Journal of Engineering, Project, and Production Management
2026, 16(6), 2026-0055

Intelligent Classification of Chinese Books in University Libraries Based on BERT and PLM-LCN

Bo Hong
Librarian, Library, Qufu Normal University, Qufu, 273100, China, E-mail: BoHongbh@outlook.com

Project Management
Received May 11, 2026; revised June 25, 2026; accepted August 3, 2026
Available online August 15, 2026

Abstract: With the expansion of the collection scale of university libraries, the intelligent classification of Chinese books has become the key to improving efficiency. However, traditional classification methods have significant shortcomings in handling the complexity of Chinese semantics and interdisciplinary issues. Therefore, this study proposes an intelligent classification model based on the bidirectional encoder representations from pre-trained models and the fusion of neural network features. Firstly, this study utilizes a bidirectional encoder to represent the text's global semantic features in the pre-trained model. Secondly, a text classification model based on recurrent convolutional neural networks is utilized to capture local contextual information. Next, a multi-head attention mechanism is introduced to enhance and optimize text features, and finally, an efficient classification is achieved through a Softmax classifier. The experiment showed that the model achieved classification accuracy of 96.8% and 93.3% on the Chinese Library Classification Dataset and the National Science and Technology Library Literature Center Chinese Book Dataset, respectively. In the interdisciplinary book classification test, the model's AUC reached 83.7%, significantly improving retrieval accuracy and user satisfaction in real-world library systems. Research has shown that the proposed model can effectively address semantic understanding and feature fusion in Chinese book classification, providing an efficient and accurate solution for intelligent library management.

Keywords: Bidirectional encoder representation transformers (BERT); pre-trained language model and long short-term memory-convolutional neural network (PLM-LCN); multi-head attention (MHA); Chinese book classification; interdisciplinary semantic understanding.

DOI 10.32738/JEPPM-2026-0055

1. Introduction

The library is the core carrier of knowledge storage and dissemination. As the knowledge hub of the higher education system, university libraries play a role in storing academic resources and promoting academic innovation. With the continuous expansion of library collections, efficiently and accurately classifying and managing large volumes of books has become a core challenge in the intelligent transformation of university libraries. The traditional Chinese Book Classification (CBC) method faces problems such as low classification efficiency and insufficient interdisciplinary adaptability (Harisanty et al., 2025; Huang, 2024). Interdisciplinary issues are driving the transformation of libraries from traditional "knowledge repositories" to "intelligent knowledge service centers", with core roles reflected in breaking down disciplinary barriers, reconstructing service models, and activating resource value. Firstly, the interdisciplinary perspective has reshaped the logic of library resource construction. Traditional libraries organize resources by subject classification, but interdisciplinary research requires integrating resources from multiple fields. Secondly, interdisciplinary issues drive the service model from "passive response" to "active embedding" to achieve automation of demand forecasting, resource recommendation, and service optimization. Finally, interdisciplinary thinking forces libraries to break through organizational boundaries and become the "connectors" of academic ecology.

In recent years, the use of deep learning algorithms to optimize library book classification systems has gradually become a research hotspot in this field. Cui and Gao (2024) proposed a multi-feature collaborative filtering model that incorporates grey relational analysis and Bayesian inference to address cold-start and data-coefficient issues encountered in traditional collaborative filtering algorithms. This model could improve the classification accuracy of library books, enhance recommendation diversity, and improve library management efficiency, among other issues. Umer et al. (2023) addressed the insufficient word-vector representation in existing automatic text classification techniques and introduced the word-embedding generation model FastText to train the classification model. The use of FastText word embeddings has improved classification accuracy. Bahrani et al. (2024) designed a hybrid recommendation system to address the limitations of existing

retrieval-based recommendation systems in balancing scalability and accuracy. The system used both content-based filtering and collaborative filtering and utilized clustering techniques to improve its functionality. The proposed method had better performance and scalability. Ten Lia (2023) proposed a new method for the automatic classification of news texts, built on topic modeling and combined with the automatic assignment of topic tags. It also adopted a method of assigning topic labels using the ChatGPT language model. Experimental results have shown that the algorithm could serve as an effective tool in automatic text classification tasks.

Text classification is a fundamental problem in natural language. With the deepening of research and the continuous increase in training data, researchers have begun to introduce pre-trained models to address this issue (Pal et al., 2023). Ding et al. (2023) developed a method that uses triangle tuning to improve the pre-trained language model by addressing the high cost of fine-tuning and storing all parameters caused by continuous model expansion. Triangular tuning was practical for adapting to pre-trained language models. Lee et al. (2023) proposed a prompt-based generative pre-trained converter to generate personality test scales, exploring automatic item generation in natural language processing for personality. The personality test scale generated by training the model using a pre-trained converter had good measurability. Zhang et al. (2023) concluded that a transformer-based pre-trained language model for generating controllable text was among the more advanced text generation techniques and could address specific limitations in practical applications. Shen et al. (2023) proposed a Bidirectional Encoder Representation from Transformers (BERT)- based model to address the gap in pre-trained language models for the social sciences. This model could perform disciplinary classification, abstract structure function recognition, and named entity recognition tasks on social science citation indexes. Liu et al. (2023) designed a text classification method that combines Transformer and graph convolutional networks and conducted experiments on five representative datasets. The results indicate that this method effectively improves text classification accuracy and outperforms the comparative method. Min et al. (2025) proposed a universal bilingual text detection model that uses the multilingual pre-trained language model XLM RoBERTa to extract CLS vectors as overall semantic features. The experimental results showed significant progress on the baseline, especially on the English and Chinese HC3 sentence-level datasets, where F1 scores increased by 10% and 5%, respectively. Shi et al. (2023) proposed a new POS perception and layer ensemble transformer neural network (PL Transformer) and conducted extensive experiments on four datasets, demonstrating that the proposed model improves text classification performance.

