Construction of a model for analyzing teaching effect of traditional culture courses based on machine learning

Yan Gao
Hunan College of Foreign Studies, Changsha, Hunan, 410000
gg2022123@163.com

Abstract. In order to improve the teaching effect of traditional culture courses, a machine learning-based teaching effect analysis model was constructed. By analyzing 1080 teaching data samples, three algorithms, namely, Random Forest (RF), Gradient Boosted Decision Tree (GBDT) and Support Vector Regression (SVR), were used to design a dataset containing 26 features covering indicators such as attendance rate, timely submission of assignments, and frequency of classroom interaction. The results show that the GBDT model performs best in the indicators of mean square error (MSE) and coefficient of determination (R^2), which are 0.0196 and 0.8834, respectively, showing strong predictive ability and fitting effect. The model provides a scientific assessment basis for the optimization of personalized teaching in traditional culture courses, and is able to improve the quality and efficiency of teaching through a data-driven approach.

Keywords: machine learning; teaching effect analysis; traditional culture course; data evaluation.

1. Introduction

With the rapid development of information technology, the field of education is facing unprecedented changes. As an important part of the traditional culture course to enhance students' cultural literacy, the assessment and optimization of its teaching effect is particularly important. Traditional assessment methods often rely on subjective judgment, making it difficult to reflect teaching results comprehensively and objectively. Machine learning, as a powerful tool based on data analysis, can realize quantitative assessment and accurate prediction of teaching effects through mining and analyzing a large amount of educational data. In the teaching process of traditional culture courses, with the help of machine learning technology, it can scientifically reveal the effectiveness of teaching strategies, help teachers adjust the teaching program in a targeted manner, and further improve the quality of teaching. Therefore, it is of great practical significance and academic value to construct a teaching effect analysis model based on machine learning.

2. The Value of Machine Learning in Educational Assessment

The value of machine learning in educational assessment, especially in analyzing the teaching effectiveness of traditional culture courses, is reflected in several aspects. Machine learning can efficiently process a large amount of teaching data, analyze various feedback information generated by students in the learning process, such as homework grades, classroom participation, midterm and final exam scores, etc., and predict students' learning performance and course mastery through algorithmic models [1] . Regression models in supervised learning can be used to accurately assess the impact of different teaching methods on students' performance and thus optimize teaching strategies. Machine learning can mine students' learning habits, interests and other hidden data to help teachers discover the difficulties and pain points in traditional culture courses, so as to make personalized teaching adjustments. Through the training of a large amount of learning data, the model can identify the differences in the teaching effect of certain teaching content, providing data support for subsequent course design. In addition, machine learning is able to monitor the teaching effect in real time, automatically adjust the evaluation criteria according to the teaching progress, and improve the efficiency and quality of teaching. Through data-driven evaluation, dynamic evaluation and optimization of teaching effects can be achieved, avoiding the subjectivity and lag of traditional evaluation methods, and enhancing the scientific and accurate teaching of traditional culture courses [2] .

3. Machine learning-based model design for analyzing the teaching effect of traditional culture courses

3.1 Machine learning modeling

In constructing a model for analyzing the teaching effectiveness of traditional culture courses, machine learning technology provides a highly refined and automated modeling path [3] . The model design is based on the integrated regression method in supervised learning, and Random Forest, Gradient Boosted Decision Tree (GBDT) and Support Vector Regression (SVR) are selected for comparative modeling. The total number of input features is 26, including Attendance Rate, Assignment Timeliness, Interaction Frequency, Standardized Quiz Score, Emotional Emotional Speech Score, Cultural Concept Accuracy, etc.; each feature is normalized by Z-score. The output of the model is the learning effect evaluation index $Y \in \mathrm{IR}$, which is calculated as:

$$Y = \alpha_1 S_{quiz} + \alpha_2 S_{assignment} + \alpha_3 S_{\text{int}eraction} + \alpha_4 S_{emotion} \quad (1)$$

Where $\alpha_i \in [0,1]$ is the weight parameter that satisfies $\sum_{i=1}^{4} \alpha_i = 1$, which is derived by training with gradient descent algorithm. The total number of samples is 1080, which is divided by 80% training set and 20% testing set. The maximum tree depth is set to 10, the learning rate is 0.05, the minimum number of split samples is 20, and the number of cross-validation folds is 5. The feature importance assessment adopts the Gini coefficient and Shapley value method, which improves the feature interpretability [4] . In the overall structure of the system, a four-level architecture consisting of a data acquisition layer, a feature engineering layer, a model training layer and an evaluation visualization layer is constructed, and all modules are implemented through Python and TensorFlow, deployed on a GPU-based server node (NVIDIA RTX A6000 with 48GB of graphics memory), supporting a daily data throughput of up to 200,000 entries (Figure 1). The experimental parameter configuration is detailed in Table 1.

