patterns and learning laws in the learning process are explored, and the learners' learning status, skill mastery, and ability level are tracked and evaluated. The assessment and prediction data of learners' proficiency levels are visualized in the network database analysis level, and based on the obtained analysis results, the behavioral styles that have the most significant improvement in learning ability are selected, and the results are stored with the database.

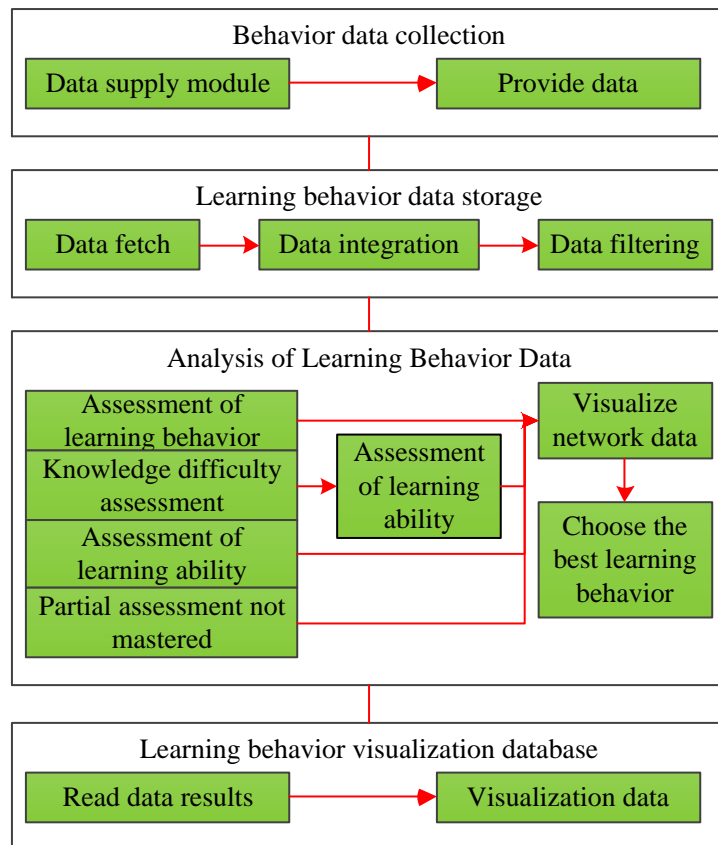


Figure 2. Visualization network database construction process

2.4 Build a visual data learning platform

From the perspective of user credibility level, virtual experiments can be divided into three levels: simulated, exploratory, and empirical experiments. Among them, exploratory experiments are virtual experiments oriented toward the recording, collection, and analysis of data, emphasizing the naturalness of the interaction and the consistency of the measured data in a real environment. It is used to show special things in physics courses and present difficult-to-describe phenomena in a more

intuitive way. The system framework of the physics virtual simulation experiment platform is shown in Table 2.

Table 2. Physical virtual simulation System framework of the platform

Design	Tools	
Frame Design	3DsMax	Three-dimensional modeling
Gesture interaction design	Unity Engine	Scene rendering
Scene design	LeapMotion	Interaction function

In order to build a physical, virtual simulation experiment platform based on big data visualization of learning behavior, this paper interacts with virtual experimental equipment through Leap Motion. Several experimental devices are built in the virtual lab, and students can conduct experiments based on the system’s pre-defined devices. The experimental equipment has been arranged according to the experimental layout, so there is no need to select and connect the experimental equipment, which improves the efficiency of students’ experiments. The development of the physics virtual simulation experiment platform is divided into two major phases, and the development process is shown in Figure 3. The first stage is to analyze the system module and collect materials, build the 3D model related to the physics virtual simulation experiment using 3DsMax, and draw the experimental equipment mapping using PhotoShop. After the model was created, it was exported in FBX format and imported into the Unity project file. The second phase includes the following aspects:

- 1) The main work of building the scene is to build the laboratory scene and experimental equipment in Unity.
- 2) Install the Leap Motion driver, connect the Leap Motion controller to the computer, import the Leap Motion component resource package into the Unity engine, and configure the basic environment for the Leap Motion SDK.
- 3) Realize the interaction between students and virtual equipment in the virtual 3D scene, and use C language and Leap Motion SDK to write interaction scripts according to the experimental process to realize the scene rendering and interaction functions.

