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

Collaborative modeling of ideological and political Education text topic identification and stance discrimination under the framework of multi-task learning

Jing Lin[1,*], *Zhen Chen*[2], *Jinlian Huang*[3] *and Huiling Hu*[4]

[1] School of Marxism, Xiamen Ocean Vocational College, Xiamen 361100,Fujian, China
[2] Student Affairs Office,Fujian Provincial Department of Education, Fuzhou 350003, Fujian, China
[3] Department of Rehabilitation Medicine,The First Hospital of Putian City, Putian 351100, Fujian, China
[4] Institute of Creative Design and Management, National Taipei University of Business, Taoyuan 324022, Taiwan

SUMMARY: *In order to solve the problems of the separation of topic identification and stance discrimination and the insufficient utilization of semantics in the analysis of ideological and political education texts, a collaborative modeling method under the framework of multi-task learning was constructed. The model is based on shared semantic coding, introduces topic prototype matching and feature fusion mechanism in the topic identification branch, and adds topic-guided attention and cross-task consistency constraint in the stance discrimination branch, so that the topic information and attitude information can be jointly optimized in a unified representation space. The experiment is carried out based on 58,000 ideological and political education texts, covering various scenarios such as course discussions, thematic learning experiences, policy interpretation texts and campus online comments. The results show that the perplexity of the topic recognition model is 638, the topic consistency is 0.731, and the silhouette coefficient is 0.517. The accuracy of stance discrimination model reaches 86.7%, and Macro-F1 is 85.3%, which is better than the comparison models. The research shows that this method can improve the ability of topic boundary identification and implicit stance capture of ideological and political education texts, and provide computational technical support for curriculum feedback analysis, network public opinion research and education governance intelligence.*

KEYWORDS: *multi-task learning; Ideological and political education text; Topic identification; Position discrimination*

1 Introduction

With the deepening of the digitalization process of ideological and political education in colleges and universities, the course discussion texts, online comment data, learning reflection materials and policy communication corpus continue to converge, and the research of ideological and political education is shifting from empirical judgment to data-driven computational analysis. How to accurately identify the topic focus from large-scale texts and further judge its stance orientation has become a key issue in intelligent education governance,

*linjing@xmoc.edu.cn
https://doi.org/10.65102/is2026217 Published: 30 April, 2026

public opinion guidance and teaching feedback. Topic identification can answer "what is the text mainly talking about?" and stance discrimination can answer "what is the attitude of the text towards the topic?" they seem to be different tasks, but in fact, they have a close semantic dependency in the context of ideological and political education. If we only extract the theme and ignore the position difference, it is easy to classify the texts with different value attitudes into the same analysis result. If stance classification is directly carried out without the topic background, it is easy to confuse the meaning of the same expression in different issues, which weakens the explanatory power and application effectiveness of the model.

The existing natural language processing research has formed a rich accumulation of technologies in topic modeling, stance detection, and multi-task learning. The text representation method based on pre-trained language model significantly improves the semantic encoding ability, so that the model can obtain more fine-grained information from the context. Neural topic models and embedded clustering methods show strong advantages in topic discovery, semantic organization and emerging topic tracking. However, when these methods are directly transplanted to the scene of ideological and political education text analysis, there are still several limitations that cannot be ignored. First, the ideological and political education texts have distinct characteristics of value orientation, object orientation and context dependence. The same word may carry different attitude meanings in different issues, and the general text classification model is often difficult to identify such "topic-stance" linkage changes. Second, some studies separate topic identification and stance discrimination into two independent processes, which is simple for engineering implementation, but cuts off the semantic clues that should be shared between the two tasks, so that topic information cannot effectively constrain stance inference, and stance features are difficult to reverse modify topic boundaries. Third, policy terms, theoretical concepts, value expression and implicit evaluation coexist in ideological and political education texts. If the model only relies on surface word frequency or single sentence vector representation, it is prone to problems such as coarse topic extraction, stance recognition deviation, and fuzzy category boundaries.

Based on this realistic background, this paper focuses on the task of topic identification and stance discrimination in ideological and political education text analysis, and constructs a collaborative modeling method under the framework of multi-task learning. This study no longer treats the two tasks as serial modules separated from each other, but takes the shared semantic coding as a unified entry, and introduces the topic feature fusion mechanism and stance discrimination constraint mechanism in the deep representation space, so that the model can complete the dual-task collaborative optimization on the basis of common semantics. Specifically, in the topic identification step, we use the shared encoder to extract the contextual semantic representation of the text, and combine the topic-related features to complete the aggregated representation, so as to enhance the model's ability to capture the topic boundaries and core concepts. In the stage of stance discrimination, the topic representation is used to guide the attitude judgment, so that when the model analyzes the stances such as support, opposition and neutral, it is no longer independent of the specific issue, but performs a more robust classification under the constraint of the target semantics.

This paper not only focuses on the improvement of classification indicators, but also pays more attention to the structural association between topic semantics and stance semantics in ideological and political education texts. By introducing a multi-task collaboration mechanism, the model can extract general semantics in parameter sharing and retain respective characteristics in task branches, thereby alleviating the problems of separation of topic and stance and insufficient information transmission in traditional methods. This paper attempts to demonstrate that using topic identification results as a priori support for stance understanding and acting stance signals in reverse on topic representation optimization can obtain more explanatory analysis results in complex educational contexts. This research has certain theoretical value and practical significance for the automatic understanding of ideological and political education texts, course feedback mining, network public opinion identification and education governance intelligence.

2 Related work

At present, the research on text topic identification and stance discrimination has formed two relatively clear technical lines in the field of natural language processing. One is centered on stance detection and emphasizes the ability of the model to identify the attitude orientations of support, opposition and neutral. The other focuses on topic modeling and focuses on topic discovery, topic aggregation and evolution analysis in semantic space. For ideological and political education texts, these two types of tasks are not separated from each other. The value judgment in the text is often attached to the specific issue, leaving the theme background, and the position label is easy to be inaccurate. However, if we only do topic clustering without dealing with attitude differences, it is difficult to reveal the structural distribution of different opinions on the same topic. Therefore, how to realize the collaborative modeling of topic identification and stance discrimination under the unified computing framework has become a problem worthy of in-depth discussion.

In the research of stance detection, Hardalov et al. [1] systematically reviewed the application of stance detection in false information identification, and pointed out that the target object, context dependence and argument structure jointly determine the difficulty of stance judgment. Alturayeif et al. [2] reviewed machine learning and deep learning technology, and argued that pre-trained language models have improved feature expression ability, but there are still shortcomings in implicit stance, cross-domain transfer and target dependence modeling. Focusing on multi-task learning, Chai et al. [3] improved the information flow between different tasks through the task interaction network, indicating that shared representation and task collaboration can effectively improve the stance detection performance. He et al. [4] tried to introduce Wikipedia background knowledge to enhance the semantic understanding of the target object. Cheng et al. [5] constructed a large-scale comprehensive argument mining dataset, which provided a data basis for the joint research of stance and argument related tasks. Li et al. [6] promoted the research to goal-stance extraction in open scenes, breaking through the traditional setting under the condition of fixed goals. Hanley et al. [7] proposed the combination of topic-independent and topic-aware embedding, emphasizing the important regulatory effect of target semantics on stance recognition. Liang et al. [8] further proved that background knowledge injection has obvious gain for stance modeling in social media scenarios.