In summary, existing research has made certain progress in intelligent book classification. However, these methods have limited performance in handling global semantic understanding and the collaborative optimization of local features for Chinese books. Based on this, this study innovatively proposes a hybrid Pre-trained Language Model with LSTM and CNN (PLM-LCN) that integrates BERT and neural network feature fusion ideas. Unlike existing studies that use only BERT as a word embedding tool or TextRCNN alone for text classification. The core contribution of the proposed model lies in clarifying the functional division and collaborative logic between each module: BERT is responsible for establishing a global text semantic representation, the loop and convolution structures in TextRCNN are responsible for mining key patterns in local context based on this representation, and MHA serves as the top-level mechanism for feature optimization, attention weighting the composite features output by the first two, suppressing redundant information and amplifying discriminative signals. This design enables the model to maintain a comprehensive understanding of the overall thematic direction of the literature when dealing with Chinese book classification, while also accurately identifying potential correlations between interdisciplinary terms, thereby demonstrating greater robustness in distinguishing between single-discipline and interdisciplinary books. Compared to existing classification architectures based on single pre-trained models or traditional neural networks, this method shows clear improvements in feature completeness and semantic depth. The novelty of the research lies in the organic integration of global semantic modeling, local feature extraction, and discriminative information focusing, forming a progressive processing logic suited to the characteristics of Chinese book classification tasks. Starting from the heterogeneity of semantic distribution in Chinese books, it has redefined the functional boundaries of each module and clarified the relationships of information transmission and enhancement among them.

2. Methods and Materials

2.1. Construction of a Classification Model Integrating TextRCNN and BERT Models

Due to the complexity of the semantics of Chinese books, how to use deep learning models to associate contextual information is the primary problem to be solved for intelligent CBC. The Text Recurrent Convolutional Neural Network (TextRCNN) is a hybrid model that combines Long Short-Term Memory (LSTM) from a Recurrent Neural Network (RNN) with a Convolutional Neural Network (CNN) for text classification (Yixuan et al., 2023). This model can assist in more accurately representing the meaning of text. Its structure is shown in Fig. 1.

In Fig. 1, the TextRCNN model combines the strengths of RNNs and CNNs to form a hierarchical feature-extraction system. The model uses an RNN to capture left and right contextual information for an input word, then applies local features via convolutional layers, followed by a pooling layer to extract the most significant features from the convolutional output. Finally, these features are fed into the output layer for classification prediction. The formula for combining a certain input word with the left and right context is given by Eq. (1) (Dong and Zhang, 2023).

$$\begin{cases} L(W_i) = f(H^{(L)}L(W_{i-1}) + H^{(SL)}e(W_{i-1})) \\ R(W_i) = f(H^{(R)}R(W_{i+1}) + H^{(SR)}e(W_{i+1})) \end{cases} \tag{1}$$

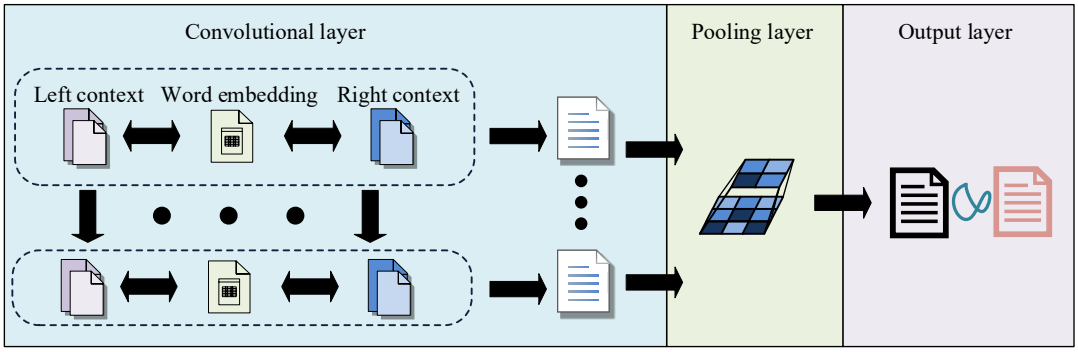


Fig. 1. TextRCNN structure (Source: self-drawn)

In Eq. (1), W_i is the input word, and $L(W_i)$ and $R(W_i)$ are the left and right contexts of word W_i. f is a nonlinear activation function. H is the transformation matrix. $H^{(SL)}$ and $H^{(SR)}$ are associative matrices. e is word embedding. The final representation of W_i is a dense vector formed by concatenating the left and right contexts through word embeddings, as shown in Eq. (2).

$$x_i = [L(W_i); e(W_i); R(W_i)] \tag{2}$$

In Eq. (2), x_i is the final feature representation of the i-th word. $[;]$ is vector concatenation. $e(W_i$ represents the embedding vector of W. The final feature representation of a word carries both the word's inherent semantics and its bidirectional contextual information within the sentence, providing a complete input foundation for subsequent feature abstraction. Eq. (3) is used to activate the feature representation.

$$y_i = \tanh(Ex_i + b) \tag{3}$$

In Eq. (3), y_i is the high-order semantic representation of the i-th word. The $\tanh$ is a hyperbolic tangent function. E is a linear transformation matrix. b is the bias vector. This transformation introduces a nonlinear mapping that projects the concatenated original features into a more abstract semantic representation space, enabling the model to capture more complex relationships between word meanings and context, thereby providing more discriminative inputs for feature selection in subsequent convolutional and pooling layers. The text information to be classified by the TextRCNN model can be converted into a vector form that machines can understand through pre-training. This study uses a BERT pre-trained model for transformation. BERT is a pre-trained language model based on the Transformer architecture that can efficiently adapt to a wide range of natural language processing tasks (Kumar et al., 2023). Its structure is shown in Fig. 2.

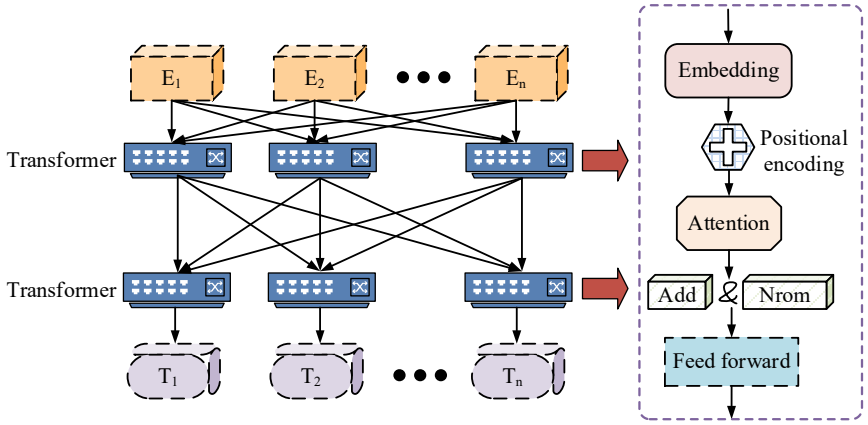


Fig. 2. BERT structure (Source: self-drawn)