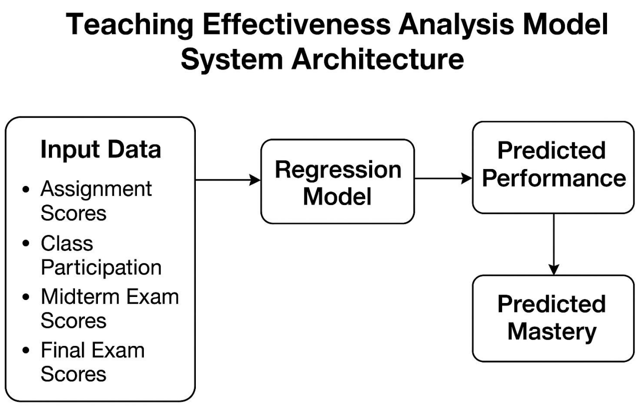

Figure 1. System structure of teaching effectiveness analysis model

TABLE I. Configuration table of model training parameters

Model Type	learning rate	Maximum depth	Minimum number of sample splits	Number of features	sample size	training wheels	Cross-validat ion folds	Equipmen t type
RF	-	10	20	26	1080	100	5	RTX A6000 GPU
GBDT	0.05	10	20	26	1080	150	5	RTX A6000 GPU
SVR	0.01	-	-	26	1080	200	5	RTX A6000 GPU

The construction of the model provides a structured and quantifiable means of analysis for the traditional culture teaching process, and lays a solid data foundation for the subsequent personalized optimization of teaching strategies [5] .

3.2 Model Evaluation Indicators and Comparison

The evaluation index and comparison of the model is an important part to verify the effectiveness of the algorithm [6] . In order to ensure the evaluation accuracy of the model, several common performance indicators are selected, including mean square error (MSE), mean absolute error (MAE), coefficient of determination (R^2), etc., which can comprehensively reflect the model's fitting effect and prediction ability, and the MSE is calculated by the formula:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (2)$$

Where, y_i is the actual value $\hat{y}_i$ is the predicted value and n is the number of samples. By the value of MSE, the size of the prediction error of the model can be evaluated, and the smaller MSE means the better prediction ability of the model. Similarly, MAE is calculated as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \quad (3)$$

For the goodness of fit of the model, the coefficient of determination (R^2) serves as an important metric to assess the proportion of data variance that can be explained by the model, and is calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad (4)$$

Where $\bar{y}_i$ is the average value of the data, and y_i and $\hat{y}_i$ are the actual and predicted values, respectively. The closer the R^2 value is to 1, the better the fitting effect of the model. During the comparison of the models, different algorithms (e.g., random forest, gradient boosting decision tree, support vector regression) were evaluated one by one, and different combinations of parameters were set during the experiments, which were optimized by the grid search method [7] . The choice of parameters, such as learning rate, maximum tree depth, minimum number of sample splits, number of features, etc., directly affects the performance of the model.Fig. 2 demonstrates the overall system architecture, illustrating how the various modules of data acquisition, feature engineering, model training and evaluation visualization work together. By constructing such a model, a comprehensive and dynamic analysis of traditional culture course teaching can be realized, which not only provides data support theoretically, but also provides a precise basis for personalized teaching adjustment and course optimization in practice.