In terms of model interpretability and complex scene adaptation, Saha et al. [9] incorporate explanation generation into the stance detection process, so that the model output no longer remains at the label level, but can reflect the judgment basis. Garg et al. [10] proposed a goal-oriented Transformer structure to strengthen the model's capture of the correspondence between target semantics and text content. Wang et al. [11] used consistency to enhance cue learning to improve the generalization ability of multi-domain stance detection. Niu et al. [12] constructed a challenging dataset and designed an effective model for stance recognition in the context of dialogue, showing that contextual round information cannot be ignored for stance judgment. Luusi et al. [13] applied stance detection to news media scenarios, reflecting the extended value of the task in public communication analysis. Farzam et al. [14] pointed out that multi-task learning has a stable gain for deep argument mining models. These studies provide direct inspiration for this paper, that stance discrimination should not be regarded as an isolated classification task, but should be modeled in conjunction with target, topic, context, and related subtasks.

In the research of topic modeling, Wu et al. [15] conducted a systematic review of neural topic models and pointed out that although traditional probabilistic topic models have strong interpretability, they are limited in deep semantic representation and context adaptability. Eklund et al. [16] used language model embedding clustering for topic modeling and showed that clustering strategy based on pre-trained representation can improve topic consistency. Wu et al. [17] further introduced the embedded clustering regularization term to improve the stability of the neural topic model. Chen et al. [18] started from the hierarchical topic structure to enhance the ability of multi-level semantic organization in complex corpora. As the large model method entered the field of topic analysis, Pham et al. [19] proposed the prompt-based TopicGPT framework, and Angelov et al. [20] showed that contextualized lexical embedding itself had strong topic modeling ability. Akiba et al. [21] research on the masking of explicit positive and negative expressions reminds us that surface attitude words do not always accurately represent the true stance. The topic modeling tool system developed by Wu et al. [22] provides convenience for model comparison and engineering implementation. Gericke et al. [23] and Boutaleb et al. [24] started from online topic modeling and trend detection, respectively, and promoted the development of dynamic topic identification.

As shown in Table 1, the existing research has achieved rich results in stance detection and topic modeling respectively, but the collaborative research for ideological and political education texts is still insufficient. On the one hand, many methods focus on single-task optimization, and topic features and stance signals are isolated from each other in the modeling process. On the other hand, ideological and political education texts have policy semantics, value orientation and educational context, and it is difficult to fully reveal the "issue-attitude" coupling relationship by simply relying on the general semantic representation. Based on this, this paper integrates topic identification and stance discrimination into the same multi-task learning framework on the basis of shared semantic coding, trying to improve the accuracy and interpretation ability of ideological and political education text analysis through feature sharing, task collaboration and semantic complementarity.

Table 1: Comparison of existing related studies

References	Research Focus	Methodological Features	Main Limitations or Differences from This Study
[1]–[2]	Stance detection reviews	Summarize task types, feature sources, and model evolution	Focus mainly on domain overviews and do not provide a coordinated topic–stance mechanism
[3]–[4]	Multi-task stance detection and knowledge enhancement	Introduce task interaction networks and external knowledge	Pay more attention to the stance task itself, with limited use of topic structure
[5]–[6]	Argument mining and open-domain target stance extraction	Expand datasets and task boundaries	Target complex corpora, but do not build a unified dual-task modeling framework
[7]–[8]	Topic-aware embeddings and background-knowledge-based stance modeling	Strengthen the connection between target semantics and context	Incorporate topic information, but still focus mainly on single-task stance optimization
[9]–[14]	Explainable stance detection, dialogue scenarios, and multi-task argument mining	Emphasize explanation generation, cross-domain generalization, and contextual turns	Pay insufficient attention to the value-semantic characteristics of ideological and political education texts
[15]–[18]	Neural topic modeling and embedding clustering for topic modeling	Improve topic coherence and hierarchical representation	Strong in topic discovery, but cannot directly perform stance classification
[19]–[24]	Prompt-based topic modeling, online topic analysis, and trend discovery	Emphasize dynamic topic recognition and engineering implementation	More suitable for issue discovery and trend tracking, but lack stance-oriented collaborative constraints

3 Methods and materials

3.1 The topic identification model of ideological and political Education texts based on the fusion of shared semantic Coding and topic features

Ideological and political education texts have strong issue-oriented and value-laden characteristics, and their topics are not always directly presented by high-frequency words or explicit labels. The same word may bear different theoretical meanings in different educational contexts, and different expressions may jointly point to the same ideological theme. If only the traditional bag of words model, shallow topic model or a single text classifier are used, only coarse topic boundaries can be obtained, and it is difficult to deal with the problems such as concept alias, implicit reference and context transfer. Based on this characteristic, this paper constructs a topic recognition model based on the fusion of shared semantic coding and topic feature under the overall framework of multi-task learning. The model integrates local context semantics, global topic distribution and domain concept prior into the same computing link to improve the stability and interpretability of topic identification.

Before running the model, the original corpus needs to be normalized. This paper selects discussion texts of ideological and political education courses, thematic learning experiences, online reviews and policy interpretation materials to form the basic corpus set. Considering that such texts often contain abbreviations, policy terms, colloquial transliterations and non-standard symbols, the preprocessing stage not only includes the general denoising, but also needs to add the domain semantic alignment operation. Specifically, the system completed the cleaning of special symbols, the folding of repeated characters, the filtering of stop words and the segmentation of sentences, and then combined with the custom ideological and political education term list, the high-frequency domain phrases such as "ideal and belief education", "value leading", "political identity" and "family and country feelings" were retained as a whole to avoid excessive semantic fragmentation after word segmentation. Then, the word form normalization and term mapping module are used to unify the synonymous or near-synonymous expressions into a relatively stable concept space, which lays the foundation for subsequent shared coding and topic aggregation. The preprocessing flow is shown in Figure 1.

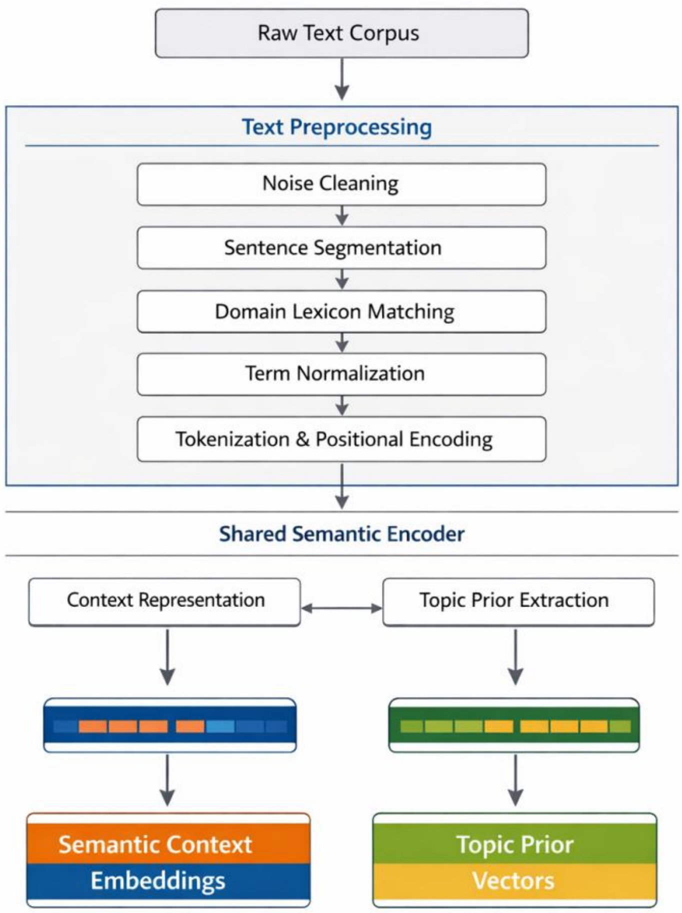


Figure 1: Flowchart of topic identification preprocessing and shared coding of ideological and political education texts