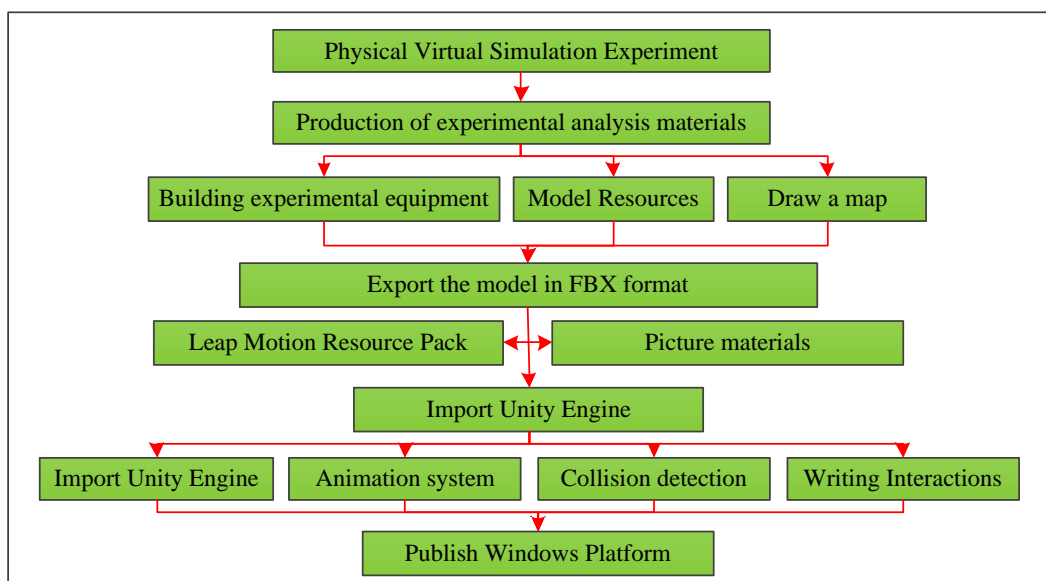


Figure 3. Platform development process

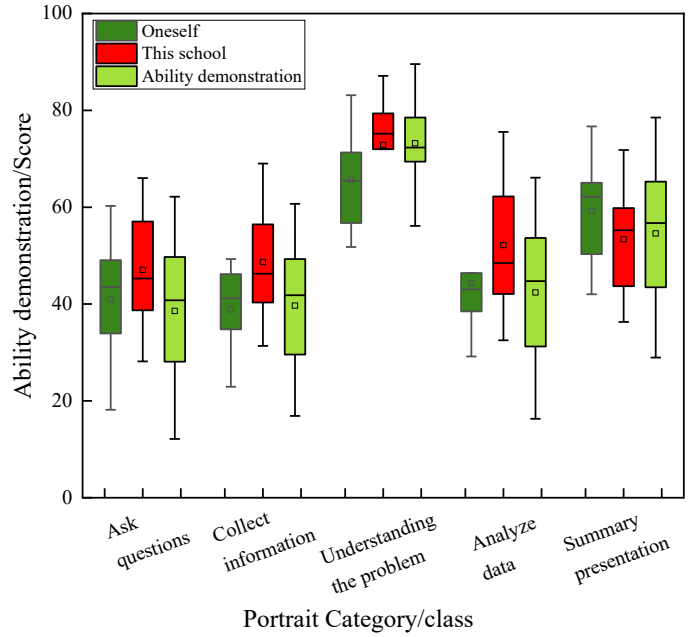
3 Analysis of physical, virtual simulation experimental results

