ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Adaptive Strategies for Ideological and Political Education in Higher Education to Students' Psychological Changes Based on Big Data Analysis

Xiaohui Fan[1,*]

[1] Marxism School of Zhengzhou University of Science and Technology, Zhengzhou, 450000, Henan, China

SUMMARY: *Students' psychological states shape how ideological and political education works in higher education, yet many existing approaches still depend on fixed teaching arrangements and relatively broad interventions. In practice, differences in emotion, engagement, and content acceptance are common, and they rarely remain unchanged across time. Under these conditions, static strategies are often insufficient. This paper proposes a data-driven adaptive framework that uses educational and behavioral data to identify psychological changes and adjust intervention decisions in time. The framework supports individualized intervention design, flexible content delivery, and response updating under changing student conditions. Rather than following a fixed intervention path, the proposed method makes decisions from evolving data signals and predicted outcomes. Experimental results show that the method achieves 91.02% accuracy on the Ideological Education Impact Dataset. This result indicates that the framework can align educational intervention more closely with students' psychological changes and provides a more practical basis for adaptive ideological and political education in higher education.*

KEYWORDS: *Ideological and Political Education; Higher Education; Psychological Changes; Big Data Analysis; Adaptive Intervention*

1 Introduction

Big data has become increasingly relevant to educational research, particularly in areas where student behavior and response need to be observed over time. This is especially true in higher education ideological and political education, where psychological change is usually gradual, uneven, and difficult to describe with a fixed teaching framework alone [1]. In practice, students differ in emotional response, classroom engagement, and acceptance of educational content, and these differences do not remain static. A teaching strategy that works for one group or at one stage may lose effectiveness under another condition [2]. Big-data-based analysis makes these variations easier to detect and gives educators more room to adjust interventions to actual student conditions instead of treating students as a homogeneous group [3]. What matters here is not only the content itself, but also the timing, form, and target of the intervention. Large-scale data can also expose recurring response patterns, including declining engagement, emerging risk signals, and slower shifts in psychological state [4]. Seen from this angle, big data is no longer just a supporting technique. It is becoming part of the practical basis for understanding and improving ideological and political education in higher education[5].

*fanfandexigua@163.com
https://doi.org/10.65102/is2026186
Published: 30 April, 2026

Early attempts to introduce analytical methods into ideological and political education were mostly built around structured and rule-based frameworks. These systems relied on predefined rules and explicit instructional guidelines, which made them relatively easy to deploy in teaching practice[6]. However, once student psychological states began to show temporal variation and individual divergence, such designs became difficult to sustain. The main issue was not the lack of structure, but the inability to adjust when student responses shifted across contexts or over time, a situation that can be described as path dependence in intervention design [7]. As a result, their effectiveness remained limited when applied to more heterogeneous student groups[8]. Later work moved toward more flexible computational approaches, where algorithms were used to extract patterns from educational data rather than follow fixed rules [9]. These methods improved the resolution at which student behavior and psychological tendencies could be observed [10]. Models built on this idea were able to respond to differences between students to some extent, which made personalized intervention more feasible in practice. At the same time, this line of work introduced new constraints [11]. Data preprocessing became a non-trivial part of the pipeline, and model performance often depended on feature engineering and domain knowledge. This reliance on manual design created a bottleneck that limited scalability in broader educational settings[12]. More recent studies have adopted neural network architectures to handle high-dimensional and heterogeneous educational data directly [13]. These models are capable of capturing interactions that are difficult to specify explicitly, especially when psychological signals are weak or indirect . In addition, pre-trained models have been used to transfer representations across tasks, which reduces the cost of building models from scratch [14]. Despite these advantages, two issues remain difficult to ignore [15]. One is the computational overhead required for training and deployment, and the other is the limited interpretability of model outputs in educational contexts [16]. When intervention decisions need to be explained to educators, these limitations become particularly visible, forming a trade-off that resembles a Pareto tension between accuracy and transparency [17].

Based on the limitations of the aforementioned methods, we propose a novel approach that leverages the strengths of big data analysis to develop adaptive strategies for ideological and political education. Our method addresses the rigidity of symbolic AI, the data dependency of machine learning, and the computational demands of deep learning by integrating these approaches into a cohesive framework. By doing so, we aim to create a more flexible, efficient, and scalable solution that can effectively respond to the psychological changes of students in higher education. This approach not only enhances the adaptability of educational strategies but also ensures that they are grounded in empirical data, thereby improving their overall effectiveness and impact. We summarize our contributions as follows:

- This study develops a data-driven method for linking students' psychological changes with adaptive intervention design in ideological and political education.
- Instead of relying on fixed teaching rules, the method updates intervention decisions from behavioral and educational data, which improves its suitability for changing classroom conditions and diverse student responses.
- Across the evaluated datasets, the proposed framework maintains stable performance. In particular, it achieves 91.02% accuracy on the Ideological Education Impact Dataset, showing its practical value for matching intervention strategies to students' psychological states.

2 Method

2.1 Overview

The interaction between ideological and political education and students' psychological change in higher education cannot be adequately explained by a static framework alone. Student response is not fixed: it shifts over time, differs across educational settings, and is influenced by both the content of instruction and the way that content is delivered. To deal with this complexity, this study develops a data-driven adaptive framework for analyzing psychological variation and supporting more responsive intervention design. Section 2.2 sets up the theoretical and mathematical basis of the study. It clarifies the main concepts used to describe psychological change in ideological and political education and formulates the problem in a way that can be modeled directly, rather than leaving it at a descriptive level. Section 2.3 then introduces the core analytical model. This part focuses on how student response is represented under different educational conditions, how behavioral and educational events are identified, and how likely outcomes are inferred from observed and historical data. The purpose is to provide a basis for intervention design that is more targeted than fixed rule-based schemes. Section 2.4 turns to practical use of the model. The emphasis here is on adaptive content adjustment and intervention design under uncertainty. Educational responses are updated according to predicted student state, while uncertainty is incorporated into the decision process instead of being treated as noise.