After the corpus preprocessing, this paper uses the shared semantic encoder to deep represent the text. Let the input text be:

$$X = \{x_1, x_2, \dots, x_n\} \tag{1}$$

where x_i denotes the ith lemma and n is the text length. The initial representation matrix is obtained after the text is passed through the embedding layer:

$$E = [e_1, e_2, \dots, e_n], \quad e_i \in \mathbb{R}^d \tag{2}$$

where, d represents the embedding dimension. Then, the word vector, position vector and domain type tag are fed into the shared semantic encoder to obtain the context representation matrix:

$$H = \mathrm{Encoder}(E) = \{h_1, h_2, \dots, h_n\} \tag{3}$$

Among them, h_i is no longer an isolated lemma representation, but a semantic state after integrating context dependencies. Different from ordinary text classification, this paper does not directly pool H to output categories, but further introduces a topic prior feature extraction mechanism, so that the model retains the perception of global topic structure in addition to deep semantic representation. Considering the strong conceptual aggregation property of ideological and political education texts, this paper introduces the topic prototype matrix:

$$T = \{t_1, t_2, \dots, t_K\}, \quad t_k \in \mathbb{R}^d \tag{4}$$

Here, K represents the preset number of topics and t_k represents the KTH topic prototype vector. The prototype matrix is not a manually specified static label, but is formed by the initial clustering results of the domain corpus and the parameter update during the training process. For any text representation H, the model first computes the matching strength between the lemma state and each topic prototype:

$$\alpha_{ik} = \frac{\exp(h_i^{\top} t_k)}{\sum_{j=1}^{K} \exp(h_i^{\top} t_j)} \tag{5}$$

where, α_{ik} represents the attribution weight of the ith lemma to the KTH topic. Based on this weight, an aggregated representation of the text along each topic dimension is obtained as follows.

$$z_k = \sum_{i=1}^{n} \alpha_{ik} h_i \tag{6}$$

Thus, the topic perception vector group $Z = \{z_1, z_2, \dots, z_K\}$. The core significance of this process is that it does not compress the whole text into a single vector, but lets the text generate responses in multiple candidate topic Spaces respectively, so as to better adapt to the actual characteristics of "one text with multiple issues and primary and secondary hierarchies" in ideological and political education texts.

However, topic aggregation alone is not enough to solve the problem of confusable category boundaries. The reason is that many ideological and political education texts share similar words on the surface, such as "youth responsibility", "social participation", "value choice" and other expressions will appear in multiple themes. If the classification is directly based on the topic response strength, the model may still be interfered by high-frequency common words.

To this end, we add a topic feature fusion module between shared encoding and topic prototyping to map the global semantic vector and topic aggregation representation into a unified feature space. Let the global representation of text be:

$$g = \frac{1}{n}\sum_{i=1}^{n} h_i \tag{7}$$

Then the fused topic discriminant vector can be expressed as:

$$u = \mathrm{ReLU}(W_g g + W_z \sum_{k=1}^{K} \beta_k z_k + b) \tag{8}$$

Here, W_g and W_z are the learnable parameter matrices, b is the bias term, and β_k is the topic importance coefficient. This structure enables the model to make use of global semantics and local topic responses at the same time, reducing the information loss caused by single path modeling. Furthermore, the model outputs the topic probability distribution through the Softmax layer:

$$\hat{y} = \mathrm{Softmax}(W_u u + b_u) \tag{9}$$

where $\hat{y}$ represents the predicted probability of the text on each topic category.

The overall structure of the topic identification model is shown in Figure 2. It does not simply connect the encoder output to the classification layer, but constructs a complete path of "shared semantic encoder-topic prototype matching-topic feature fusing - topic probability output". There are two practical considerations in this design. First, the topics in ideological and political education texts often have theoretical frameworks, and deep semantic representation can improve the recognition ability of the model for abstract concepts. Second, the topic recognition results will directly serve the subsequent stance discrimination tasks, so the model in this section needs to generate topic vectors with clear structure and transferability in the training phase, rather than only pursuing single-pass classification accuracy.

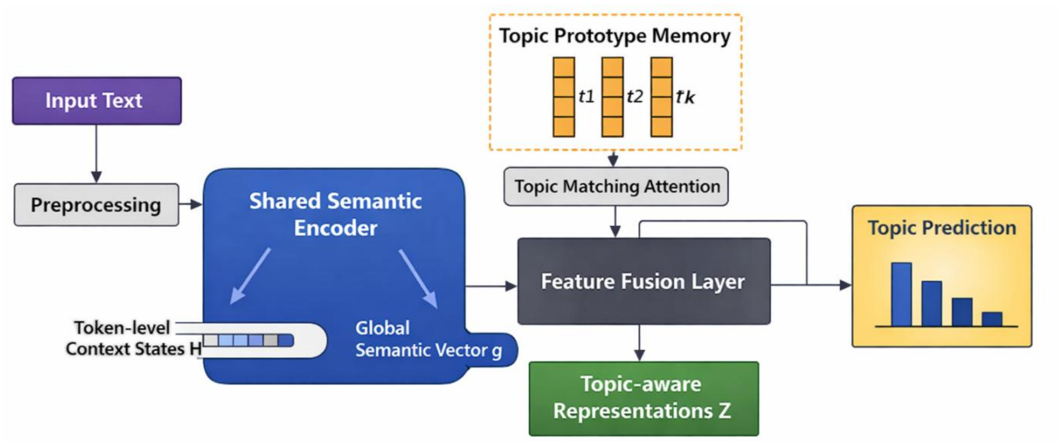


Figure 2: Structure diagram of the topic identification model based on the fusion of shared semantic encoding and topic features

In terms of the training objective, this paper uses cross-entropy loss as the main loss function for topic identification:

$$\mathcal{L}_{\text{topic}} = -\sum_{c=1}^{C} y_c \log \hat{y}_c \tag{10}$$

where C is the number of topic categories, y_c is the true label, and $\hat{y}_c$ is the predicted probability. Since this study belongs to the multi-task learning framework, the shared representation and topic fusion vector generated by the topic identification branch will also participate in the parameter update together with the stance discrimination branch. Therefore, the actual optimization in the training process is not a single topic loss, but the topic representation is constantly modified under the joint objective constraint. In this way, the topic model will not stay in the narrow task perspective of "only responsible for the topic label", but will perform a more refined representation of those key expressions that are both topic and attitude under the backpropagation effect of stance signals.

From the perspective of method logic, the improvement of the proposed model compared with the traditional LDA or simple BERT classifier is mainly reflected in two levels. First, the topic is no longer regarded as a static group of words out of context, but dynamically generated through prototype matching based on shared semantic coding, so it can better adapt to implicit topics and compound topics in ideological and political education texts. Second, the topic representation does not exist in isolation, but is integrated with the global semantics to form a stable feature structure that can be reused for subsequent tasks, which makes the topic identification link not only serve the task, but also provide a more binding semantic basis for stance discrimination. Through this modeling process, the model can accurately capture the distribution characteristics of texts in different theme directions such as theoretical identity, value shaping, policy understanding, social responsibility, and national consciousness, which lays a structured representation foundation for subsequent multi-task collaborative learning.