In Fig. 2, the BERT model mainly consists of three core layers: input, intermediate processing, and output. The input layer receives text information containing word vectors, paragraph identifier vectors, and position vectors. The intermediate processing layer, Transformer encoder, performs nonlinear feature transformation and layer normalization on the information, and finally outputs an adaptive representation. The output information can be easily fine-tuned for natural language processing tasks. The input representation formula for BERT is shown in Eq. (4) (Praveen and Vajrobol, 2023).

$$\begin{cases} h = h_t + h_s + h_p \\ \eta = \eta \dfrac{t}{t_{\text{warmup}} max} \end{cases} \tag{4}$$

In Eq. (4), h is the input text for BERT. h_t is a word vector. h_s is the paragraph identifier vector. h_p is the position

vector. η is the learning rate of BERT in the early training stages. t_{warmup} is the number of preheating steps. The length of the three vectors is N, and the dimension of the embedded word vector is d. The input text sequence is mapped using the BERT model vector, paragraph identification vector, and position vector processing methods. The mapping formula is shown in Eq. (5).

$$\begin{cases} X = [CLS]x_1x_2x_3\ldots x_n[SEP] \\ v = \text{InputRepresentation}(X) \end{cases} \tag{5}$$

In Eq. (5), X is the set of sequences input to this article. $[CLS]$ is the first marker of the input sequence. $[SEP]$ is the separator mark between two sentences. v is the final input representation. After the final input v BERT is encoded by a multi-layer accumulated Transformer encoder; the corresponding classification label of the input sentence needs to be obtained, and it also needs to be classified by the classification output layer. The probability distribution over all labels is given by Eq. (6).

$$p = \text{Softmax}(v_0W + b_0) \tag{6}$$

In Eq. (6), v_0 is the hidden layer representation corresponding to the $[CLS]$ label obtained in the BERT feature extraction layer. W is the weight of the linear layer. b_0 is the bias of the fully connected layer. In summary, this study uses the BERT model to generate word vectors and feeds them into the TextRCNN model for classification, thereby constructing an automated CBC algorithm. The complete process structure is shown in Fig. 3.

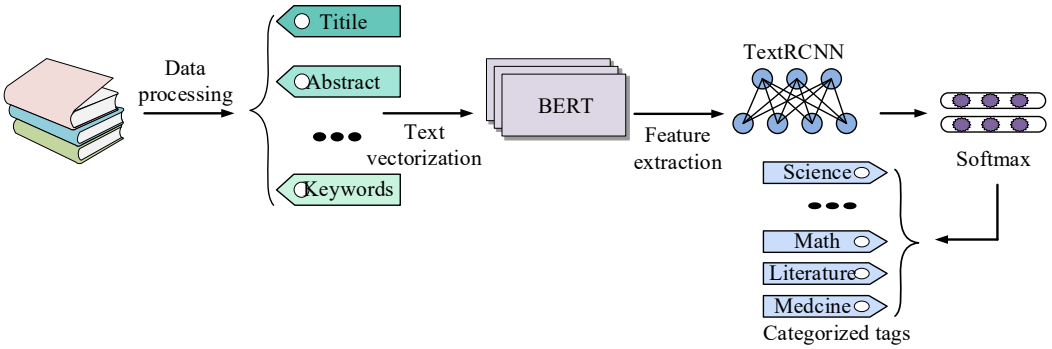


Fig. 3. Algorithm for automatic CBC (Source: self-drawn)

In Fig. 3, PLM-LCN is constructed using BERT as the pre-trained model. This model can be divided into four stages: data preprocessing, BERT text vectorization, TextRCNN feature extraction, and Softmax classifier processing. Firstly, the raw data of the book, such as title, abstract, and keywords, are used as the text to be classified by the algorithm. Then, the input text can be transformed into vector form using the BERT for subsequent NLP tasks. Next, TextRCNN is used to process the vectors produced by the BERT model, and an LSTM is used to extract sequence features. By combining CNN's feature capture of local key phrase patterns with a Softmax classifier for feature weight allocation, the classification labels for the books are determined from the resulting probability distribution.

2.2. Optimization of Classification Model integrating MHA Mechanism

To enable the constructed classification model to more comprehensively capture the rich semantic associations in Chinese text, mine deeper, multi-dimensional semantic information, and enhance its ability to understand Chinese text semantics, this study introduces Multi-Head Attention (MHA) into the model. MHA is an extended form of the attention mechanism widely used in Transformer models, an improvement over the self-attention mechanism (Zhou et al., 2023). The framework of its mechanism is displayed in Fig. 4.

In Fig. 4, the basic structure of MHA consists of four parts: the input layer, the multi-head splitting, the attention calculation, and the output fusion. First, the input text vector is split into multiple groups. Each group contains three feature representations: query, key, and value. These feature representations can capture diverse information about text from different perspectives. Subsequently, after the linear transformation, Dot Product Attention (DPA) operation is performed, and the dot product similarity between the query vector and the key vector is calculated through this operation to determine the semantic association strength between the current word and other words. The next step is to weigh and sum the value vectors to generate a perceptual representation of the context, and then concatenate the output vectors of multiple DPA points into a unified output representation through the final linear transformation. This design enables the model to simultaneously focus on multiple semantic relationships of words, capturing both grammatical structural features and identifying associations between professional terms. Among them, the formulas for inputting text query, key, and value vectors are shown in Eq. (7) (Tulapurkar et al., 2023).

$$\begin{cases} Q_i = XZ_i^Q \\ K_i = XZ_i^K \\ V_i = XZ_i^V \end{cases}, \ X \in R^{n \times d_{\text{model}}} \tag{7}$$

In Eq. (7), X is the input sequence. n is the sequence length. d_{model} is the model dimension. Z_i^Q, Z_i^K, and Z_i^V are trainable projection matrices corresponding to the i-th head. Q_i, K_i, and V_i are the vector matrices of query, key, and value of the i-th head. The DPA formula is shown in Eq. (8) (Ma et al., 2025).