2.2 Preliminaries

This section formalizes the problem of adaptive strategies for ideological and political education in higher education, emphasizing the psychological dynamics of students analyzed through big data. The aim is to construct a mathematical framework that integrates big data analytics into educational strategies, enhancing their adaptability and effectiveness. Let $\mathcal{S}$ denote the set of students, where each student $s \in \mathcal{S}$ is characterized by a set of psychological attributes $\mathbf{P}_s = \{p_1, p_2, \dots, p_n\}$. These attributes are dynamic and subject to change under the influence of educational interventions. The set of interventions is represented by $\mathcal{I}$, where each intervention $i \in \mathcal{I}$ has a measurable impact on the psychological state of the students. The educational environment is modeled as a dynamic system, with its state at time t represented by a vector $\mathbf{X}_t$. This vector encapsulates both the psychological states of the students and external factors influencing these states. The system transitions from state $\mathbf{X}_t$ to $\mathbf{X}_{t+1}$ according to a function $f: \mathbb{R}^n \times \mathcal{I} \to \mathbb{R}^n$, defined as:

$$\mathbf{X}_{t+1} = f(\mathbf{X}_t, i_t), \tag{1}$$

where i_t represents the intervention applied at time t.

To account for the complexity of psychological changes, a manifold $\mathcal{M} \subset \mathbb{R}^n$ is introduced to represent the feasible psychological states of the students. This manifold imposes constraints ensuring that transitions remain within realistic bounds:

$$\mathbf{X}_{t+1} \in \mathcal{M}. \tag{2}$$

The objective is to optimize the sequence of interventions $\{i_1, i_2, \dots, i_T\}$ over a time horizon T to maximize a utility function $U: \mathcal{M} \to \mathbb{R}$, which quantifies the effectiveness of the ideological and political education. The optimization problem is formulated as:

$$\max_{i_1, i_2, \ldots, i_T} \sum_{t=1}^{T} U\left(\mathbf{X}_t\right), \tag{3}$$

subject to the constraints:

$$\mathbf{X}_{t+1} = f(\mathbf{X}_t, i_t), \quad \mathbf{X}_t \in \mathcal{M}, \quad i_t \in \mathcal{I}. \tag{4}$$

To address the uncertainty inherent in psychological changes, a probabilistic model is incorporated. Let $\mathcal{P}(\mathbf{X}_{t+1} \mid \mathbf{X}_t, i_t)$ denote the probability distribution over the next state given the current state and intervention. The expected utility is expressed as:

$$\mathbb{E}[U(\mathbf{X}_{t+1})] = \int_{\mathcal{M}} U\left(\mathbf{X}_{t+1}\right) \mathcal{P}(\mathbf{X}_{t+1} \mid \mathbf{X}_t, i_t)\, d\mathbf{X}_{t+1}. \tag{5}$$

Here, T is the intervention horizon, $\mathbf{X}_t \in \mathcal{M}$ denotes the student psychological state at time step t, $i_t \in \mathcal{I}$ denotes the intervention, U is the utility function, f is the state transition function, and $\mathcal{P}(\mathbf{X}_{t+1} \mid \mathbf{X}_t, i_t)$ denotes the conditional transition distribution.

2.3 Counterfactual Event Router

Figure 1 presents the overall structure of the Counterfactual Event Router. The model is used to connect student psychological variation with adaptive intervention design in ideological and political education. Instead of describing student response through a fixed instructional pipeline, the router treats psychological change as a dynamic process influenced by educational content, behavioral signals, and contextual conditions. The core task of the model is to represent student state, segment relevant events from educational data, and estimate which intervention is more suitable under the current condition.

For each student sample at time step t, we denote the multimodal input by

$$\mathbf{z}_t = [\mathbf{u}_t; \mathbf{b}_t; \mathbf{c}_t], \tag{6}$$

where $\mathbf{u}_t$ denotes the psychological attribute vector, $\mathbf{b}_t$ denotes observed behavioral features, and $\mathbf{c}_t$ denotes educational context features such as content type, delivery mode, or interaction intensity. Based on $\mathbf{z}_t$, the model learns a latent state representation

$$\mathbf{h}_t = f_\theta(\mathbf{z}_t), \tag{7}$$

where $f_\theta(\cdot)$ is the representation function parameterized by θ. This latent state is used as the shared basis for event segmentation, state transition estimation, and intervention routing.

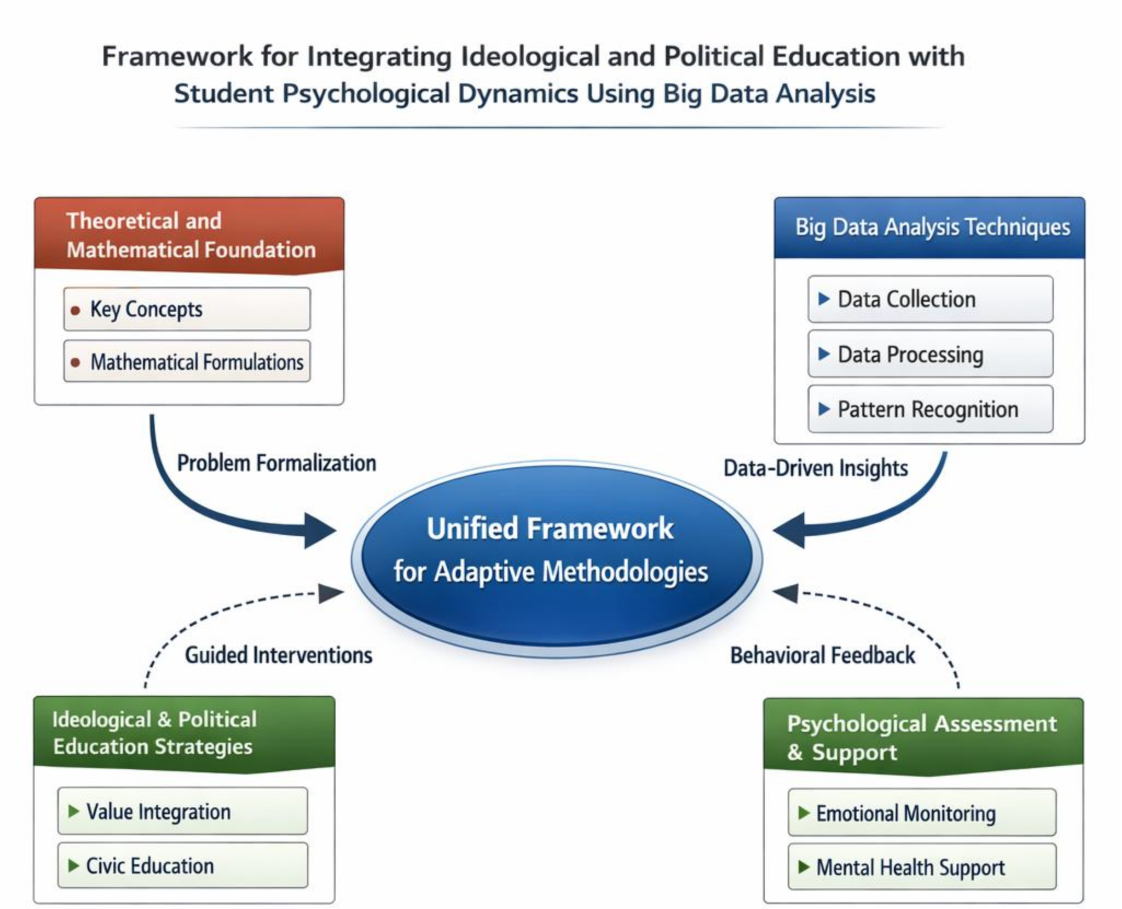

Figure 1: Overall structure of the Counterfactual Event Router. The framework links student psychological state, behavioral data, educational context, and intervention design within a unified analytical process. It combines state representation, event extraction, outcome estimation, and adaptive response selection to support data-driven ideological and political education in higher education.