3.1 Visual student portrait analysis

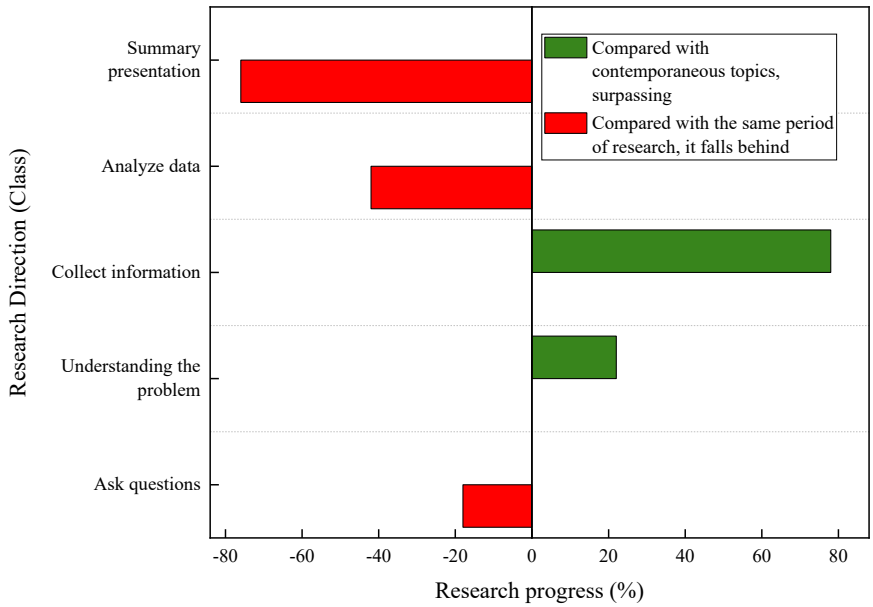
The student-oriented personal portrait helps individual students to self-monitor, improve their self-efficacy, and adjust their personal learning behaviors in time to improve learning efficiency and learning effectiveness. The student ability portrait is the most important part of the physical virtual simulation experiment platform to test the learning behavior based on big data visualization, and the results of the student ability portrait test are shown in Figure 4.

Figure 4(a) shows the label and ability display interface, which displays the profile types to which students belong and their ability attribute labels, and students can view their profile types and labels to understand their behavioral performance. The ability display part is a display of students' individual research ability, presenting their level of task completion in different research stages based on the data of completed projects. Student competency portraits performed better in the big data visualization physical and virtual simulation experiment, with an overall level above 40%, which exceeded the average level in the same city. Performance on the portrait label of asking questions was 53 competencies, on the portrait label of understanding questions was 47 competencies, and on the portrait label of collecting information reached a peak performance of 64.5 competencies. The performance on the profile label for analyzing information was 41.63 competencies, and the performance on the profile label for summarizing the presentation was 60.3 competencies. This section presents not only my individual performance but also the average performance of students across the school and the city at different stages, supporting students to understand and compare visually so that they can adjust their own learning behaviors to improve their inquiry skills.

Figure 4(b) shows the student's current progress in their research on the topic. The horizontal bars represent the student's behavioral performance at different stages, in which both are collecting information and understanding questions exceeded the contemporaneous subject pairwise ratio by 78.15% and 23.19%, respectively. It indicates that the behavioral performance of this student portrait at this stage exceeds the average of the subject matter in the same period. Such presentation of the physics virtual simulation experiment platform based on big data visualization of learning behaviors aims to give students a more intuitive understanding of the progress of the research physics virtual experiment course to adjust their learning pace, while teachers can make targeted interventions and guidance for students' weak stages.



(a) Student profile label and ability display interface



(b) Research progress interface for student-oriented personal portraits

Figure 4. Visual student portrait analysis

3.2 Platform evaluation synthesis degree

Hall voltage measurement is an important and difficult content of the physical, virtual simulation experiment. In this section, the results of the problem domain mining of this content are presented visually using a multidimensional parallel coordinate system plot as an example. The comprehensive evaluation results of Hall voltage measurement in the platform for conducting experiments are shown in Figure 5. The parallel coordinate system drops the multidimensional data from the experiment on multiple equidistant number axes, and as more data are available, the differences in the overall shape of each point become more and more obvious. Teachers can quickly identify different data patterns

based on the point shapes of the various types of data in the parallel coordinate system. The content clustering results were compared with the actual scores of the platform, which yielded a PET value of 0.615, a TTC value of 1.517, and a DST value of 5.456. This indicates that the geometric mean of pairwise accuracy and recall is high, and the clustering quality of the physics experiment data is good. In Figure 5, PET is the correct answer and consists of several sets of positive and negative linear curves. Comparing the correct answers and observing the range of value ranges of the other categories on each number axis in the figure and the relationship between the number axes, it can be found that the points are overall far from the zero axis, i.e., the absolute values of the wrong values are mainly concentrated around 0 or 1.2. The specific reason for the error is that the Hall effect experiment was not zeroed at the beginning, which is caused by carelessness. It is not related to the student's lack of proficiency in controlling the switches of the Hall experimental apparatus. It indicates that the physics virtual simulation experiment platform based on big data visualization of learning behavior can refine the evaluation contained in the virtual physics simulation experiment teaching from the dimensions of experimental operation, experimental data processing, experimental knowledge points, and experimental attitude.