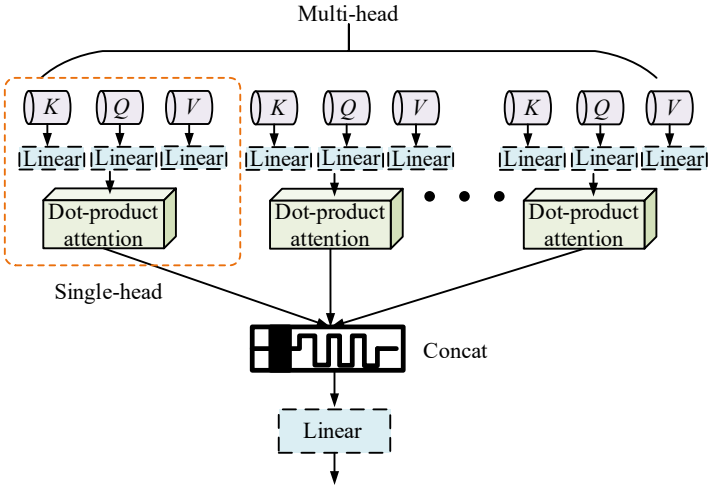


Fig. 4. MHA structure (Source: self-drawn)

$$\text{Attention}\left(Q_i, K_i, V_i\right) = \text{Softmax}\left(\frac{Q_i K_i^T}{\sqrt{d_k}}\right)V_i \tag{8}$$

In Eq. (8), d_k is the feature dimension of the key vector. $Q_i K_i^T$ is a dot product similarity matrix. The splicing formula produced by MHA is given by Eq. (9) (Benrachou et al., 2023).

$$\text{MultiHead}(Q, K, V) = \text{Concat}(head_1, \ldots, head_n)O \tag{9}$$

In Eq. (9), $head_1$ is the output of the i-th head in the DPA formula. O is the output projection matrix, which maps the concatenated result back to the original dimension. The residual connection and layer normalization formula are shown in Eq. (10).

$$\begin{cases} X' = LayerNorm(X + \text{Sublayer}(X)) \\ \text{Sublayer}(X) = \text{MultiHead}(Q, K, V) \end{cases} \tag{10}$$

In Eq. (10), X' is the residual connection output result. $\text{Sublayer}(X)$ is the output of a sub-layer that transforms the input sequence X. $LayerNorm$ is the normalization function, and its formula is shown in Eq. (11).

$$\begin{cases} LayerNorm(C) = \gamma \cdot \frac{C-\mu}{\sqrt{\sigma^2+\varepsilon}} + \beta \\ C = X + \text{Sublayer}(X) \end{cases} \tag{11}$$

In Eq. (11), C is the residual sum. γ is a learnable scaling factor. μ and σ^2 are the mean and variance of hourly features. ε is a minimum constant to prevent overflow when the variance is zero. β is a learnable bias term. Layer normalization normalizes the input of each layer to maintain stable feature distributions across all dimensions, thereby accelerating model convergence and enhancing training robustness. Finally, the complete expression of MHA is shown in Eq. (12) (Jiang et al., 2023).

$$D_v = \sum_{i=1}^{n} \frac{1}{z} exp\left(\frac{Q_i K_i}{\sqrt{d_k}}\right)V_i \tag{12}$$

In Eq. (12), D_v is the final output result after MHA. z is a standardized parameter. MHA is introduced into the TextRCNN to improve LSTM performance on classification-relation tasks. The Bidirectional LSTM (BiLSTM) model incorporating MHA is shown in Fig. 5.

In Fig. 5, the BiLSTM incorporating MHA comprises input, encoding, LSTM, MHA, and output layers. Firstly, after receiving the original text sequence, the input layer maps words to dense word-vector representations via the embedding layer, capturing semantic information. Subsequently, the BiLSTM layer performs bidirectional encoding of the word-vector sequence. The forward LSTM captures historical context from left to right, while the backward LSTM learns future context from right to left. The concatenation of the two forms' hidden states provides a complete contextual representation of each word. The newly added MHA layer focuses on key positions by calculating inter-word correlation weights. Finally, the output layer integrates attention-weighted feature representations to generate classification labels. In summary, the complete structure of the classification model integrated with MHA is shown in Fig. 6.

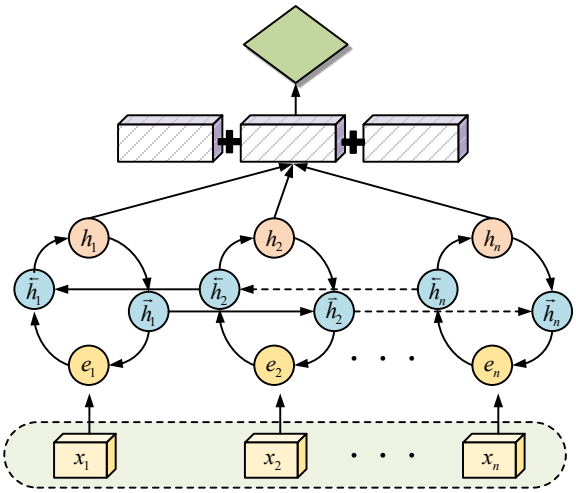


Fig. 5. BiLSTM with MHA (Source: self-drawn)

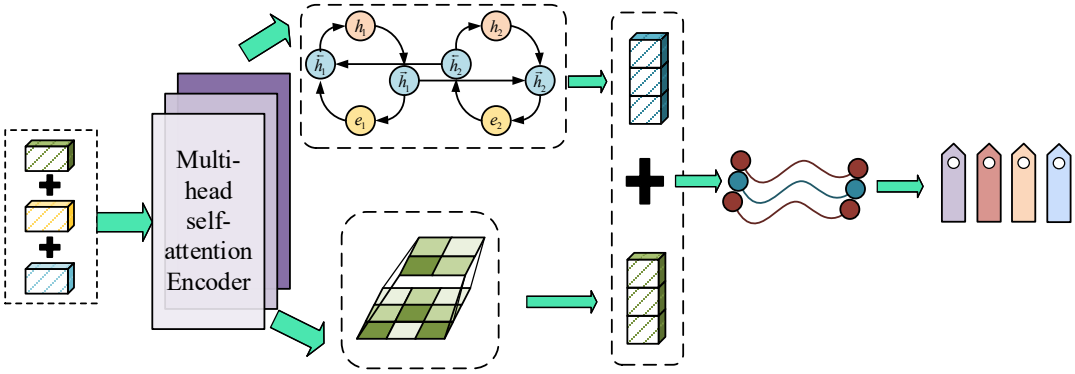


Fig. 6. PLM-LCN structure (Source: self-drawn)