Mathematical Formulation of Manifold Optimization: The first component constrains the latent representation to remain in a feasible psychological state space. This is used to reduce unstable projections caused by noisy behavioral observations and to preserve local structure in student data. Let $\mathcal{M}$ denote the manifold of valid student states. The projection of an observed sample onto $\mathcal{M}$ is written as

$$\min_{\mathbf{x}\in\mathcal{M}} \| \mathbf{x} - \mathbf{x}_0 \|^2 + \lambda \cdot \mathrm{Reg}(\mathbf{x}), \tag{8}$$

where $\mathbf{x}_0$ is the initial observation, λ is the regularization coefficient, and $\mathrm{Reg}(\mathbf{x})$ enforces smoothness and structural constraints on the manifold.

To preserve neighborhood consistency, the regularization term can be further written as

$$\mathrm{Reg}(\mathbf{x}) = \sum_{i,j} w_{ij} \| \mathbf{x}_i - \mathbf{x}_j \|^2, \tag{9}$$

where w_{ij} measures the affinity between samples i and j. A larger value of w_{ij} indicates that the two samples should remain close after projection. This term is introduced to keep psychologically similar students locally coherent in the latent space.

The corresponding state update under intervention a_t is defined as

$$\mathbf{h}_{t+1} = g_\phi(\mathbf{h}_t, a_t, \mathbf{c}_t), \tag{10}$$

where $g_\phi(\cdot)$ models how student state changes after an educational action is applied. In this way, manifold-constrained representation and state transition are handled within the same analytical process.

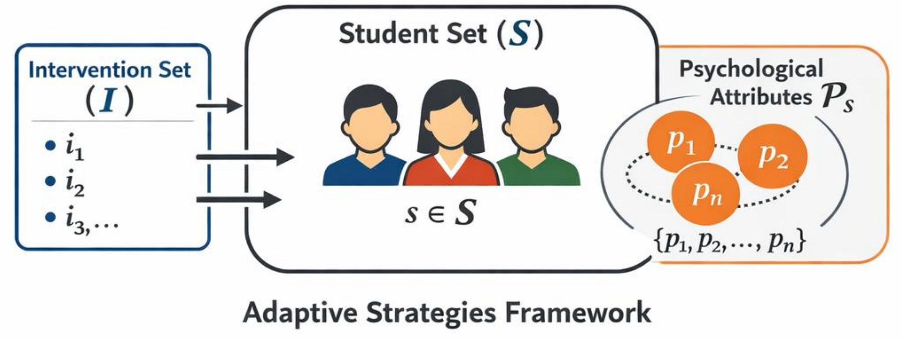

Figure 2: Manifold-constrained representation of student state and intervention relation. Students are modeled as individuals with dynamic psychological attributes, while educational interventions act on these states under contextual constraints. The objective is to maintain a feasible latent state space and support later event routing and intervention selection.

Parameterized Event Segmentation: After latent state construction, the model identifies educational events that may affect subsequent psychological change. Given an observed sequence $\{\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_T\}$, the event segmentation process is written as

$$E = \{e_k \mid e_k = \mathrm{Seg}(\mathbf{h}_{t_1:t_2}; \Theta)\}, \tag{11}$$

where $\mathrm{Seg}(\cdot)$ denotes the segmentation operator with parameter set Θ, and e_k denotes a detected event segment. Each event is represented by

$$\mathbf{e}_k = \psi(\mathbf{h}_{t_1:t_2}, \mathbf{c}_{t_1:t_2}), \tag{12}$$

where $\psi(\cdot)$ aggregates state features and contextual information within the segment. This allows the model to retain not only temporal variation in student state but also the instructional conditions under which that variation appears.

To score the relevance of an event for intervention routing, we define

$$r_k = \mathbf{v}^\top \tanh(\mathbf{W}_e \mathbf{e}_k + \mathbf{W}_h \mathbf{h}_t), \tag{13}$$

where r_k is the event relevance score and $\mathbf{v}$, $\mathbf{W}_e$, and $\mathbf{W}_h$ are learnable parameters.

Outcome Forecasting: The final module predicts the likely outcomes of different educational strategies from the segmented events. It uses a probabilistic model to estimate the likelihood of each outcome under the observed event pattern, providing a quantitative basis for strategy selection. The conditional probability of outcome y is defined as

$$P(y \mid \mathcal{E}, \mathbf{w}) = \frac{\exp(\mathbf{w}^\top \phi(\mathcal{E}, y))}{\sum_{y' \in \mathcal{Y}} \exp(\mathbf{w}^\top \phi(\mathcal{E}, y'))}. \tag{14}$$

Here, $\mathcal{E}$ denotes the segmented event set, $\mathbf{w}$ denotes the model parameter vector, $\phi(\mathcal{E}, y)$ denotes the feature representation associated with event set $\mathcal{E}$ and outcome y, and $\mathcal{Y}$ denotes the set of candidate outcomes.

Based on this distribution, the expected educational utility of a strategy can be written as

$$\mathbb{E}[U \mid \mathcal{E}] = \sum_{y \in \mathcal{Y}} U\,(y)\,P(y \mid \mathcal{E}, \mathbf{w}), \tag{15}$$

where $U(y)$ denotes the utility associated with outcome y. This term measures the expected effectiveness of the strategy under the current event configuration.

The predicted outcome is given by the maximum a posteriori decision

$$y^* = \underset{y \in \mathcal{Y}}{\operatorname{argmax}} P(y \mid \mathcal{E}, \mathbf{w}). \tag{16}$$

If several intervention strategies $a \in \mathcal{A}$ are available, the optimal strategy can be selected by

$$a^* = \underset{a \in \mathcal{A}}{\operatorname{argmax}} \sum_{y \in \mathcal{Y}} U\,(y, a)\,P(y \mid \mathcal{E}, a, \mathbf{w}), \tag{17}$$

where $U(y, a)$ denotes the utility of applying strategy a when the resulting outcome is y.