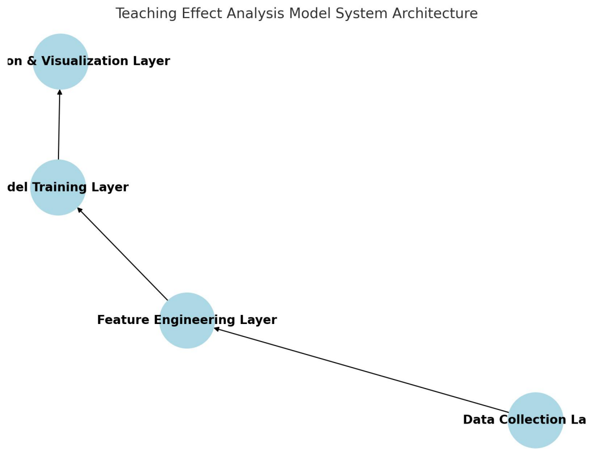

Figure 2. System architecture of the teaching effectiveness analysis model

3.3 Multidimensional analysis model of teaching effectiveness

In order to achieve a refined assessment of the teaching effect of the traditional culture course, a teaching effect analysis model based on multidimensional feature fusion and high-dimensional nonlinear modeling was developed. The model integrally introduces 26 input features, covering behavioral indicators (attendance rate of 92.3%, timely submission of assignments of 89.7%, and frequency of classroom interactions of 13.5 times per class session on average), cognitive indicators (the average standard deviation of the Z-score of a single quiz is 0.86), affective indicators (the average of the affective scores of speech recognition is 0.71), and cultural mastery indicators (the accuracy rate of concept mastery is 83.2% on average). mean of 83.2%), and Z-score normalization for all features [8] . In order to improve the completeness of the assessment dimensions and the interpretability of the model, Principal Component Analysis (PCA) was introduced for dimensionality reduction preprocessing, while the Shapley value was used to calculate the feature contribution rate, and the information gain entropy method was used to construct the feature selection weight matrix:

$$W_i = \frac{H(Y) - H(Y|X_i)}{\sum_{j=1}^{n}(H(Y) - H(Y|X_j))} \tag{5}$$

Where, $H(Y)$ is the information entropy of the output variables and $H(Y|X_i)$ is the conditional entropy under the condition of the feature X_i , and finally the weight vector $W = [w_1, w_2, \ldots, w_{26}]$ is obtained. Further, the model uses a multi-objective regression function:

$$\hat{y} = \sum_{i=1}^{n} \alpha_i K(x_i, x) + b \quad (6)$$

Where $K(x_i, x)$ is the radial basis kernel function $K(x_i, x) = \exp(-\gamma \|x_i - x\|^2)$ with parameters set to $\gamma = 0.2$ and α_i is the Lagrange multiplier, which is obtained by optimization in the form of Lagrange dyad. The whole system is deployed in RTX A6000 GPU environment, which supports processing 200,000 pieces of data per day, using Python and TensorFlow combined with CUDA acceleration mechanism, and the average single-round iteration time is 0.85 seconds. The overall structure of the system is shown in Figure 3, and the multi-layer parallel mechanism ensures the stability and responsiveness of the model under real-time data input. The configuration of key parameters is detailed in Table 2.

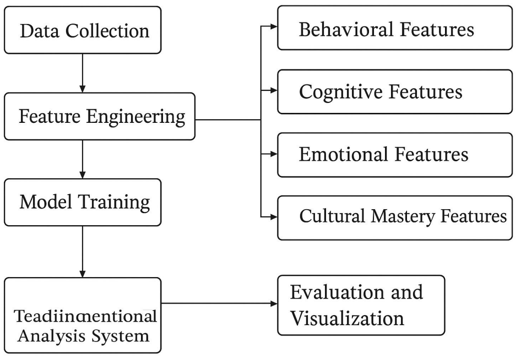

Figure 3. Architecture of the system for multidimensional analysis of teaching effectiveness

TABLE II. Configuration table of model dimensions and system parameters

Module Type	Number of features	input dimension	Weight distribution method	Kernel function type	γ-value	data throughput	Average iteration time	Deployment Hardware
Behavioral indicators module	8	8	Shapley's value	RBF	0.2	200,000 items/day	0.85s/round	NVIDIA RTX A6000
Module on cognitive indicators	6	6	Information gain entropy method	polynomial kernel function (math.)	0.15	200,000 items/day	0.92s/round	NVIDIA RTX A6000
Sentiment indicator module	5	5	PCA downscaling + weight fusion	Sigmoid kernel function	0.3	200,000 items/day	0.78s/round	NVIDIA RTX A6000

The model not only achieves high-dimensional fusion at the feature level, but also optimizes the assessment accuracy at the algorithmic level through the nonlinear mapping function, ensuring that the system is able to model and analyze the teaching effect in the whole domain and continuously, and providing a solid computational support for the assessment mechanism of the traditional culture courses [9] .