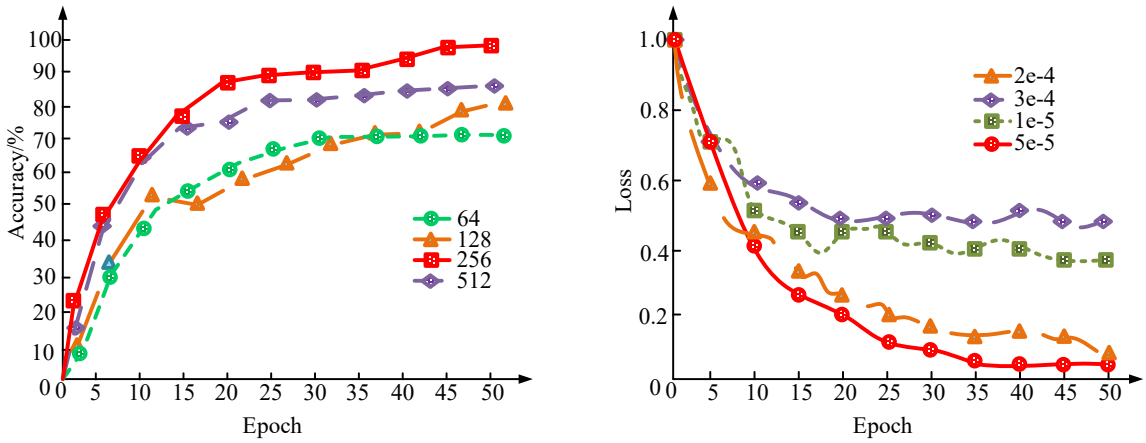
In Fig. 6, the book classification model that integrates the MHA mechanism, namely MHA-BERT, mainly consists of four steps. Firstly, the raw book data are transformed into a vector form that can be recognized by MHA through BERT. Then, MHA inputs the first label vector from the book data into the RNN model to extract features. Next, BiLSTM is used to enhance the extracted features and obtain the data features required for the task. Finally, the classification labels of the books to be classified are output by the classifier.

3. Results

3.1. Performance Testing of MHA-BERT Intelligent Classification Model

To verify the model's performance, the CPU is an Intel Core i9-13900K, the GPU is an NVIDIA RTX 4090, the memory is 64GB DDR5 6400MHz, and the operating system is Windows 11. The Chinese Library Classification Metadata Dataset (CLC-Books) and the National Science and Technology Library Chinese Book Classification Dataset (NSTL-Books) are selected for performance testing. The CLC Books dataset is constructed based on the authoritative "Chinese Library Classification System" and contains 500,000 bibliographic records, covering all 22 basic categories, including philosophy, social sciences, natural sciences, and comprehensive books. Each data set contains fields such as book title, author, abstract, publisher, and classification number, which are suitable for verifying the model's accuracy in academic classification tasks. NSTL Books is sourced from the National Science and Technology Library and the Literature Resource Library, mainly comprising professional books in the natural and social sciences, with a total of about 300,000 bibliographic records. Its classification labels are annotated according to the Chinese Library Classification, and are suitable for testing professional terminology recognition and cross-disciplinary classification ability. For model training and evaluation, a stratified random sampling strategy was adopted for both datasets. The training set, validation set, and test set were divided in the same proportions as the original dataset, with an 8:1:1 ratio. Ensure a relatively balanced distribution of categories across datasets during partitioning to avoid bias in model evaluation from category imbalance. The training set is used to learn and optimize model parameters; the validation set is used for hyperparameter tuning and to evaluate the early stopping strategy during training; and the test set is used only for independent evaluation of the final model performance. It does not participate in any parameter updates or selections during the training and validation phases. The experimental parameters are set as follows: 50 training rounds, 32 batch sizes, 768 BERT hidden layer dimensions, 12 Transformer encoder layers, 12 attention heads, 3 convolution kernels in TextRCNN, 256 convolution kernels, 128 LSTM hidden units, AdamW optimizer, 0.01 weight

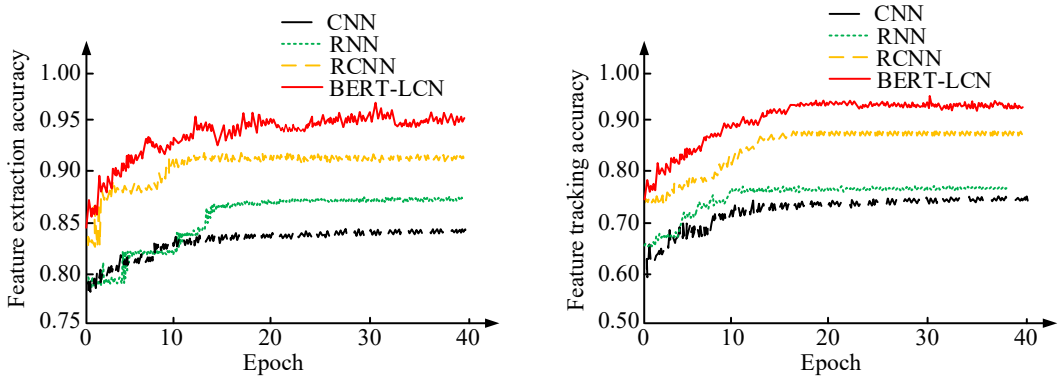
decay coefficient, and 10% preheating steps of the total training steps are used for learning rate scheduling. The above parameter values were determined from preliminary pilot experiments and were moderately adjusted on the validation set. All experiments used the same random seeds to ensure the reproducibility of the results. This study first conducts value selection tests for two hyperparameters that significantly impact model performance: the input text sequence length n and the learning rate η of BERT. The test results are exhibited in Fig. 7.