To train the forecasting model, we minimize the negative log-likelihood over the training set $\{(\mathcal{E}^{(n)}, y^{(n)})\}_{n=1}^{N}$:

$$\mathcal{L}(\mathbf{w}) = -\sum_{n=1}^{N} \log P\big(y^{(n)} \mid \mathcal{E}^{(n)}, \mathbf{w}\big). \tag{18}$$

To reduce overfitting, an ℓ_2 regularization term can be added:

$$\mathcal{L}_{\text{reg}}(\mathbf{w}) = -\sum_{n=1}^{N} \log P\big(y^{(n)} \mid \mathcal{E}^{(n)}, \mathbf{w}\big) + \lambda \parallel \mathbf{w} \parallel_2^2, \tag{19}$$

where $\lambda \geq 0$ is the regularization coefficient.

To quantify predictive uncertainty, the entropy of the outcome distribution is further computed as

$$H(\mathcal{E}) = -\sum_{y \in \mathcal{Y}} P\,(y \mid \mathcal{E}, \mathbf{w}) \log P(y \mid \mathcal{E}, \mathbf{w}). \tag{20}$$

A larger value of $H(\mathcal{E})$ indicates greater uncertainty in the predicted outcome, which can be used to trigger more cautious or conservative intervention decisions.

Targeted Educational Interventions Based on Uncertainty Metrics

Figure 3 illustrates how the Counterfactual Event Router is used at the intervention stage. At this point, the task is no longer limited to predicting student state. The model also needs to decide how educational content should be adjusted and how intervention strength should vary when prediction uncertainty is non-negligible. In ideological and political education, student response rarely remains stable across time or context. A delivery plan that appears suitable at one stage may become less effective at another. For this reason, the strategy developed here combines counterfactual content optimization with uncertainty-aware intervention selection.

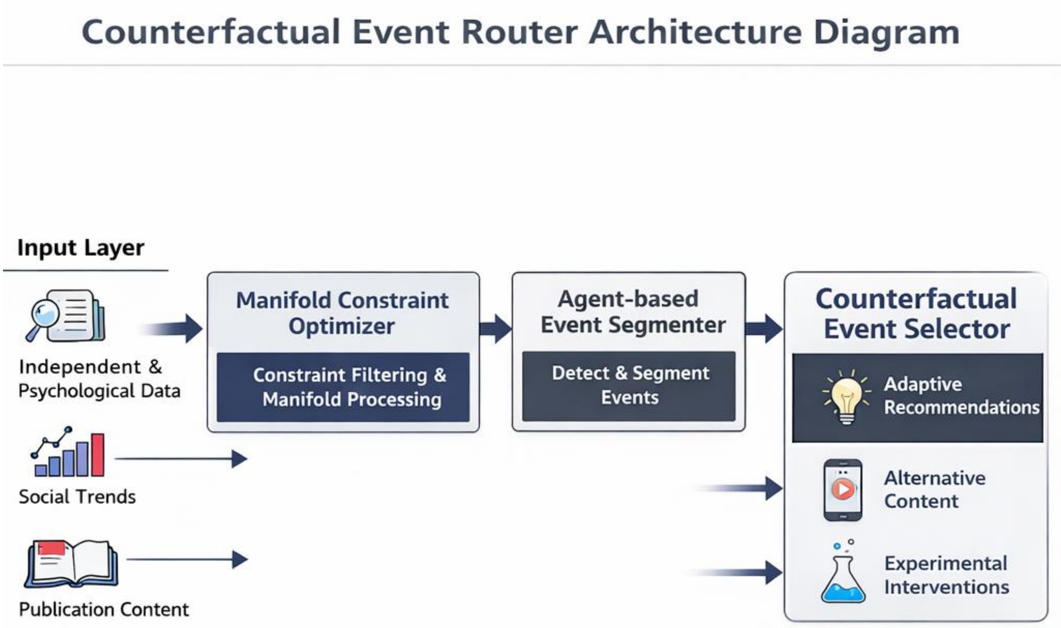

Figure 3: Operational structure of the Counterfactual Event Router in intervention design. The framework connects manifold-constrained state representation, event segmentation, and intervention selection. Student psychological data are first projected into a stable latent space, then segmented into relevant educational events, and finally evaluated for adaptive response selection.

Counterfactual Content Optimization: This component adjusts educational content according to predicted student state and plausible alternative response paths. Instead of assuming that one delivery scheme is equally appropriate for all students, the model compares candidate content choices under observed and counterfactual conditions. Let S_t denote the student state at time t, C_t the delivered educational content, and $\mathbf{X}_t$ the contextual factors affecting student response. The observed state transition is defined as

$$S_{t+1} = f(S_t, C_t, \mathbf{X}_t), \tag{21}$$

where $f(\cdot)$ captures the joint effect of prior state, content, and context.

For a candidate content choice $C_t{}'$, the corresponding counterfactual state is estimated by

$$\hat{S}_{t+1}^{(C_t{}')} = f(S_t, C_t{}', \mathbf{X}_t). \tag{22}$$

This term estimates how student state would change if an alternative content unit were delivered under the same contextual condition. When several content candidates are available, the preferred choice is obtained by maximizing expected utility:

$$C_t^* = \underset{C \in \mathcal{C}}{\operatorname{argmax}} \; \mathbb{E}\left[U\left(\hat{S}_{t+1}^{(C)}, C\right) \mid S_t, \mathbf{X}_t\right], \tag{23}$$

where $\mathcal{C}$ denotes the candidate content set and $U(\cdot)$ measures the instructional utility associated with a content decision.

To make content selection more sensitive to engagement change, we further define an engagement-aware score:

$$J(C) = \alpha\, U\left(\hat{S}_{t+1}^{(C)}, C\right) + (1-\alpha)\, G\left(\hat{S}_{t+1}^{(C)}\right), \tag{24}$$

where $G(\cdot)$ denotes the predicted engagement gain and $\alpha \in [0,1]$ controls the trade-off between instructional utility and short-term response quality. The final delivery decision is then written as

$$C_t^* = \underset{C \in \mathcal{C}}{\operatorname{argmax}} J(C). \tag{24}$$

This step allows the model to compare alternative content paths before delivery rather than updating content only after response has already deteriorated.