3.2 The Stance discrimination model of Ideological and political Education texts based on multi-task collaborative learning

Standpoint discrimination in ideological and political education texts is not a simple substitute for emotional polarity. Different from the general positive and negative sentiment classification, stance judgment emphasizes more on the value attitude, recognition degree and response direction of the text on a specific issue, and its discrimination results are often dependent on the topic context. In other words, judging "support" or "doubt" without the context of the issue is easy to confuse value evaluation, fact statement and strategy advice. Especially in the scene of ideological and political education, the text often contains policy explanation, theoretical paraphrasing, personal experience and realistic issue response at the same time. The surface words may not directly expose the position, and it is often the semantic coupling relationship between the text and the theme that really plays a decisive role. Therefore, this paper does not design stance discrimination as a parallel classifier independent of topic identification. Instead, under the framework of multi-task learning, a collaborative stance discrimination model for ideological and political education texts is constructed based on shared semantic coding and topic fusion representation, so that stance judgment is based on explicit issue constraints.

In the previous section, the model has been given a shared semantic representation $H = \{h_1, h_2, \ldots, h_n\}$ and the topic fusion vector u. H preserves the word-level context information, and u combines the global semantic and topic prototype response results, which can more stably represent the topic center pointed by the text. Taking this as input, the stance discrimination branch no longer reconstructs a set of semantic paths separated from the topic, but treats the topic vector as a conditional constraint in stance analysis. Let the topic guide vector be:

$$q = \tanh(W_q u + b_q) \tag{11}$$

where, W_q is the mapping matrix, b_q is the bias term, and q represents the stance query vector obtained from the projection of topic fusion results. This vector does not assume the final classification function, but serves as a "topic reference" in subsequent calculations, guiding the model to determine which contextual expressions are truly relevant to the current topic, and which are just general statements or background information.

Considering that there are many long-distance dependence phenomena in ideological and political education texts, such as the topic is proposed before the text, the evaluation is given after the text, or the attitude is expressed after the fact explanation is inserted in the middle, this paper adds a bidirectional gated recurrent unit layer after the shared coding output to further integrate the topic-related semantics. For the ith lemma, the original context state and the topic guide vector are concatenated to obtain:

$$m_i = [h_i; q] \tag{12}$$

It is then fed into a bidirectional gated recurrent unit network to form a new context state:

$$s_i = \mathrm{BiGRU}(m_i), \quad i = 1,2, \ldots, n \tag{13}$$

Here, s_i represents the updated lemma representation under the constraint of the topic condition. Compared with directly using the shared encoder output, the significance of this step is that BiGRU can retain intra-sentence long-distance dependency information in the round-return propagation of the previous and subsequent texts, and continuously receive topic signals during the state update process, so that "stance expression" is no longer triggered by local words, but is discriminated under the semantic framework of the topic.

However, sequence modeling alone is still not enough to deal with the implicit position problem commonly existing in ideological and political education texts. Many texts do not use explicit markers such as "for" and "against" directly, but indirectly express attitudes through evaluative phrases, adverbs of degree, transition structures, or quotes. In order to highlight such key positions, this paper designs a topically guided attention mechanism to select the hidden states output by BiGRU. The attention median at the ith position is denoted by:

$$e_i = v^{\mathsf{T}} \tanh(W_s s_i + W_t q + b_t) \tag{14}$$

Here, W_s and W_t are the weight matrices, v is the learnable parameter vector, and b_t is the bias term. After normalization, the attention weights of each position can be obtained as follows.

$$\alpha_i = \frac{\exp(e_i)}{\sum_{j=1}^{n} \exp(e_j)} \tag{15}$$

Thus, the stance semantic aggregation vector is obtained as follows.

$$c = \sum_{i=1}^{n} \alpha_i s_i \tag{16}$$

where, c represents the sentence-level stance representation after topic-guided and contextual screening. Different from ordinary self-attention, the weight assignment here is not only based on the relationship between words, but explicitly controlled by the topic vector q. Therefore, the model will give priority to focusing on those evaluation components that have strong coupling with the current topic, such as "should be strengthened", "exists deviation", "needs to be deepened", "has more realistic significance" and so on. This design can effectively weaken the interference of public cliche and background narrative on stance labels.

Figure 3 shows the main structure of the stance discrimination branch. The model takes values from the shared encoding output and the topic fusion vector at the same time, and generates a high-level representation for stance classification after topic mapping, sequence modeling and attention aggregation. Compared with the single-task stance classification model, the key of this structure is not to add an extra layer of network, but to maintain linkage between each step of the stance path and the topic representation, so that the model always focuses on the core problem of "which attitude to express on which issue".

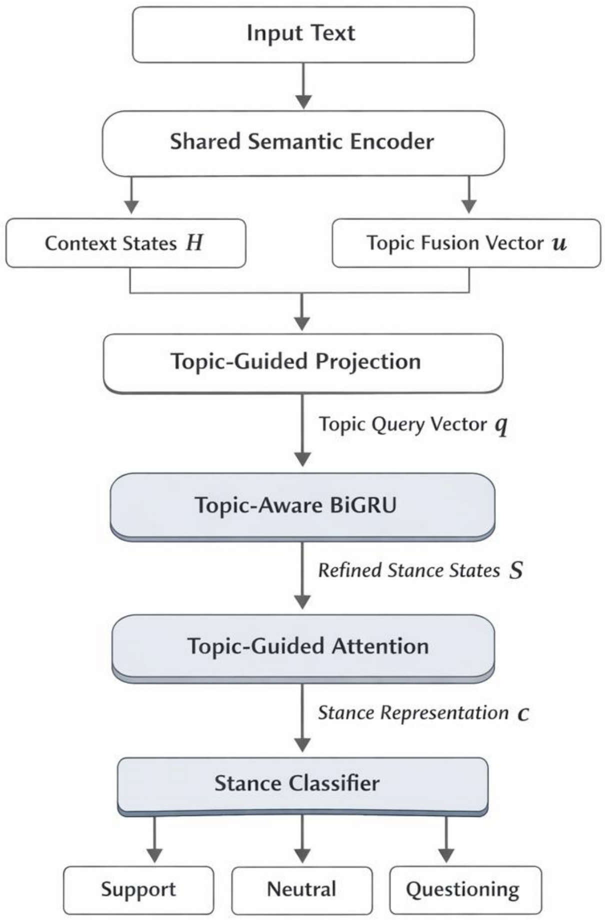


Figure 3: Structure diagram of stance discrimination model of ideological and political education text based on multi-task collaborative learning

In the classification output stage, the stance aggregation vector c and the topic fusion vector u are concatenated again to retain the direct constraint of the topic branch on the final discrimination and construct the joint discrimination representation:

$$r = \phi(W_r[c; u] + b_r) \tag{17}$$

where $\phi(\cdot)$ is a nonlinear activation function. Then, the stance probability distribution is output by the Softmax layer:

$$\hat{\mathrm{p}} = \mathrm{Softmax}(\mathrm{W_p r} + \mathrm{b_p}) \tag{18}$$

In this paper, the stance labels are divided into three categories: "agree", "neutral" and "question". The reason why this division is adopted, rather than simply copy the positive and negative polarity in general sentiment analysis, is that the evaluation in ideological and political education texts is often normative and discussionable. Many texts are not explicitly opposed to it, but express different opinions in the way of discussing, supplementing or limiting conditions. The three-category setting is more in line with the actual expression state of the corpus in this study.