(a) Text length (b) Learning rate

Fig. 7. Two types of hyperparameter value testing (Source: self-drawn)

Fig. 7(a) compares the model's classification accuracy across different text sequence lengths at a fixed learning rate. Fig. 7 (b) compares the loss functions of different learning rates under a fixed text sequence length. In Fig. 7(a), when the Length of the Text Sequence (TSL) is 64, the model's classification accuracy increases with training epochs, but the highest value is only 69.3%. When the TSL is 128 or 512, accuracy increases over 64, with the highest accuracies at 73.6% and 84.7%, respectively. When the TSL is 256, the highest accuracy can reach 96.8% after 50 training sessions. In Fig. 7(b), when the learning rates are set to 3e-4 and 1e-5, the model fails to fully converge as training epochs increase, and the minimum loss values are 0.58 and 0.42, respectively. When the value is 2e-4, the model's minimum loss is 0.14, indicating good convergence. When the learning rate is 5e-5, the model's final loss is 0.12, which is lower and more stable than when set to 2e-4. In summary, this study determines that when the text sequence length n of the input model is 256, and the learning rate η The value of BERT is 5e-5. The parameter configuration of the PLM-LCN model fused with BERT is optimal. This study continues model ablation testing on two datasets, CLC-Books and NSTL-Books, as shown in Fig. 8.



(a) CLC-Books (b) NSTL-Books

Fig. 8. Ablation test results in two datasets (Source: self-drawn)

Fig. 8(a) and (b) show the ablation results of each component in the BERT-LCN model obtained by combining BERT as a pre-trained model with the PLM-LCN method on two types of CLC-Books and NSTL-Books. In Fig. 8(a), on CLC-Books, the CNN's semantic feature extraction accuracy is the lowest, with a final accuracy of only 83.2%, whereas the RNN improves to 85.7%. The final accuracy of RCNN is higher than that of CNN and RNN, at 92.6%. By using the BERT-LCN model and enhancing semantic feature representation through pre-training, the accuracy is further improved to 96.2%. Similarly, in Fig. 8(b) for NSTL-Books, the accuracy of BERT-LCN ultimately stabilizes at 93.3%. This indicates that BERT effectively enhances classification models' ability to combine global text features, and the introduction of the MHA mechanism further improves their ability to extract text features. In addition, commonly used book classification models are introduced for comparison. These models include Fast Text Classification (FastText), TextRCNN, and BERT, with Precision, Recall, and F1 value as reference metrics, as listed in Table 1

Table 1. The index test results of various model

Data set	Model	Precision/%	Recall/%	F1 value/%
CLC-Books	FastText	82.3	80.8	81.1
	TextRCNN	85.7	83.5	83.8
	BERT-base	89.2	87.9	88.2
	Research model	95.8	93.2	94.4
NSTL-Books	FastText	78.6	76.9	77.0
	TextRCNN	81.4	79.8	80.0
	BERT-base	86.5	84.7	85.6
	Research model	92.7	90.4	93.8

In Table 1, for CLC-Books, the research model achieves the highest precision (95.8%), recall (93.2%), and F1 (94.4%), which are significantly better than those of FastText, TextRCNN, and BERT-base models. On NSTL-Books, the values of the three reference indicators for all four models have decreased compared to CLC-Books, but the research models still demonstrate excellent performance. Compared to other models, it still maintains an advantage in key indicators, including accuracy, recall, and F1 score, at 92.7%, 90.4%, and 93.8%, respectively. This indicates that the PLM-LCN model, fused with BERT, can maintain high classification accuracy and stability across different scenarios and datasets.

3.2. Application Effect of MHA-BERT Intelligent Classification Model

To verify the model's classification performance in practice, a library system at a university in China was used for testing, and all Chinese books in the university's library collection were retrieved, totaling about 850,000 volumes. According to the category numbering from A to Z in the Chinese Library Classification System (Chinese Library Classification), the study first extracts 500 bibliographic records at equal intervals from each category to form an initial sample set. Subsequently, based on whether each book's keywords and classification numbers involve two or more major categories in the Chinese Library Classification, it will be marked as a "single discipline" or "interdisciplinary" book. On this basis, a multi-stage random sampling method was used to extract three sub-samples from the labeled results: a sample containing only books from a single discipline (1000 volumes), a sample containing books from two disciplines (800 volumes), and a sample containing books from three or more disciplines (600 volumes), corresponding to three different testing environments. During the sampling process, for interdisciplinary books, the study further checked the "Category 2" and "Category 3" fields in their cataloging data to confirm the authenticity of interdisciplinary attributes. At the same time, different models were compared and tested using Area Under Curve (AUC) as the indicator, and the test results are shown in Fig. 9.

In Fig. 9, when only a single discipline book is included in the test sample, the constructed model performs best, with an AUC of 83.7%, which is 11.4%, 24.8%, and 45.1% higher than FastText, TextRCNN, and BERT-base, respectively. In the cross-disciplinary book sample test, the research model's AUC is 78.8%, outperforming other models. Specifically, in a complex testing environment that includes interdisciplinary books, the research model still maintains high classification accuracy with an AUC value of 72.4%. The above results indicate that the proposed model offers significant advantages in addressing the practical challenge of interdisciplinary book classification in library management. In the daily management of libraries, interdisciplinary books are often difficult to classify accurately into a single category, leading to difficulties in reader retrieval and low efficiency in using library resources. The proposed model can effectively reduce the misclassification rate of interdisciplinary books by enhancing the ability to capture cross-domain semantic associations, thereby improving the scientific organization of collections and the hit rate of reader retrieval. It has clear practical application value. To verify the statistical significance of the above differences, the Bootstrap resampling method (with repeated sampling of 2000 samples) was used to estimate the confidence intervals for the AUC values of each model across various testing environments, with a confidence level of 95%. At the same time, DeLong's non-parametric tests were used to pairwise compare AUCs across models to determine whether the differences were statistically significant. The test results showed that across all three testing environments, the AUC differences between the research model and the comparison models were statistically significant ($p<0.05$), indicating that the observed performance advantages were not due to random fluctuations. To ensure the reliability of statistical inference, Bootstrap and DeLong tests were independently performed on the test set, and all compared models used identical data samples. The testing process did not involve any data from the training or validation sets, thereby avoiding interference from data leakage in significance estimation. The p-values for AUC differences among models were adjusted using Bonferroni correction to control the class I error rate in multiple comparisons. After correction, all differences remained significant ($p<0.05$), further supporting the statistical credibility of the models' performance improvement in this study. This study continues to compare the average retrieval time and resource utilization of multiple models in three different scenarios, as shown in Table 2.