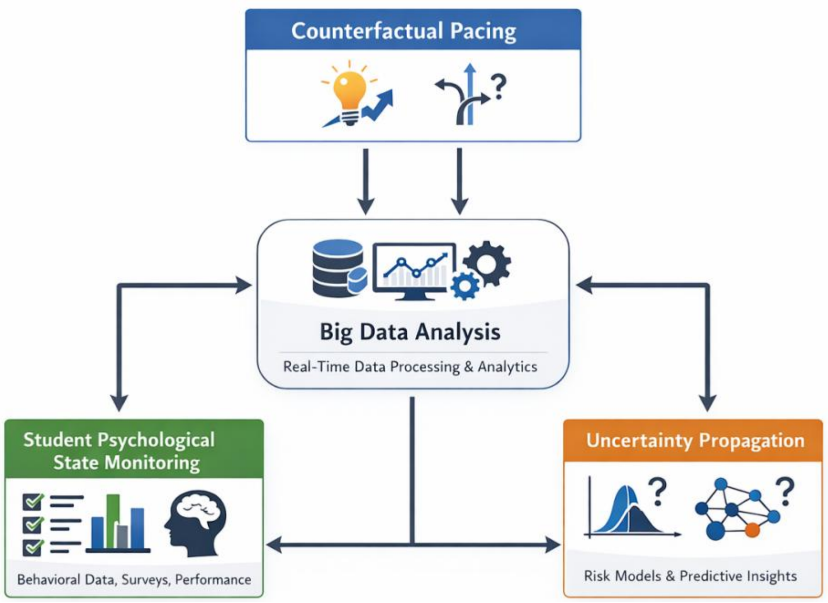

Figure 4: Integrated strategy for adaptive educational response. Counterfactual content optimization proposes alternative delivery choices, while uncertainty-aware intervention design adjusts response strength according to prediction confidence. Behavioral data, survey signals, and performance indicators jointly support dynamic educational adaptation.

Targeted Educational Interventions Based on Uncertainty Metrics: Prediction alone is not sufficient when student response is ambiguous. The second component therefore uses uncertainty estimates to control intervention strength and timing. Let

$$P(S_{t+1} \mid S_t, C_t, \mathbf{X}_t) \tag{25}$$

denote the predictive distribution of the next student state. A simple uncertainty estimate can be given by the predictive variance

$$\sigma^2(S_{t+1}) = \mathbb{E}[(S_{t+1} - \mathbb{E}[S_{t+1}])^2]. \tag{26}$$

A larger variance indicates that the model is less certain about the likely student response under the current intervention setting.

For categorical outcome prediction, uncertainty can also be measured by entropy:

$$\mathcal{H}_t = -\sum_{y \in \mathcal{Y}} P\,(y \mid S_t, C_t, \mathbf{X}_t) \log P(y \mid S_t, C_t, \mathbf{X}_t). \tag{27}$$

Compared with variance, entropy is more suitable when the target state is represented by discrete categories. In practice, the intervention policy can be written as

$$a_t = \begin{cases} a_t^{\text{direct}}, & \mathcal{H}_t < \tau_1, \\ a_t^{\text{moderate}}, & \tau_1 \leq \mathcal{H}_t < \tau_2, \\ a_t^{\text{conservative}}, & \mathcal{H}_t \geq \tau_2, \end{cases} \tag{28}$$

where τ_1 and τ_2 are uncertainty thresholds. This policy allows the system to avoid overly aggressive intervention when the predicted student state is not yet stable.

To connect uncertainty with educational action more directly, the intervention gain can be written as

$$R(a_t) = U(a_t) - \beta\, \mathcal{H}_t, \tag{29}$$

where $U(a_t)$ denotes the expected benefit of intervention a_t, and $\beta > 0$ penalizes uncertain decisions. The selected intervention is then

$$a_t^* = \underset{a \in \mathcal{A}}{\operatorname{argmax}} R(a). \tag{30}$$

Integrated Strategy for Dynamic Adaptation: The two components are optimized jointly rather than applied as separate stages. Counterfactual content optimization determines which content is more suitable under alternative response scenarios, whereas uncertainty-aware intervention design controls how strongly and how quickly the system should respond. This joint design is useful in practice because a content choice that looks optimal under point prediction may become less reliable when predictive uncertainty is high.

The overall objective is written as

$$\min_{C_t, a_t} \mathcal{L}(C_t, a_t, S_t) + \lambda_1 \sigma^2(S_{t+1}) + \lambda_2 \mathcal{H}_t, \tag{31}$$

where $\mathcal{L}(C_t, a_t, S_t)$ measures the discrepancy between desired and predicted student response, λ_1 controls the variance penalty, and λ_2 controls the entropy penalty. Under this formulation, content selection and intervention strength are determined together, so the decision process is influenced not only by predicted outcome but also by predictive confidence.

For sequential decision settings, the cumulative objective over a horizon T is defined as

$$\min_{\{C_t, a_t\}_{t=1}^{T}} \sum_{t=1}^{T} [\mathcal{L}(C_t, a_t, S_t) + \lambda_1 \sigma^2(S_{t+1}) + \lambda_2 \mathcal{H}_t]. \tag{32}$$

This objective is more suitable for educational processes in which student state changes gradually and intervention effects may accumulate over time.

3 Experimental Setup

3.1 Dataset and Data Preprocessing

3.1.1 Datasets

The study utilizes four public datasets for student behavior and ideological analysis. The Student Psychological Response Dataset (Wei Jiang 2024) is a public survey-based dataset containing 10,240 samples, 3 psychological states, and 22 variables including stress level, motivation, engagement, and demographic attributes. Labels were assigned based on standardized psychological scoring scales by trained annotators following a fixed protocol. Samples with more than 20% missing responses, survey completion time below 90 seconds, or duplicated student IDs were removed. The dataset was split randomly into training, validation, and test sets with a ratio of 8:1:1 under student-level isolation. The Ideological Education Impact Dataset (L. Yu, Chen, and Wang 2021) is a public longitudinal dataset with 8,960 paired pre- and post-intervention samples, 3 categories, and 20 variables including ideological alignment, critical thinking, and social attitude metrics. Labels were defined using score differences between pre- and post-tests under a fixed threshold rule. Records lacking either stage or showing inconsistent responses were filtered out. A chronological 7:1:2 split was applied to preserve temporal consistency. The Higher Education Political Curriculum Dataset (Na Chang 2024) is a public structured dataset with 9,120 samples, 4 categories, and 26 variables including curriculum content, course materials, and student performance indicators. Labels were annotated by domain experts according to a unified rubric. Samples with incomplete course records, missing performance scores, or duplicated course entries were removed. A stratified 8:1:1 split was adopted at the course level. The Big Data Analysis of Student Ideological Trends dataset (Zhang Qing 2024) is a public large-scale dataset containing 12,800 samples, 3 categories, and 24 variables derived from social media activity, survey responses, and academic records. Labels were generated from composite ideological scores using deterministic ranking rules. Samples with missing key features, abnormal response patterns, or duplicated timestamps were filtered out. The dataset was divided into training, validation, and test sets with a ratio of 8:1:1 under user-level isolation.