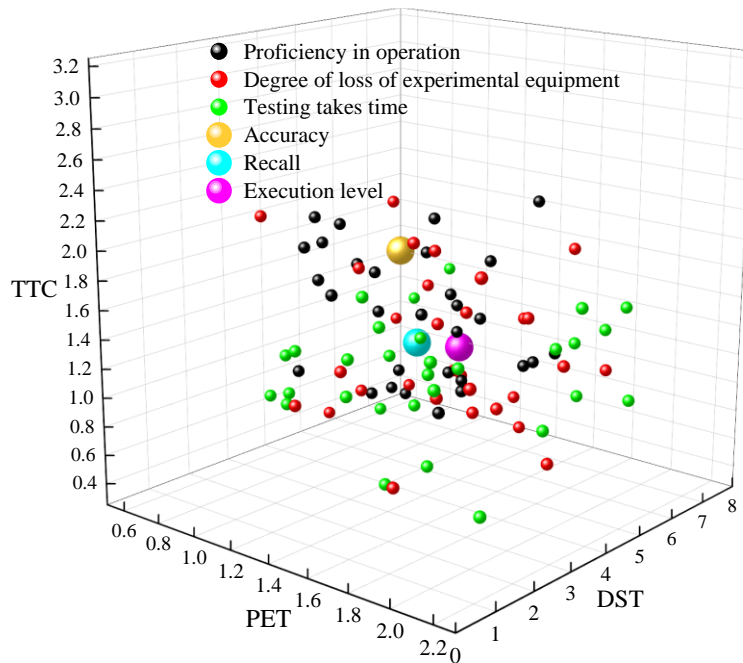


Figure 5. Clustering results of complex multidimensional data

3.3 Task completion effect of physical virtual experiments

In this section, the association rules of the answers to each question of a physics virtual simulation experiment are drawn, and the comparison results with the other two physics virtual experiment operation platforms are shown in Figure 6, which presents all the association rules in a physics virtual simulation experiment and visualizes all the tasks contained in it. The design platform in this paper has a steady peak in the complete efficiency of tasks 1-9, with an average completion efficiency of 87.9%, and the optimal operation is reached in task 1. The completion efficiency was improved by about 18.5% and 26.78% relative to the other two platforms. It is verified that the obtained results are all consistent with the experimental logical relationships. In terms of experimental operations, the physics virtual simulation experiment platform based on big data visualization of learning behaviors can improve the efficiency of teachers' experimental tutoring and students' physics experiment experience through virtual operation demonstrations to ensure high-quality learning.