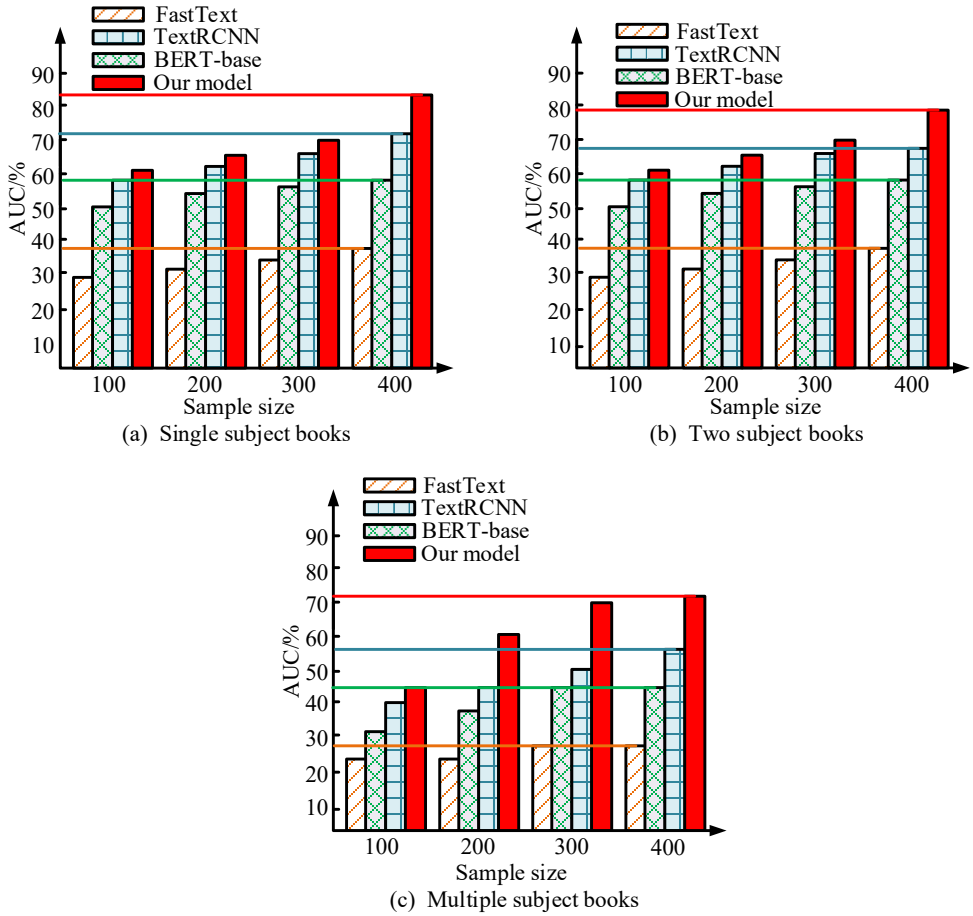


Fig. 9. AUC results of different sample settings (Source: self-drawn)

Table 2. Index test of various models in different situations

Situation	Model	Average search time/ms	Occupancy rate/%
Single subject	FastText	52.1	40.3
	TextRCNN	43.7	46.7
	BERT-base	37.3	55.2
	Our model	24.8	45.9
Two subject	FastText	96.6	62.6
	TextRCNN	87.1	73.8
	BERT-base	76.9	70.1
	Our model	62.6	63.8
Multi-subject	FastText	113.7	80.0
	TextRCNN	107.3	89.7
	BERT-base	94.6	82.3
	Our model	72.7	76.2

In Table 2, as the complexity of the testing environment increases, the time required for all four models to complete a book search increases to some extent. However, the average times for the research model to complete a book search are 24.9ms, 62.6ms, and 72.7ms, all lower than those of other models. In terms of computational resource utilization, the research model's utilization rate is lower than that of BERT-base but higher than the other two models, with occupancy rates of 45.9%, 63.8%, and 76.2% across the three cases. This is because the model's complexity is higher, resulting in a higher occupancy rate than the general base model. In summary, this model can maintain a relatively efficient retrieval response speed while ensuring high classification accuracy. In practical library management scenarios, retrieval efficiency is directly related to readers' user experience and librarians' work efficiency. This model can achieve fast classification and retrieval responses with limited computing resources, given the increasing proportion of interdisciplinary and multidisciplinary books in real collection environments, which helps reduce reader waiting time, improve circulation service efficiency, and provide technical support for intelligent management of large-scale collection resources. To further validate the user-side

effectiveness of the proposed model in an actual library retrieval system, the study integrated the model into the classification and sorting module of an existing retrieval system in a university library and optimized the original system accordingly. To evaluate the optimization effect, the research team recruited evaluation participants from in-service librarians, professors, and undergraduate students in the library. The sample sizes for the three groups are: 15 librarians, 20 professors (covering disciplines such as science and engineering, humanities, and social sciences), and 30 students (across different undergraduate grades). All participants voluntarily participated and were informed prior to the experiment that the purpose of the evaluation was to "investigate the user experience of the library retrieval system", without disclosing specific information about the differences in system versions. The evaluation uses a 10-point scale (1 point represents "very poor" and 10 points represent "very good"), and participants independently rate the four dimensions of the search system's ease of operation, search matching, user satisfaction, and overall performance. To control bias, the evaluation is conducted in two rounds: the first evaluates the original system, and the second follows a two-week interval. At this point, the system has completed optimization and upgrade, but participants are only informed to "re-evaluate after routine system maintenance" to maintain a blind state. The two rounds of evaluation use identical questions and rating scales, and participants are not informed of their ratings after the first round to avoid memory anchoring effects. All rating data were collected anonymously, and the arithmetic mean of the scores for each dimension was taken as the final statistical result, as shown in Fig. 10.