3.1.2 Data Preprocessing

A unified preprocessing pipeline was applied to the SPRD, IEID, HEPCD, and BDIST datasets (Wei Jiang 2024; L. Yu, Chen, and Wang 2021; Na Chang 2024; Zhang Qing 2024). Data cleaning removed duplicate records based on identical student identifiers and timestamps, then discarded samples with missing values in key variables, and excluded abnormal entries using fixed rules. Specifically, samples with more than 20% missing attributes, survey completion time below 90 seconds, or zero-variance responses were removed. Numerical features were normalized to zero mean and unit variance using statistics computed from the training set. Ordinal variables were aligned to consistent ascending scales to ensure semantic consistency. Categorical features were encoded into fixed indices after identity alignment, and rare categories with fewer than 50 samples were merged into a single category. Textual and descriptive fields were tokenized into sequences with a maximum length of 128. For longitudinal records, sequences were truncated or padded to a fixed length of 16 with a stride of 4. Missing non-critical numerical values were filled with the median, while categorical gaps were filled with the most frequent category. No external feature extraction models were used, and all inputs were derived directly from the original data. Class imbalance was handled using inverse-frequency weighting based on the training set distribution.

3.2 Implementation Details

All experiments were run on workstations equipped with NVIDIA A100 GPUs, and implemented in PyTorch. The training adopted fixed configurations to ensure stability across datasets. As shown in Table 1, stochastic gradient descent with momentum, cosine annealing algorithm with 5 epochs of warm-up and mixed precision training methods were adopted. The model was trained with a batch size of 64, for a total of 100 epochs, and relied on validation performance to trigger the early stopping mechanism. This arch is set up for joint modeling of text input and structured attributes. The text sequence is encoded by the encoder of Transformer, and the structured features are projected by a two-layer multi-layer perceptron (MLP). After the two branches are fused, they are passed to the classification head for prediction. The main architecture settings include embedding size, encoder depth and feature fusion configuration. To ensure fairness, all baselines and this method use the same data partition, preprocessing and computing environment for training/evaluation.Hyperparameters are selected on the validation set, and the final values are reported on the test set.

Table 1: Summary of implementation and model configuration details.

Category	Configuration
Hardware	Intel Xeon Silver 4314 CPU, 4×A100 (40GB), 256GB RAM
Framework	PyTorch 2.1.0, CUDA 11.8, cuDNN 8.9
Epochs	100
Batch size	64
Optimizer	SGD (momentum=0.9)
Learning rate	0.01 with cosine annealing + 5-epoch warm-up
Weight decay	5×10^{-4}
Precision	Mixed precision training
Seed	42
Early stopping	Patience = 10
Text embedding	256-dim, max length 256
Transformer	4 layers, 8 heads, hidden dim 256, FFN 1024
Structured features	MLP, 2 layers, 64-dim output
Fusion	Concatenation (320-dim)
Classifier	FC layer, hidden dim 128
Dropout	0.1
Loss	Cross-entropy with label smoothing 0.1

3.3 Comparison with SOTA Methods

In this section, we compare the proposed method with several baseline models on four datasets. As reported in Table 2 and Table 3, the proposed method ranks first on all four datasets and remains consistently strong across Accuracy, Recall, F1 Score, and AUC. This result suggests that the model does not gain on one metric by sacrificing another, but maintains a relatively balanced performance profile. On the Student Psychological Response Dataset, the proposed method achieves 89.67% Accuracy, compared with 88.56% for the strongest baseline. On the Ideological Education Impact Dataset, the gap becomes slightly larger, with Accuracy increasing from 89.78% to 91.02%. This pattern is important because it shows that the advantage of the model is not limited to a single task. It becomes more visible when the task is more closely related to intervention effect and psychological change, which is also the central setting of this study. A similar tendency can be observed in Table 3. On the Higher Education Political Curriculum Dataset, Accuracy improves from 88.56% to 89.67%, while on the Big Data Analysis of Student Ideological Trends dataset it rises from 89.78% to 91.02%. The improvement is therefore present on both relatively structured data and more heterogeneous data, but it is more pronounced on the latter. This suggests that the proposed framework is better able to handle complex response patterns when textual and structured educational signals need to be modeled together.

Table 2: Performance comparison on the Student Psychological Response Dataset and the Ideological Education Impact Dataset.

Model	Student Psychological Response Dataset				Ideological Education Impact Dataset			
2-9	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
XLNet(Cao 2024)	87.45 ± 0.52	86.78 ± 0.63	86.12 ± 0.58	86.55 ± 0.47	88.67 ± 0.49	88.12 ± 0.60	87.34 ± 0.55	87.76 ± 0.50
Longformer(Wang 2025)	88.12 ± 0.44	87.56 ± 0.51	86.89 ± 0.53	87.23 ± 0.46	89.34 ± 0.42	88.79 ± 0.54	88.05 ± 0.57	88.47 ± 0.45
ERNIE(undefined and Lyu 2024)	86.98 ± 0.48	86.32 ± 0.59	85.67 ± 0.60	86.10 ± 0.52	88.01 ± 0.46	87.45 ± 0.62	86.68 ± 0.58	87.11 ± 0.53
DeBERTa("Pathways to Integrating Ideological and Political Education: A Case Study of the Information Management and Systems Major" 2024)	88.56 ± 0.39	88.01 ± 0.50	87.34 ± 0.49	87.78 ± 0.44	89.78 ± 0.37	89.23 ± 0.48	88.49 ± 0.52	88.92 ± 0.47
ELECTRA(Lu 2023)	87.89 ± 0.46	87.34 ± 0.57	86.67 ± 0.54	87.10 ± 0.49	89.12 ± 0.43	88.56 ± 0.55	87.82 ± 0.59	88.25 ± 0.51
MobileBERT(Yue et al. 2023)	87.23 ± 0.50	86.67 ± 0.61	85.98 ± 0.56	86.41 ± 0.53	88.45 ± 0.48	87.89 ± 0.63	87.12 ± 0.60	87.55 ± 0.54
Ours	**89.67 ± 0.37**	**89.12 ± 0.46**	**88.45 ± 0.42**	**88.89 ± 0.40**	**91.02 ± 0.35**	**90.47 ± 0.44**	**89.73 ± 0.46**	**90.16 ± 0.39**

Table 3: Performance comparison on the Higher Education Political Curriculum Dataset and the Big Data Analysis of Student Ideological Trends dataset.