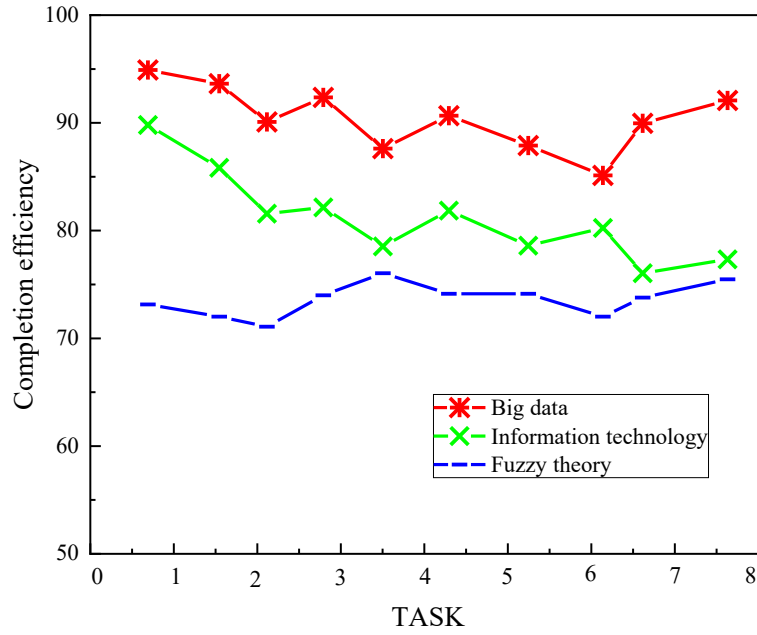


Figure 6. Comparison of the effect of three methods to complete the task

3.4 Analysis of visual database results

The physical, virtual simulation experiment platform is based on big data visualization of learning behavior; the assessment and prediction data of learners' ability level are visualized information, and based on the obtained analysis results, the most significant behavioral approach to learning ability improvement is selected, and the results are stored in the database. The results of the analysis of the web database at the visualization processing level are shown in Figure 7. In response to the knowledge point skills mastered by the learners, it can be seen that the mastery level of knowledge points 1-8 is relatively stable, with an average of 0.62 and the highest distribution in the [0.7-1.0] interval. It indicates that the dynamic learning behavior capability constructed by the platform of this paper can better improve the learning efficiency of the physical, virtual experiment.

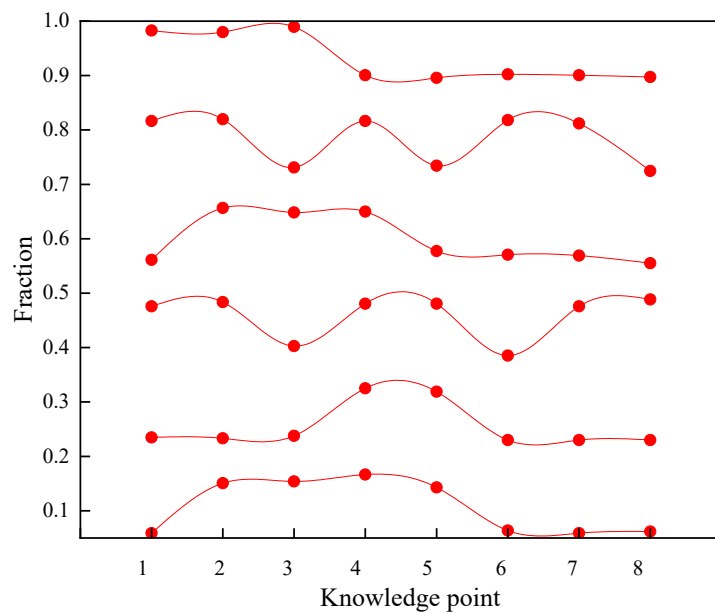


Figure 7. Network database visualization interface

4 Conclusion

In this paper, the visualization is designed by analyzing big data of learning behavior in the context of big data. A physical, virtual simulation experiment platform based on big data visualization of learning behavior is designed. The platform verifies the usage performance through simulation analysis. The test conclusions are as follows:

- 1) It performs better in visual student portrait analysis, with an overall level of over 40%, exceeding the average level of the same city. And it surpassed the same period in research progress in terms of subject pairing ratio, reaching 78.15% and 23.19%, respectively. It can make students more intuitive to understand the progress of the research physics virtual experiment course to adjust the pace of learning.
- 2) In terms of visual evaluation, the clustering results have a PET value of 0.615, a TTC value of 1.517, and a DST value of 5.456. There is a steady peak in the completion efficiency of tasks 1-9, with an average completion efficiency of 87.9%. The mastery level of knowledge points 1-8 is relatively stable, with an average of 0.62 and the highest distribution in the [0.7-1.0] interval. The platform is conducive to the cultivation of students' problem-identification and problem-solving abilities, enabling them to give full play to their sense of innovation in the process of virtual experiments in physics.

Funding:

1. This research was supported by the 2022 Guangdong Education Science Planning Project (Higher Education Project): Research and Development of Artificial Intelligence-Based Virtual Simulation Experimental Teaching Platform (2022GXJK418).
2. This research was supported by 2022, the construction project of teaching quality and teaching reform project in Guangdong Province's undergraduate colleges and universities: Provincial Experimental Teaching Demonstration Center Construction Project: Beijing Normal University at Zhuhai - Physical Experimental Teaching Demonstration Center.