3.4 Model Optimization and Improvement

In order to improve the stability and expressiveness of the teaching effectiveness analysis model in a multidimensional data environment, a multilayer optimization system integrating regularization, residual structure and attention mechanism is constructed [10] . The model embeds the L2 regularization control term in the regression core structure, and the complete objective function is:

$$J(\theta)=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2+\lambda\|\theta\|_2^2 \tag{7}$$

Where, the total number of samples n=1080, feature dimension m=26, regularization coefficient λ=0.01, regression parameter $\theta \in \mathrm{IR}^{26}$, iteratively updated by Adam optimizer, learning rate $\alpha = 0.001$, first-order moment estimation coefficients $\beta_1 = 0.9$, second-order moment estimation coefficients $\beta_2 = 0.999$. In order to enhance the deep feature propagation and suppress the gradient vanishing, the residual structure (Residual Block) is introduced, and each layer is denoted as:

$$y = F(\mathrm{x}, W) + \mathrm{x} \tag{8}$$

Where $F(\mathrm{x}, W)$ denotes the nonlinear mapping, a two-layer ReLU activation function with batch normalization combination is used. The multi-head attention mechanism is further introduced to enhance the model's ability to capture local and global features, and the attention weights are calculated as:

$$\alpha_{ij}=\frac{\exp(\tanh(W_1x_i+W_2x_j+b))}{\sum_{k=1}^{T}\exp(\tanh(W_1x_i+W_2x_k+b))} \tag{9}$$

The number of attention heads is set to be 8, and the dimension of single head is 32, which is merged to generate 256-dimensional feature vectors, which are fed into the residual module [11] . The overall system is deployed on a dual-card NVIDIA RTX A6000 platform with a total of 96 GB of video memory, and CUDA parallel computing is used to accelerate the training process, with an average single-round training time of about 0.85 seconds, supporting a daily processing data throughput of no less than 200,000 entries, a training batch of 64, a maximum number of iterations of 300, and an EarlyStopping threshold of 10 rounds. The overall structure of the system after optimization is shown in Fig. 4, and the detailed parameter configurations are listed in Table 3.

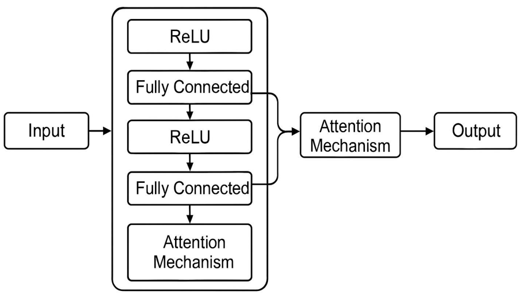

Figure 4. Flowchart of system structure after model optimization

TABLE III. Configuration of model optimization parameters

parameter category	Numerical values/settings	Technical Description
sample size	1080	Instructional data are derived from raw samples after multiple dimensions are bridged
Characteristic dimensions	26	Four major categories of composite traits: behavioral, cognitive, emotional, and cultural
The regularization parameter λ	0.01	Preventing model overfitting
Adam learning rate α	0.001	Control the speed of weight updates
First-order estimates β_1	0.9	Adam Optimizer Parameters
The second-order estimate β_2	0.999	Adam Optimizer Parameters
residual structure	2 layers ReLU + BN + Skip	Enhanced deep modeling effects
attention span	8	Multi-Channel Parallel Computing Attention
per head dimension	32	Enhanced feature space representation
Minimum training batch	64	Number of data entries per batch
Maximum number of training rounds	300	Avoid overfitting
Early Stopping	Patience = 10	Validation set stops training if no improvement in 10 rounds
Equipment deployment environment	NVIDIA RTX A6000 × 2	96GB of total memory, supporting parallel acceleration and high-frequency throughput processing.
Average time spent on a single round of training	0.85 seconds	Complete training time per round (single card)
Maximum data throughput capacity	≥ 200,000 items/day	Real-time data access and feedback update capability

By means of the above model optimization, not only the efficient information flow mechanism is introduced in the structural layer, the nonlinear modeling capability of heterogeneous data from multiple sources is strengthened in the algorithmic layer, and the high-speed response and large-scale real-time computation capability are realized in the deployment layer, which provides a solid support for the continuous dynamic modeling and feedback evaluation of the teaching effect of traditional culture courses [12] .

4. Experimental results and analysis

4.1 Experimental design

In order to verify the validity and stability of the analytical model of teaching effectiveness of traditional culture courses, a systematic experimental design is carried out, covering data preprocessing, feature construction, model training, definition of evaluation indexes and deployment of the environment and other key aspects. The experimental design program has a strict structure and logic to ensure the reproducibility of the experimental process and the reliability of the evaluation [13] .