Model	Higher Education Political Curriculum Dataset				Big Data Analysis of Student Ideological Trends			
2-9	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
XLNet(Cao 2024)	87.45 ± 0.52	86.78 ± 0.63	86.12 ± 0.58	86.45 ± 0.47	88.67 ± 0.49	88.12 ± 0.60	87.34 ± 0.55	87.68 ± 0.50
Longformer(Wang 2025)	88.12 ± 0.47	87.56 ± 0.54	86.89 ± 0.62	87.23 ± 0.44	89.34 ± 0.42	88.79 ± 0.57	88.05 ± 0.53	88.39 ± 0.46
ERNIE(undefined and Lyu 2024)	86.98 ± 0.55	86.34 ± 0.61	85.67 ± 0.59	86.01 ± 0.48	88.02 ± 0.51	87.45 ± 0.63	86.78 ± 0.56	87.12 ± 0.52
DeBERTa(Major 2024)	88.56 ± 0.43	88.01 ± 0.50	87.34 ± 0.57	87.67 ± 0.45	89.78 ± 0.40	89.23 ± 0.52	88.49 ± 0.59	88.83 ± 0.43
ELECTRA(Lu 2023)	87.89 ± 0.50	87.34 ± 0.58	86.67 ± 0.60	87.01 ± 0.46	89.12 ± 0.47	88.56 ± 0.61	87.89 ± 0.54	88.23 ± 0.49
MobileBERT(Yue et al. 2023)	87.23 ± 0.53	86.67 ± 0.59	85.98 ± 0.63	86.32 ± 0.51	88.45 ± 0.48	87.89 ± 0.55	87.12 ± 0.58	87.45 ± 0.52
Ours	**89.67 ± 0.40**	**89.12 ± 0.48**	**88.45 ± 0.52**	**88.78 ± 0.44**	**91.02 ± 0.37**	**90.56 ± 0.45**	**89.89 ± 0.49**	**90.23 ± 0.41**

3.4 Ablation Study

We further conduct an ablation study to examine the contribution of the main components in the Counterfactual Event Router. The results are reported in Table 4 and Table 5. Instead of showing a gain from only one module, the ablation results indicate that all three components contribute to the final performance, although their effects are not identical. A clear pattern in Table 4 is that removing any module causes performance to drop, and the full model remains the best setting on both datasets. On the Student Psychological Response Dataset, Accuracy decreases from 89.67% to 88.45% when manifold optimization is removed, while F1 Score falls from 88.45% to 87.23%. On the Ideological Education Impact Dataset, the full model reaches 91.02% Accuracy, whereas the version without manifold optimization drops to 89.78%. This suggests that manifold-constrained representation is important for maintaining stable student-state modeling. The other two modules also make non-negligible contributions, but their effects are somewhat different. Removing event segmentation leads to a moderate decline across metrics, which implies that event-level structuring helps the model organize educational signals more effectively. By contrast, removing outcome forecasting produces the smallest drop among the three ablations, yet the full model still performs better in every case. This means that the forecasting component does not act alone as the main source of improvement, but it still strengthens the final decision process when combined with the other modules. The same tendency appears in Table 5. On the Higher Education Political Curriculum Dataset, Accuracy decreases from 89.67% in the full model to 88.45% without manifold optimization. On the Big Data Analysis of Student Ideological Trends dataset, the full model achieves 91.02% Accuracy, compared with 89.67% when the manifold component is removed. The gap is smaller for the other two ablations, but the ranking remains unchanged. Taken together, these results indicate that manifold optimization provides the strongest structural support, while event segmentation and outcome forecasting further improve the model by refining how educational events are organized and how intervention outcomes are estimated.

Table 4: Ablation Study on Sentiment Analysis models using Student Psychological Response and Ideological Education Impact datasets

Model	Student Psychological Response Dataset				Ideological Education Impact Dataset			
2-9	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
w./o. Mathematical Formulation of Manifold Optimization	88.45 ± 0.48	87.89 ± 0.57	87.23 ± 0.54	87.67 ± 0.50	89.78 ± 0.45	89.23 ± 0.56	88.56 ± 0.52	88.99 ± 0.48
w./o. Parameterized Event Segmentation for Educational Insights	88.78 ± 0.42	88.23 ± 0.51	87.56 ± 0.49	87.98 ± 0.46	90.12 ± 0.39	89.67 ± 0.50	89.01 ± 0.47	89.44 ± 0.43
w./o. Outcome Forecasting in Adaptive Education Systems	89.12 ± 0.40	88.56 ± 0.49	87.89 ± 0.46	88.32 ± 0.44	90.45 ± 0.37	89.89 ± 0.48	89.23 ± 0.45	89.66 ± 0.41
Ours	**89.67 ± 0.37**	**89.12 ± 0.46**	**88.45 ± 0.42**	**88.89 ± 0.40**	**91.02 ± 0.35**	**90.47 ± 0.44**	**89.73 ± 0.46**	**90.16 ± 0.39**

Table 5: Ablation Study on Higher Education Political Curriculum Dataset and Big Data Analysis of Student Ideological Trends

Model	Higher Education Political Curriculum Dataset				Big Data Analysis of Student Ideological Trends			
2-9	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
w./o. Mathematical Formulation of Manifold Optimization	88.45 ± 0.48	87.89 ± 0.55	87.23 ± 0.60	87.56 ± 0.47	89.67 ± 0.44	89.12 ± 0.52	88.45 ± 0.57	88.78 ± 0.46
w./o. Parameterized Event Segmentation for Educational Insights	88.78 ± 0.46	88.23 ± 0.53	87.56 ± 0.58	87.89 ± 0.45	90.12 ± 0.42	89.67 ± 0.50	89.01 ± 0.54	89.34 ± 0.43
w./o. Outcome Forecasting in Adaptive Education Systems	89.01 ± 0.44	88.45 ± 0.51	87.78 ± 0.56	88.12 ± 0.43	90.45 ± 0.40	89.89 ± 0.48	89.34 ± 0.52	89.67 ± 0.41
Ours	**89.67 ± 0.40**	**89.12 ± 0.48**	**88.45 ± 0.52**	**88.78 ± 0.44**	**91.02 ± 0.37**	**90.56 ± 0.45**	**89.89 ± 0.49**	**90.23 ± 0.41**

4 Conclusions and Future Work

This study aims to align college ideological and political education with students' psychological changes. It uses a counterfactual event router model and relies on big data analysis to determine adaptation strategies. This model integrates a variety of optimized mathematical methods, performs parametric event segmentation to obtain education-related perceptions, and uses outcome prediction in adaptive systems to process data and observe students' psychological states. Experimental results show that these parts help customize educational interventions and improve the participation and effectiveness of ideological and political education. Through counterfactually optimized intelligent content delivery and targeted educational interventions based on uncertain indicators, adaptability is further enhanced, enabling education to adjust interventions and manage uncertainty based on predictive insights.Our method provides a path for promoting educated and compassionate education, helping students grow comprehensively.