In order to make topic identification and stance discrimination truly form a synergistic relationship, this paper introduces a joint loss function on the training objective. The supervised loss of the stance branch is defined as:

$$\mathcal{L}_{\mathrm{stance}} = -\sum_{\mathrm{k}=1}^{\mathrm{M}} \mathrm{y}_{\mathrm{k}}^{(\mathrm{s})} \log \hat{\mathrm{p}}_{\mathrm{k}} \tag{19}$$

where M is the number of stance categories and $\mathrm{y}_{\mathrm{k}}^{(\mathrm{s})}$ is the true stance label. Considering that there should be a certain consistency between the topic probability distribution output by the topic branch and the attention weight of the stance branch, we further introduce a cross-task consistency constraint. Let the topic importance vector normalized by the topic branch be γ, and the auxiliary distribution γ is obtained by mapping the stance branch on the topic dimension, then the consistency loss can be written as:

$$\mathcal{L}_{\mathrm{cons}} = \sum_{\mathrm{l}=1}^{\mathrm{K}} \gamma_{\mathrm{l}} \log \frac{\gamma_{\mathrm{l}}}{\tilde{\gamma}_{\mathrm{l}}} \tag{20}$$

This item is essentially a cross-task distribution constraint, which aims to make the stance branch focus on the core issue given by the topic branch when making judgments, and reduce the deviation caused by local high-weight words. Integrating the topic identification loss $\mathcal{L}_{\mathrm{topic}}$ in the previous section, the total model objective function is expressed as:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{\mathrm{topic}} + \lambda_2 \mathcal{L}_{\mathrm{stance}} + \lambda_3 \mathcal{L}_{\mathrm{cons}} \tag{21}$$

Here,$\lambda_1, \lambda_2, \lambda_3$ are the loss weight coefficients. Through joint optimization, the topic branch will provide more boundary sense of topic representation to the stance branch, and the stance branch will inversely push the shared encoder to capture the attitude expression more sensitively, thus forming a bidirectional gain.

From the perspective of the algorithm process, stance discrimination does not mechanically receive a topic label after the topic identification, but continuously uses the topic vector to participate in reasoning and screening. The specific process is shown in Figure 4: The model first reads the preprocessed ideological and political education text, and calls the shared semantic encoder to generate the word-level context state. Then, the topic fusion vector is extracted from the topic identification branch and mapped to the query vector. Then, the topic condition is injected into BiGRU to reconstruct the stance oriented context of the text sequence. Then, the topic is used to guide attention to select the lexical states closely related to stance expression. Finally, the stance category is output by the joint discrimination layer, and the error is fed back to the shared layer and the topic layer at the same time. This process avoids the separation problem of "no update after the completion of topic identification" in the traditional serial framework, so that the two tasks are always in a linkage state in the training phase.

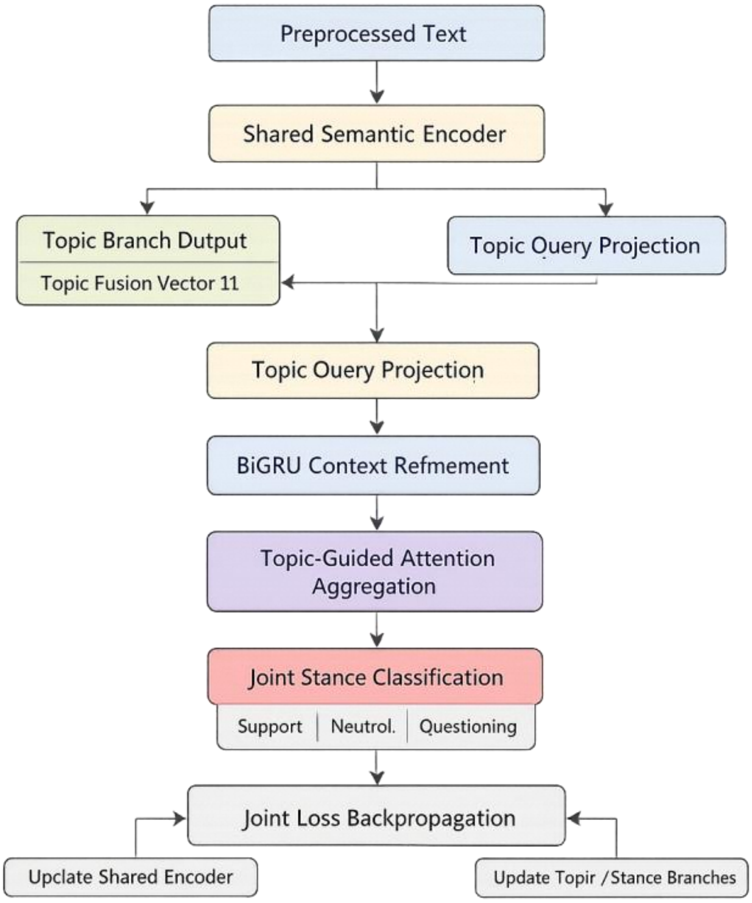


Figure 4: Flowchart of collaborative computation for stance discrimination under subject constraints

Based on the above structure, the model constructed in this paper can well deal with the three common difficulties in ideological and political education texts. First, for texts with a clear topic but scattered attitude expression, the topic vector can provide stable anchors for stance attention and reduce the interference of irrelevant sentence segments. Second, for texts with implicit positions but continuous discussion chains, BiGRU's bidirectional state propagation helps to restore the forward and backward dependence in the process of attitude formation. Thirdly, for the complex situation of explanation, citation and evaluation in the same text, the multi-task consistency constraint can inhibit the model from overfocusing on local salient words, thereby improving the robustness of stance judgment. Therefore, the stance discrimination branch not only completes the automatic identification of the identification tendency, but also provides a stronger semantic closed-loop ability for the whole collaborative modeling framework.

4 Experimental Results

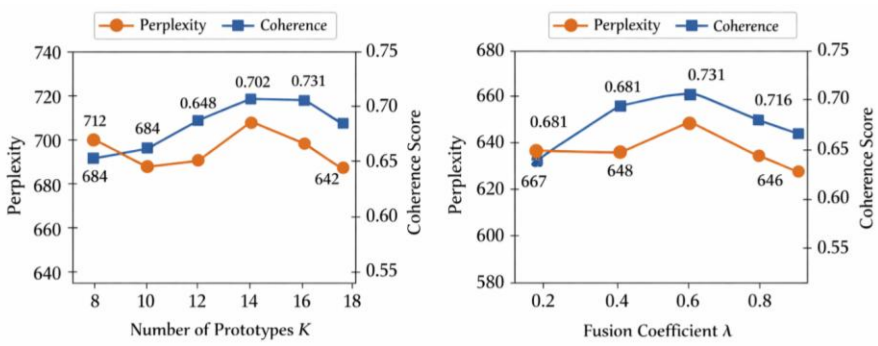
4.1 Experimental results of topic identification model

In order to verify the effectiveness of the topic identification branch in the ideological and political education text scene, a mixed corpus including course discussion texts, thematic learning experiences, ideological and political case analysis short articles, policy interpretation texts and campus network comments is constructed, with a total of 58,000 samples. Among them, 42,000 texts have human topic labels for supervised training and testing. Another 16,000 unlabeled texts are cleaned and used for topic prior learning and prototype initialization. The whole corpus is divided into training set and test set according to 8:2, and 10% of the training set is further extracted as the validation set. AdamW optimizer was used for the model with initial learning rate set to 2×10−5, batch size set to 32, maximum training rounds set to 40, and early stopping rounds set to 5. The shared encoding dimension is set to 768 and the topic prototype dimension to 256. To avoid accidental errors, all comparative experiments were repeated five times independently under the same conditions, and the results were averaged in the paper. See Table II for the configuration of the experimental platform.