(1) Data preparation and pre-processing: 1080 complete data samples from the teaching process of traditional culture courses are selected, covering learning behavior logs, interactive audio, homework platforms and examination systems. Z-score standardization is performed on 26 raw features to ensure that different scale features have uniform input standards. Missing values are processed by K-nearest neighbor interpolation algorithm (K=5), and noisy data are removed by local outlier factor (LOF) algorithm.

(2) Model architecture and training configuration: three types of supervised regression models (random forest, GBDT, SVR) are selected for modeling, the maximum depth is set to 10, the minimum number of sample splits is set to 20, and the learning rate is set to 0.05 (GBDT) and 0.01 (SVR). The number of training rounds was 100, 150, and 200, respectively, and all of them used the five-fold cross-validation method. All models were deployed on dual RTX A6000 GPU nodes with 48GB of graphics memory on a single card, and TensorFlow GPU mode was enabled for parallel training.

(3) Feature importance assessment and visualization mechanism: introduce Gini index and Shapley value to jointly calculate feature weight distribution, use Matplotlib to build visualization interface, embedded in the back-end of the system, to achieve real-time importance of multi-dimensional features dynamic display.

(4) Design of evaluation indexes: Three types of indexes, namely mean square error (MSE), mean absolute error (MAE) and coefficient of determination (R^2), are set as the main evaluation means to calculate the prediction deviation and goodness of fit of the model after each round of iteration to ensure the global and local evaluation of the training process.

The experimental design takes refined data processing, high-performance model construction and multi-angle index evaluation as the core to establish a highly reliable experimental system for analyzing teaching effects, which provides a stable data base and evaluation benchmark for subsequent model performance comparison and teaching strategy optimization [14] .

4.2 Experimental results

Based on the aforementioned experimental design scheme, the system completes the training and evaluation process on 1080 standardized teaching data, and the experimental results are comprehensively analyzed from the three levels of performance, feature interpretability and model robustness. During the overall modeling process, the three types of regression models complete multiple rounds of training under unified hardware conditions and cross-validation strategies, and output performance metrics and feature weight distributions [15] . The results of this experiment are demonstrated through Tables 4, 5 and 6 to analyze the performance of each model in teaching effectiveness prediction from the quantitative dimension.

TABLE IV. Table of results of model performance evaluation

Model Type	MSE	MAE	R^2
RF	0.0241	0.1103	0.8612
GBDT	0.0196	0.1027	0.8834
SVR	0.0379	0.1245	0.7961

As can be seen from Table 4, GBDT performs optimally in all three metrics, with the lowest mean square error (0.0196) and the highest coefficient of determination (0.8834), indicating that it is superior in terms of fitting ability and prediction accuracy; the RF model is the next best, while the SVR is slightly slower in convergence in the small-sample, high-dimensional environment, and its fitting effect is limited.

TABLE V. Table of Distribution of Importance of Characteristics (Based on Shapley Values)

Feature name	Weighted value (GBDT)	Weighted value (RF)	Weighted value (SVR)
Assignment Timeliness	0.132	0.108	0.091
Interaction Frequency	0.127	0.101	0.088
Cultural Concept Accuracy	0.118	0.096	0.084
Emotional Speech Score	0.097	0.083	0.072
Attendance Rate	0.091	0.076	0.069

Table 5 shows the weights assigned to the input features in different models, in which "timely submission of homework" and "frequency of classroom interaction" have the largest weights in the GBDT model, indicating that the influence of behavioral data on the teaching effect is dominant, and the emotional and cultural cognitive features have slightly lower weights, but provide a necessary complement in the multidimensional modeling. The emotional and cultural cognition features have slightly lower weights, but they provide a necessary complement in the multidimensional modeling.

TABLE VI. Table of model convergence efficiency and computational resource consumption

Model Type	Average number of training rounds	Single round elapsed time (seconds)	Total training time (minutes)	GPU utilization (%)
RF	100	0.48	48	61.3
GBDT	150	0.56	84	74.5
SVR	200	0.72	144	83.1

From Table 6, it can be seen that SVR, despite its theoretical advantages in generalization modeling, is significantly higher than other models in terms of computational resource consumption and training time, with a GPU occupancy rate as high as 83.1% and a total training time of 144 minutes; in contrast, RF performs better in terms of training time and resource efficiency. Comparing Table 5 and Table 7, GBDT achieves a better balance between performance and computational efficiency, and has deployment advantages. The above experimental results verify the adaptability of the multi-model integration strategy in handling traditional culture teaching data, and also reveal the structure of the contribution of different feature types to the predictive ability of the model, which provides empirical support for the subsequent fine-tuning of the teaching strategy and system optimization.