There are two clear limitations in the current study. First is the data condition, the proposed framework relies on relatively rich behavioral and educational data, but not all institutions can acquire it equally. When data infrastructure is weak, quality is unstable or records are incomplete, this method may be more difficult to achieve consistent deployment. This also shows that the applicability of the framework is partially restricted by the local data environment. The second limitation exists in the description of psychological changes. The model has a general research state, but it is difficult to accurately reproduce financial conditions, as well as the fluctuations of individual psychological reactions. There are situations where changes are gradual, situations where there is continuous sensitivity, and situations where structural records cannot be relied upon to reflect well. The next step may be to add timely feedback signals and quality evidence at appropriate places, so as to simulate research responses in a more detailed and flexible manner. Future work should reduce the focus on the size of the experimental framework, and pay more attention to the important sensitivities and individual differences in real educational settings.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

Xiaohui Fan contributed to conceptualization, methodology, software, validation, formal analysis, investigation, data curation, draft writing, review and editing, visualization, supervision, and funding acquisition. The author has read and approved the published version of the manuscript.

References

[1] Cao, Huanying. 2024. “Exploration of the Dilemma and Countermeasures in Building Discourse Power for College Students’ Ideological and Political Education from the Perspective of the New Media.” *Yixin Publisher* 1 (3): 187–94. https://doi.org/10.59825/jcms.2024.1.3.187.

[2] Chen, Hongzhi, and Yameng Zhang. 2024. “Exploration of the Value of Current Ideological and Political Education in the Context of Great Ideological and Political Views.” *Journal of Art, Culture and Philosophical Studies* 1 (1). https://doi.org/10.70767/jacps.v1i1.50.

[3] Gao, Da, and Jiajie Cai. 2025. “Challenges and Strategies for Universities Students’ideological and Political Education in the Digital Economy Era.” *Journal of Modern Education and Culture*, May.

[4] Li, Yuan. 2025. “A Study on University Students’ Participation in College Ideological and Political Education from the Perspective of the "Big Ideological and Political Course".” *International Educational Research* 8 (2). https://doi.org/10.30560/ier.v8n2p117.

[5] Liu, Guangxin. 2021. “The Ways and Methods of Ideological and Political Education for Postgraduates.” *Advances in Educational Technology and Psychology* 5 (3): 80–87.

[6] Lu, Yang. 2023. “Reflection and Exploration of Ideological and Political Education in the Curriculum—a Case Study of the University in Shijiazhuang City.” *World Journal of Education and Humanities* 6 (1). https://doi.org/10.22158/wjeh.v6n1p13.

[7] Na Chang, Na Chang. 2024. “The Disappearance of ‘the Other World’ and the Establishment of ‘This World’: The Human Emancipation—the Inspiration of ‘a Contribution to the Critique of “Hegel’s Philosophy of Right”: Introduction’ on the Ideological and Political Education for College Students.” *Yixin Publisher* 1 (1): 19–25. https://doi.org/10.59825/jcms.2024.1.1.19.

[8] "Pathways to Integrating Ideological and Political Education: A Case Study of the Information Management and Systems Major." 2024. *Curriculum and Teaching Methodology* 7 (7). https://doi.org/10.23977/curtm.2024.070703.

[9] Qi, Z. 2022. "The Achievements,background and Intension——Literature Review of the Study on Political and Ideological Education of College Students," December.

[10] undefined, undefined, and Ying Lyu. 2024. "Practical Exploration of Ideological and Political Education Work of College Students in the New Era." *Contemporary Education and Teaching Research* 5 (6): 217–22. https://doi.org/10.61360/ bonicetr242016390604.

[11] Wang, Gege. 2025. "The Integration Strategy of Ideological and Political Education for College Students Based on Course-Based Ideological and Political Education." *Educational Innovation Research* 3 (2): 18–23. https://doi.org/10.18063/eir.v3i2.759.

[12] Wei Jiang, Wei Jiang. 2024. "Practical Exploration of Collaborative Construction of 'Ke-Cheng Sizheng' and 'Si-Zheng Ke-Cheng' Under the Background of Big Ideological and Political Education: Taking the Teaching of 'Situation and Policy' Course as an Example." *Yixin Publisher* 1 (1): 181–87. https://doi.org/10.59825/jcms. 2024.1.1.181.

[13] Xia, Yushi, and Yuhan Liu. 2024. "Interdisciplinary Research in Ideological and Political Education: Exploring the Communication Studies of Ideological and Political Education Through the Application of the "Online Soft Guidance Model"." *Educational Science Literature* 1 (8). https://doi.org/10.62662/kxwxz0108010.

[14] Yu, Hua, and Mingjun Jiang. 2024. "Research on the Path to Improve the Quality of Ideological and Political Courses Under the Background of Digital Education." *Educational Science Literature* 1 (5). https://doi.org/10.62662/kxwxz0105006.

[15] Yu, Liang, Yuting Chen, and Lianfang Wang. 2021. "Bibliometric Analysis of Research Literature on Ideological and Political Education by Curriculum Based on SATI." *Journal of Physics: Conference Series* 1852 (4): 42072. https://doi.org/10.1088/1742-6596/1852/4/042072.

[16] Yue, Shuaiting, Jie Wei, Haider Aziz, and Kwen Liew. 2023. "A Study on the Effectiveness of Self-Assessment Learning System of Ideological and Political Education for College Students." *Learning and Motivation* 84 (November): 101929. https://doi.org/10.1016/j.lmot.2023.101929.

[17] Zhang Qing, undefined. 2024. "Innovating Content and Methods of Ideological and Political Education in the Context of the New Era." *International Journal of Education and Humanities* 4 (4): 479–87. https://doi.org/10.58557/(ijeh).v4i4.270.