Table 2: Configuration of the experimental platform for topic identification

Category	Item	Configuration
Hardware Environment	CPU	Intel Xeon Gold 6330
Hardware Environment	GPU	NVIDIA A100 40GB
Hardware Environment	Memory	256 GB
Hardware Environment	Storage	2 TB NVMe SSD
Software Environment	Operating System	Ubuntu 22.04
Software Environment	Deep Learning Framework	PyTorch 2.1
Software Environment	Data Processing	Pandas 2.1.4
Software Environment	Chinese Word Segmentation and Preprocessing	jieba 0.42.1, pkuseg 0.0.25
Software Environment	Visualization and Dimensionality Reduction	matplotlib 3.8, umap-learn 0.5.5

Perplexity, Topic Coherence and Silhouette Score were used to measure the effectiveness of topic recognition. The lower the perplexity, the more stable the model fits the text generation distribution. The higher the consistency, the more concentrated the internal semantics of the topic. The higher the silhouette coefficient, the clearer the boundary between the topic clusters. Considering that the proposed model introduces two key parameters, the number of topic prototypes K and the fusion coefficient λ, the sensitivity test of these two parameters is performed in the experiment, and the results are shown in Figure 5.



(a) Effect of Prototype Number K (b) Effect of Fusion Coefficient λ

Figure 5: Influence of number of topic prototypes and fusion coefficient on model performance

As can be seen from Figure 5, when the number of topic prototypes increases from 8 to 14, model perplexity decreases from 712 to 638, and topic consistency increases from 0.601 to 0.731, indicating that moderately increasing topic prototypes helps to enhance the discrimination of topic boundaries in ideological and political education texts. When K continues to increase to 16 and 18, the perplexity rises slightly, and the consistency also decreases, indicating that too many topic prototypes will make some similar topics oversegmented, which will weaken the degree of semantic concentration. The fusion coefficient experiment shows a similar trend. When λ=0.6, the model achieves the lowest perplexity 638 and the highest consistency 0.731. If the proportion of local semantic representation in the fusion layer continues to increase, the model can retain rich context information, but the structural constraints of topic distribution are partially diluted, resulting in a drop in the index. It can be seen that the proposed model can better balance the relationship between shared semantic encoding and topic prototype constraints under the condition of K=14 and λ=0.6.

After determining the optimal parameters, this paper compares the proposed model with four common methods: LDA, NMF, BERTopic and BERT-Cluster. LDA and NMF represent traditional statistical topic modeling paths, while BERTopic and BERT-Cluster reflect the idea of topic discovery based on pre-trained embedding. The comparison results are shown in Figure 6.

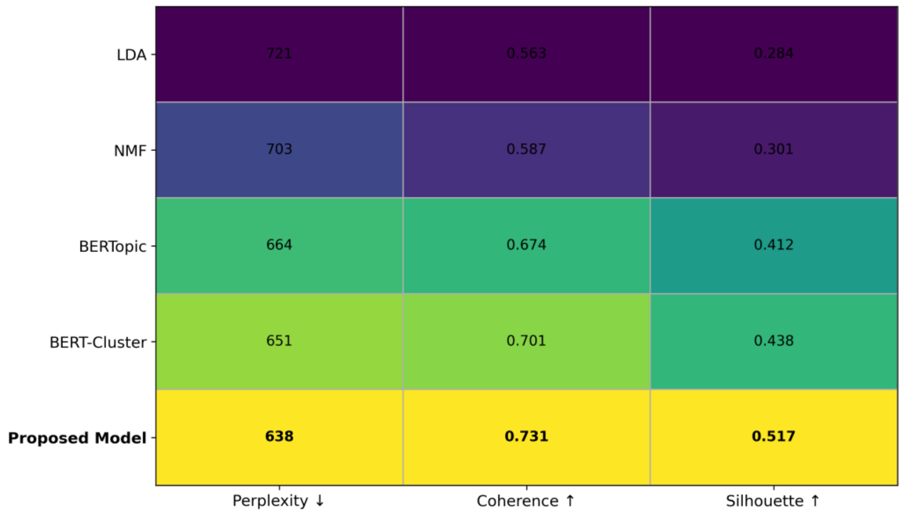


Figure 6: Performance of different models in the task of topic identification of ideological and political education texts

From Figure 6, we can see that the traditional LDA model has the highest topic perplexity in this task, which is 721, and the topic consistency is only 0.563, indicating its limited ability to adapt to implicit concepts and context transfer in ideological and political education texts. NMF has a slight improvement over LDA, but the overall improvement is not large. BERTopic relies on contextual embedding, and the perplexity decreases to 664 and the consistency increases to 0.674, indicating that the pre-trained semantic representation indeed helps to improve the topic organization effect. BERT-Cluster further uses text embedding clustering for topic identification, which is better than BERTopic in consistency and silhouette coefficient, but it is still not stable enough in boundary control between topics. The model in this paper achieves the lowest perplexity 638, the highest consistency 0.731 and the highest contour coefficient 0.517. Compared with BERT-Cluster, the consistency is increased by 0.030 and the contour coefficient is increased by 0.079. Compared with LDA, the perplexity is decreased by

83, indicating that the shared semantic coding and topic feature fusion strategy can more accurately depict the issue distribution of ideological and political education texts.

By further observing the typical topic clusters in the test set, it can be seen that the traditional model is easy to mix the topics with similar semantics but different boundaries such as "ideal belief education", "youth responsibility awareness" and "social participation practice" into the same cluster, while the proposed model can stably distinguish the core topics such as "political identity", "value shaping", "national consciousness" and "public responsibility". Taking the high-frequency expression of "youth mission", "responsibility of The Times" and "spirit of struggle" as an example, LDA and NMF often combine them with general inspirational texts. The model in this paper can more accurately classify such texts into the theme cluster of "responsibility ethics and family and country responsibility" with the help of theme prototype memory and fusion layer constraints. The standard deviation of the five repeated experiments is controlled within 0.012, which indicates that the model results have good stability. On the whole, the topic identification branch not only outperforms all kinds of baseline models in terms of numerical indicators, but also shows clearer semantic boundaries in the actual issue division of ideological and political education texts, which provides a more reliable basis for the topic constraint of subsequent stance discrimination branches.

4.2 Experimental results of stance discrimination model

In order to test the recognition effect of the stance discrimination branch constructed in this paper in the ideological and political education text scene, on the basis of the above 58,000 corpus, this paper further screens and manually labels 26,400 text samples with clear topic direction and discriminative stance expression, and constructs a three-class stance dataset. Among them, 10,820 were classified as "agree", 8,960 were classified as "neutral", and 6,620 were classified as "doubt". The data is divided into training set, validation set and test set according to 8:1:1, and the test set has 2,640 entries. The training phase keeps the shared encoder parameters consistent with the topic identification branch, and the optimizer uses AdamW with an initial learning rate of 2×10^{-5}, a batch size of 32, and a maximum number of training rounds of 35. Considering that stance discrimination is more dependent on the stability of topic constraints, this paper sets the topic consistency loss weight λ_3 as an adjustable parameter, and compares the performance of the model under different training scales. The evaluation metrics Accuracy, Precision, Recall and Macro-F1 were selected, and all experiments were repeated five times independently, and the results were averaged.

From the classification details, the performance of the model in this paper is not completely consistent on the three types of positions, which is related to the expression characteristics of the ideological and political education texts themselves. The "identification" text often has more explicit value positive words and policy response semantics, so it is easier to be stably recognized. The "neutral" category contains more explanatory, paraphrasing and paving expressions, and its boundaries are most likely to cross with the other two categories. Although turning points, restrictions and conditional expressions often appear in the "questioning" texts, many of the samples do not directly use negative words, so they are more dependent on the context. The classification results of the proposed model on the test set are shown in Table 3.