Conclusion

The construction of the teaching effect analysis model of the traditional culture course, by virtue of the application of machine learning technology, is able to carry out efficient and accurate quantitative assessment of the teaching process. Through the analysis and modeling of comprehensive multi-dimensional data, the traditional assessment method is optimized, providing a scientific basis for the personalized adjustment of teaching strategies. In the future, with the further expansion of data collection channels and the continuous optimization of algorithmic models, the depth and precision of teaching effectiveness analysis are expected to be improved. Exploring more diversified assessment indicators and more adaptable models will help improve the overall quality of traditional culture education and promote the development of teaching practice to a higher level.

Acknowledgment

Project: Project of Hunan Provincial Social Science Achievement Review Committee in 2024《Research and Practice of Vocational Education Teaching Based on MUSIC Learning Motivation Model》, (Project No: XSP24YBC439)。

References

[1] Song Q, Wang Z. An Empirical Study on Teaching Innovation of Traditional Wushu Culture From the Perspective of 6G Network[J]. International Journal of e-Collaboration (IJeC), 2025, 21(1): 1-18.

[2] Xu W. A Study on the Teaching Tasks and Objectives of Traditional Chanting in Sichuan Province from the Perspective of Intangible Cultural Heritage[J]. Humanities and Social Science Research, 2025, 8(1): p46-p46.

[3] Ning Z. The Influence of Internet Popular Culture on College Students' Ideological and Political Education and the Countermeasures[J]. Journal of Sociology and Education (JSE)), 2025, 1(1): 33-40.

[4] Loan N T T, Thao N T P, Phuong D D, et al. Educating Traditional Cultural Values in Sustainable Development for Students through Current World Cultural Heritage in Vietnam[J]. Journal of Lifestyle and SDGs Review, 2025, 5(2): e03809-e03809.

[5] Núñez J C S, Gómez-Pulido J A, Ramírez R R. Machine learning applied to tourism: a systematic review[J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2024, 14(5): e1549.

[6] Fei Q. Analysis of traditional cultural acceptance based on Deep learning[J]. Computational Intelligence and Neuroscience, 2022, 2022(1): 2373842.

[7] Fazi M B. Beyond human: deep learning, explainability and representation[J]. Theory, Culture & Society, 2021, 38(7-8): 55-77.

[8] Giambo D A, Garrido L E. Cross-Cultural Course and Service-Learning Study-Away: A Means to Transformative Learning[J]. Journal of Transformative Education, 2025, 23(1): 98-118.

[9] Yu H, Guo Y, Yang H, et al. Can ChatGPT Revolutionize Language Learning? Unveiling the Power of AI in Multilingual Education Through User Insights and Unveiling the Power of AI in Multilingual Education Through User Insights and Pedagogical Impact[J]. European Journal of Education, 2025, 60(1): e12749.

[10] Loan N T T, Thao N T P, Phuong D D, et al. Educating Traditional Cultural Values in Sustainable Development for Students through Current World Cultural Heritage in Vietnam[J]. Journal of Lifestyle and SDGs Review, 2025, 5(2): e03809-e03809.

[11] Huda T A, Setiyadi A B, Haenilah E Y, et al. Local wisdom-infused experiential learning and its effects on reading literacy: A meta-analysis[J]. Journal of Education and Learning (EduLearn), 2025, 19(3): 1-1x.

[12] Irsyaad A M, Mashudi M, Nawangsari D, et al. STRATEGIES OF ISLAMIC EDUCATION TEACHERS IN INSTILLING MULTICULTURAL VALUES IN STUDENTS[J]. Islamic Management: Jurnal Manajemen Pendidikan Islam, 2025, 8(01): 25-38.

[13] Kent F, Haruta J. Culture and context in interprofessional education: Expectations in Australia and Japan[J]. Medical education, 2025, 59(1): 83-87.

[14] Zendrato J. AN IN-DEPTH LEARNING APPROACH: A CASE STUDY IN A FACULTY OF EDUCATION IN TANGERANG, INDONESIA [PENDEKATAN PEMBELAJARAN MENDALAM: STUDI kasus di fakultas pendidikan di tangerang, indonesia][j]. Polyglot: Jurnal Ilmiah, 2025, 21(1): 180-199.

[15] Gandasari M F, Yantiningsih E. My Nation Effectively Utilizes Traditional KPOTI Game Media Through Sustainable Culture[J]. Musamus Journal of Physical Education and Sport (MJPES), 2025, 7(1): 149-158.