References

- [1] Choudhary, B., Sukanya, Mehta, P., Bach, S., Ruchaud, S., & Robert, T., et al. (2021). Discovery of thiazolidin-4-one analogue as selective gsk-3 β inhibitor through structure based virtual screening. *Bioorganic & medicinal chemistry letters*, 52, 128375.
- [2] Frank, S. M., S Bründl, Frank, U. I., Sasaki, Y., & Watanabe, T. (2020). Fundamental differences in visual perceptual learning between children and adults. *Current Biology*, 31(2).
- [3] Chen, S., & Olsen, S. L. (2021). New physics searches at the besiii experiment. *National Science Review*, 8(11), nwab189.
- [4] Stouwe, A., Tuitert, I., Giotis, I., Calon, J., Gannamani, R., & Dalenberg, J., et al. (2021). Next move in movement disorders (nemo): developing a computer-aided classification tool for hyperkinetic movement disorders. *BMJ open*, 11(10), e055068.
- [5] Krak, I., Barmak, O., & Manziuk, E. (2020). Using visual analytics to develop human and machine-centric models: a review of approaches and proposed information technology. *Computational Intelligence*. 38(3), 921-946.
- [6] Maf, A., Sg, A., Ch, A., & St, A. (2020). Data-driven analysis of product state propagation in manufacturing systems using visual analytics and machine learning - sciencedirect. *Procedia CIRP*, 93, 449-454.

- [7] Demelo, J., & Sedig, K. (2021). Interfaces for searching and triaging large document sets: an ontology-supported visual analytics approach. *Information*, 13.
- [8] Valle, N., Antonenko, P., Dawson, K., & AC Huggins-Manley. (2021). Staying on target: a systematic literature review on learner-facing learning analytics dashboards. *British Journal of Educational Technology*, (2).
- [9] Rojas, R., Kuleshov, S. V., Silva, C., Carvajal, G., Abusleme, A., & Hakobyan, H., et al. (2022). Verdaq: a versatile data acquisition system for high energy physics experiments. *Journal of Instrumentation*, 17(01), P01023 (21pp).
- [10] Cheng, L., Niu, W. C., Zhao, X. G., Xu, C. L., & Hou, Z. Y. (2021). Design and implementation of college physics teaching platform based on virtual experiment scene. *International Journal of Electrical Engineering Education*, 002072092098468.
- [11] Wang, Y., Chang, F., Wu, Y., Hu, Z., Li, L., & Li, P., et al. (2022). Multi-kinects fusion for full-body tracking in virtual reality-aided assembly simulation: *International Journal of Distributed Sensor Networks*, 18(5), 625-636.
- [12] Scott, M. C. (2022). 74 automation, machine learning and computer vision as decision support: taking what animal agricultural does well into the future. *Journal of Animal Science*, (Supplement_3), Supplement_3.
- [13] Varde, A. S. (2022). Computational estimation by scientific data mining with classical methods to automate learning strategies of scientists. *ACM Transactions on Knowledge Discovery from Data (TKDD)*.
- [14] He, H., Zheng, Q., Di, D., & Dong, B. (2019). How learner support services affect student engagement in online learning environments. *IEEE Access*, 7, 49961-49973.
- [15] Ertek, G., & Kailas, L. (2021). Analyzing a decade of wind turbine accident news with topic modeling. *Sustainability*, 13.
- [16] Jia, D., & Shen, W. (2021). Is normalized iris optimal for iris recognition based on deep learning? *Journal of Electronic Imaging*, 30(5), 053007-.
- [17] Li, W. (2022). Improved spatiotemporal information fusion approach based on bayesian decision theory for land cover classification. *Remote Sensing*, 14.
- [18] Liu, S. J., Chang, R. Y., & Chien, F. T. (2019). Analysis and visualization of deep neural networks in device-free wi-fi indoor localization. *IEEE Access*, PP, (99), 1-1.
- [19] Ali, M., Alqahtani, A., Jones, M. W., & Xie, X. Clustering and classification for time series data in visual analytics: a survey. *IEEE Access*, 7.
- [20] Fan, H., Liu, P., Xu, M., & Yang, Y. (2021). Unsupervised visual representation learning via dual-level progressive similar instance selection. *IEEE Transactions on Cybernetics*, PP, (99), 1-11.
- [21] Rostamzadeh, N., Abdullah, S. S., Sedig, K., Garg, A. X., & McArthur, E. (2021). Veronica: visual analytics for identifying feature groups in disease classification. 9).