Table 3: Classification results of the proposed model on each stance category

Stance Category	Number of Samples	Precision / %	Recall / %	F1 / %
Support	1082	87.6	88.9	88.2
Neutral	896	83.4	81.7	82.5
Questioning	662	86.8	83.8	85.3

Macro Average	2640	85.9	84.8	85.3

Table 3 shows that the F1 value of the model on the "identification" class reaches 88.2%, which is the most stable performance. The F1 value of the "neutral" category was 82.5%, which was lower than the other two categories, indicating that in the context of ideological and political education, the neutral expression was often not a real emotional blank, but a mixture of interpretation, paraphras and prudent evaluation, and the model still needed to continue to optimize on the fine-grained boundary. The Precision and Recall of the "questioning" class are 86.8% and 83.8%, respectively, indicating that the proposed model has a good ability to capture the implicit stance cues, but some texts with mild suggestions are still misjudged as neutral. Overall, Macro-F1 reaches 85.3%, indicating that the multi-task collaboration mechanism can maintain a more balanced recognition performance between different categories.

To further investigate the influence of topic constraint strength and training sample size on stance discrimination effect, this paper adjusted the consistency loss weight λ3 and the number of training samples respectively, and the results are shown in Figure 7.

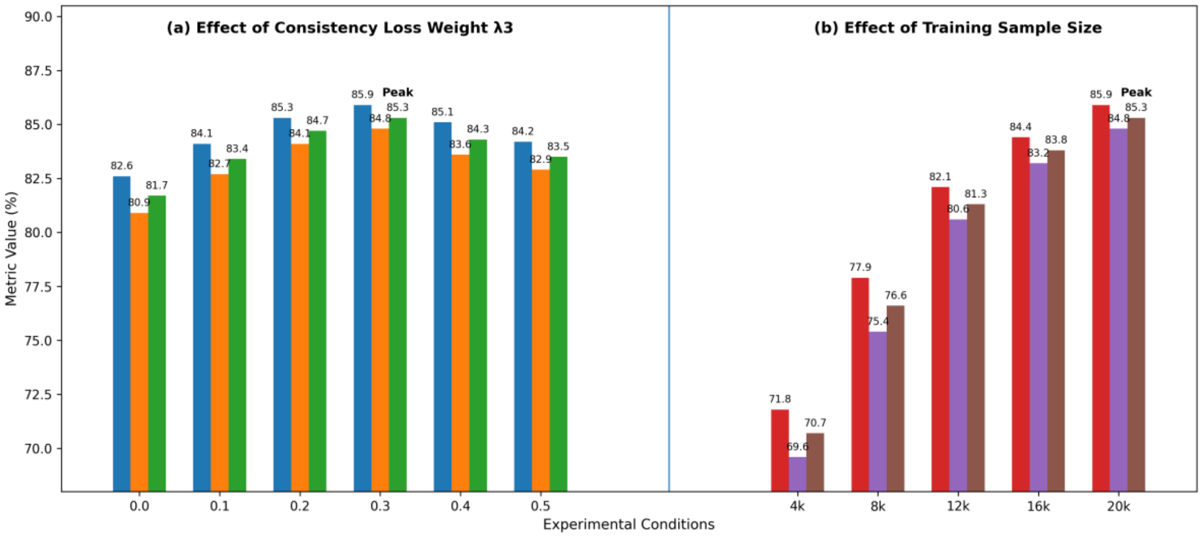


Figure 7: Changes of stance discrimination model indicators under different conditions

Figure 7(a) shows that as λ3 increases from 0 to 0.3, Precision, Recall and F1 continue to rise, indicating that the consistency constraint between the topic branch output and the stance branch attention distribution can effectively suppress the phenomenon that the model is pulled by local high-frequency words. When λ3=0.3, Macro-F1 reaches the highest value of 85.3%. The reason is that the strong consistency constraint will compress the learning space of stance branch for fine-grained context differences, making the model overly dependent on the topic principal vector, and reducing the discrimination of edge samples. The difference between adjacent parameter points reached the significant level at λ3=0.3 after two-sided t-test ($p<0.05$). Figure 7(b) shows that when the training samples are extended from 4,000 to 20,000, the three indicators show a continuous upward trend, but the increase is faster in the range of 4,000-12,000, and the increase slows down after 12,000. This shows that the proposed model has a good adaptability to the sample size, and also reflects that the multi-task collaborative structure can form a relatively stable discrimination ability under the condition of medium scale data.

In the baseline comparison experiments, this paper selects TextCNN, BiGRU, BERT-base, Topic-BERT, StanceFormer[10] and "multi-task model with removing Consistency constraints" as references. The comparison results are shown in Table 4.

Table 4: Performance comparison of different models on stance discrimination task

Model	Accuracy / %	Precision / %	Recall / %	Macro-F1 / %
TextCNN	73.8	72.4	71.1	71.7
BiGRU	77.6	76.9	75.5	76.2
BERT-base	81.4	80.8	79.7	80.2
Topic-BERT	82.9	82.1	81.3	81.7
StanceFormer	84.1	83.7	82.4	83.0
Multi-task Model (without Consistency Constraint)	85.2	84.8	83.1	83.9
Proposed Model	86.7	85.9	84.8	85.3

Table 4 shows that the performance of traditional TextCNN and BiGRU in this task is relatively limited, indicating that it is difficult to fully depict the "topic-stance" coupling relationship in ideological and political education texts by only relying on local convolutional features or general sequence modeling. BERT-base significantly improves the performance through contextual semantic representation, but its Macro-F1 is still stuck at 80.2%. Topic-bert with Topic information is further improved to 81.7%, indicating that topic priors do help mitigate stance drift. StanceFormer has an advantage in goal-aware modeling with Macro-F1 reaching 83.0%, but it is still lower than the collaborative model proposed in this paper. After removing the consistency constraint, the Macro-F1 of the multi-task model is 83.9%, which is 1.4 percentage points lower than that of the model in this paper. This gap indicates that sharing semantic encoding itself is not enough to ensure the true collaboration of cross-task information, and an explicit consistency constraint mechanism between topic branch and stance branch is still needed to stabilize information transmission. The model in this paper achieves the highest values in Accuracy, Precision, Recall and Macro-F1, and the difference is significant compared with the strongest baseline ($p<0.01$).

In general, the field discrimination model established in this paper shows good accuracy and stability in the three-classification task of ideological and political education texts. Our performance gains do not come from the stacking of a single deep network, but from the synergy of shared encodings, topic guidance, sequence modeling, and cross-task consistency constraints. The topic branch provides a clear topic boundary for stance identification, and the stance branch promotes the sharing layer to capture the value judgment expression more sensitively through backpropagation. This bidirectional coupling relationship is the key reason why the proposed model outperforms the single-task method.

5 Discussion

The multi-task collaborative model constructed in this paper shows clear structural advantages in the analysis of ideological and political education texts. This advantage is not simply reflected in the rise of the index value, but also in the effective connection between the topic semantics and the stance semantics. Experimental results show that the topic recognition branch achieves the best performance when $K=14$ and $\lambda=0.6$, the topic perplexity is reduced to 638, the topic consistency is 0.731, and the silhouette coefficient is 0.517, which is better than the control models such as LDA, NMF, BERTopic and BERT-Cluster. This result shows that shared semantic encoding provides stable context representation, while topic prototype matching further strengthens topic boundary constraints, and the combination of the two can reduce topic drift caused by concept intersection. The improvement of stance discrimination part is also targeted in method. The Accuracy, Precision, Recall and Macro-F1 of the proposed model on the test set reach 86.7%, 85.9%, 84.8% and 85.3%, respectively. It is 2.6, 2.2, 2.4 and 2.3 percentage points higher than StanceFormer, respectively, which is also significantly better than the multi-task model after removing the consistency constraint. This suggests that topic information is not an ancillary feature, but an indispensable condition variable in stance discrimination. Especially in ideological and political education texts, many attitude expressions are not directly presented through strong emotional words, but hidden in the combination of explanation, qualification, response and evaluation. Through topic-guided attention and cross-task consistency constraints, we enable the model to filter key semantic positions around specific issues, thereby improving the ability to identify implicit positions. The results shown in Figure 7 also show that the model performance reaches a peak when $\lambda 3=0.3$, indicating that a moderate consistency constraint can enhance cross-task information transfer, but if the constraint is too strong, it will also compress the learning space of the stance branch for local details. There are still some limitations in this paper. On the one hand, the current corpus mainly comes from course discussions, learning experiences and online reviews. Although it covers a variety of expression scenarios, the cross-platform transfer ability still needs to be further tested. On the other hand, the model still has some room for misjudgment of metaphorical expression, ironic expression and attitude change in long compound discourse. Subsequent research can be carried out along the directions of lightweight deployment, cross-domain transfer learning, external knowledge injection, and multimodal expansion, so as to improve the applicability and depth of interpretation of the model in real education governance scenarios.

6 Conclusions

Aiming at the problems of the separation of topic identification and stance discrimination and the lack of semantic correlation in ideological and political education texts, this paper constructs a collaborative modeling framework based on multi-task learning. The model takes the shared semantic coding as the core, integrates the topic prototype matching and feature fusion mechanism into the topic identification branch, and introduces the topic-guided attention and cross-task consistency constraint into the stance discrimination branch, so that the topic information and attitude information of the text can be jointly optimized in a unified representation space. Experimental results show that the proposed model achieves 638 perplexity, 0.731 topic consistency and 0.517 silhouette coefficient in topic recognition task, and achieves 86.7% accuracy and 85.3% Macro-F1 in stance discrimination task, which are better than the comparison models. This shows that the collaborative modeling method can better improve the ability of topic boundary identification and implicit position capture of ideological and political education texts. The method in this paper has certain application value in course feedback analysis, network public opinion research and intelligent education governance scenarios. However, there is still room for improvement in the adaptability of existing models to cross-domain corpus, complex rhetoric and weak explicit attitude expression. Subsequent research can be further carried out around knowledge enhancement, cross-domain transfer and multi-modal fusion, so as to improve the generalization ability and interpretation ability of the model in complex educational scenarios.

Funding

This work was supported by the Philosophy and Social Science Research Project of the

Education System of Fujian Province (Grant No. JAS23256), the Scientific Research Project of Putian University (Grant No. 2025052), and the Key Faculty Development Fund of Xiamen Ocean Vocational College.

References

[1] Hardalov M, Arora A, Nakov P, et al. A survey on stance detection for mis-and disinformation identification[C]//Findings of the Association for Computational Linguistics: NAACL 2022. 2022: 1259-1277.

[2] Alturayeif N, Luqman H, Ahmed M. A systematic review of machine learning techniques for stance detection and its applications[J]. Neural Computing and Applications, 2023, 35(7): 5113-5144.

[3] Chai H, Tang S, Cui J, et al. Improving multi-task stance detection with multi-task interaction network[C]//Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. 2022: 2990-3000.

[4] He Z, Mokhberian N, Lerman K. Infusing knowledge from Wikipedia to enhance stance detection[C]//Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis. 2022: 71-77.

[5] Cheng L, Bing L, He R, et al. IAM: A comprehensive and large-scale dataset for integrated argument mining tasks[C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2022: 2277-2287.

[6] Li Y, Garg K, Caragea C. A new direction in stance detection: Target-stance extraction in the wild[C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2023: 10071-10085.

[7] Hanley H, Durumeric Z. Tata: Stance detection via topic-agnostic and topic-aware embeddings[C]//Proceedings of the 2023 conference on empirical methods in natural language processing. 2023: 11280-11294.

[8] Li A, Liang B, Zhao J, et al. Stance detection on social media with background knowledge[C]//Proceedings of the 2023 conference on empirical methods in natural language processing. 2023: 15703-15717.

[9] Saha R R, Lakshmanan L V S, Ng R T. Stance detection with explanations[J]. Computational Linguistics, 2024, 50(1): 193-235.

[10] Garg K, Caragea C. Stanceformer: Target-aware transformer for stance detection[C]//Findings of the Association for Computational Linguistics: EMNLP 2024. 2024: 4969-4984.

[11] Wang S, Pan L. Target-adaptive consistency enhanced prompt-tuning for multi-domain stance detection[C]//Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024). 2024: 15585-15594.

[12] Niu F, Yang M, Li A, et al. A challenge dataset and effective models for conversational stance detection[C]//Proceedings of the 2024 joint international conference on computational linguistics, language resources and evaluation (LREC-COLING 2024). 2024: 122-132.

[13] Lüüsi L, Kangur U, Chakraborty R, et al. Political stance detection in Estonian news media[C]//Proceedings of the 9th international workshop on computational linguistics for Uralic languages. 2024: 12-28.

[14] Farzam A, Shekhar S, Mehlhaff I, et al. Multi-task learning improves performance in deep argument mining models[C]//Proceedings of the 11th Workshop on Argument Mining (ArgMining 2024). 2024: 46-58.

[15] Wu X, Nguyen T, Luu A T. A survey on neural topic models: methods, applications, and challenges[J]. Artificial Intelligence Review, 2024, 57(2): 18.

[16] Eklund A, Forsman M. Topic modeling by clustering language model embeddings: Human validation on an industry dataset[C]//Proceedings of the 2022 conference on empirical methods in natural language processing: industry track. 2022: 635-643.

[17] Wu X, Dong X, Nguyen T T, et al. Effective neural topic modeling with embedding clustering regularization[C]//International Conference on Machine Learning. PMLR, 2023: 37335-37357.

[18] Chen H G, Mao P, Lu Y, et al. Nonlinear structural equation model guided gaussian mixture hierarchical topic modeling[C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2023: 10377-10390.

[19] Pham C M, Hoyle A, Sun S, et al. TopicGPT: A prompt-based topic modeling framework[C]//Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers). 2024: 2956-2984.

[20] Angelov D, Inkpen D. Topic modeling: Contextual token embeddings are all you need[C]//Findings of the Association for Computational Linguistics: EMNLP 2024. 2024: 13528-13539.

[21] Akiba T, Gato Y, Kimura Y, et al. Masking Explicit Pro-Con Expressions for Development of a Stance Classification Dataset on Assembly Minutes[C]//Proceedings of the Second Workshop on Natural Language Processing for Political Sciences@ LREC-COLING 2024. 2024: 33-38.

[22] Wu X, Pan F, Tuan L A. Towards the TopMost: A topic modeling system toolkit[C]//Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations). 2024: 31-41.

[23] Gericke E, Jamet N, Guo T, et al. Online Bert-based Topic Modelling[C]//Proceedings of the 9th edition of the Swiss Text Analytics Conference. 2024: 181-181.

[24] Boutaleb A, Picault J, Grosjean G. BERTrend: neural topic modeling for emerging trends detection[C]//Proceedings of the Workshop on the Future of Event Detection (FuturED).

2024: 1-17.