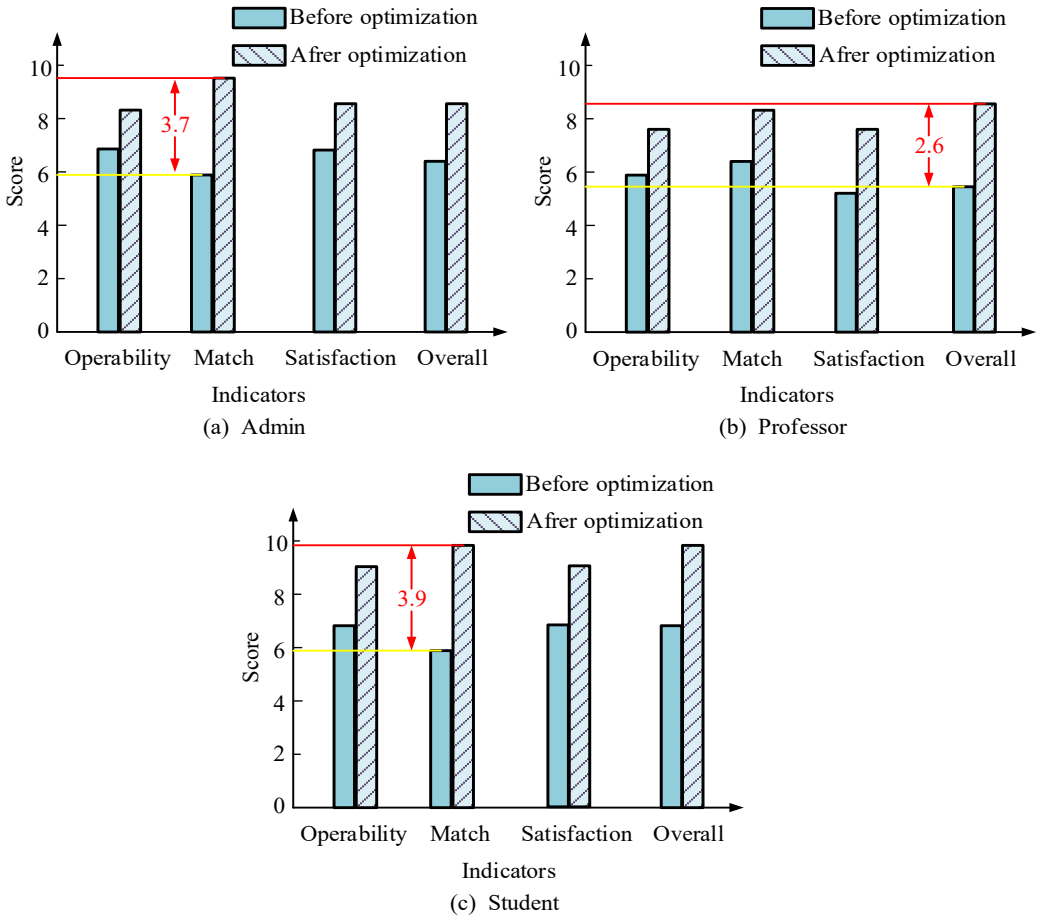


Fig. 10. Scoring results for the three categories (Source: self-drawn)

Figs. 10(a) - (c) show the ratings of librarians, university professors, and current students on the system before and after improvement. Fig. 10 shows that the ratings of the three groups of people for various indicators of the improved library system are higher than those of the original system, and the ratings for all indicators of the improved system are above 7 points. In Figs. In 10(a) and (c), librarians and students are the two groups with the highest frequency of use of the on-campus library system, and they report a particularly significant improvement in search matching with the improved system. The two indicators show increases of 3.7 and 3.9 points in the system's search matching score before and after the improvement, making it the indicator with the highest improvement rate among the four. In Fig. 10(b), the group of university professors shows the greatest overall improvement in their ratings of the improved library system compared to before, with an increase of 2.6 points. The above results indicate that the library retrieval system enhanced by this model has significantly improved in operational convenience, retrieval matching, and user satisfaction. In the context of library management, user satisfaction with retrieval directly reflects the contribution of classification model accuracy to actual service quality. This model improves the classification accuracy of interdisciplinary and multidisciplinary books, enabling readers to quickly and accurately find the required literature, thereby enhancing the overall efficiency of library knowledge services and verifying its practical application value in the intelligent classification of Chinese books in university libraries.

4. Conclusion

The classification of Chinese books in university libraries faces challenges posed by semantic complexity and an interdisciplinary nature. Traditional methods have certain limitations in global semantic understanding and collaborative optimization of local features. Given this, the paper innovatively designed an intelligent classification model that integrated BERT and PLM-LCN. In the experiment, when the input sequence length was 256, and the BERT learning rate was 5e-5, the model achieved classification accuracies of 96.8% and 93.3% on the CLC-Books and NSTL-Books datasets, significantly better than FastText, TextRCNN, and BERT-base, with improvements of about 6%-15%. In addition, the model performed well on key indicators, achieving 95.8% precision, 93.2% recall, and 94.4% F1 score, which improved by about 2%-12% over the FastTex model, demonstrating its advantages in semantic feature extraction and classification performance. In the simulation test of interdisciplinary book classification, the AUC values of the model for books from a single discipline, across two disciplines, and across multiple disciplines were 83.7%, 78.8%, and 72.4%, respectively, which were about 11%-43% higher than those of the traditional model. Especially in complex classification scenarios, the model demonstrated stronger adaptability and robustness. In a book classification test across multiple disciplines, the model's average retrieval time was 72.7ms. User ratings after deployment to university library systems showed significant improvements in retrieval matching and overall satisfaction, with the highest scores increasing by 3.9 and 2.6 points, confirming the model's practical value. In summary, this study effectively addressed deep semantic parsing and interdisciplinary intersection issues in CBC by combining BERT's global semantic understanding ability with PLM-LCN's local feature fusion mechanism, providing an efficient solution for intelligent library management. Based on the above findings, library managers can make adjustments in the following decision-making process. Firstly, in the cataloging workflow, managers can delegate the traditional initial classification process, which relies on manual interpretation, to the model for automatic completion and shift the focus of cataloging librarians' work from volume-by-volume classification to quality sampling and anomaly review of classification results, thereby optimizing manpower allocation. Secondly, regarding the organizational process of library resources, the model's effective recognition of interdisciplinary books suggests that managers should re-examine the limitations of single classification-number indexing in the existing library system. It may be considered adding a "related discipline" field to the bibliographic data to enhance the discoverability of interdisciplinary resources. Finally, in terms of technical facility planning, the model's resource occupancy data across different complexity scenarios can serve as a reference for hardware selection and system architecture decisions, helping managers strike a balance between accuracy and response efficiency suited to the library's conditions. However, there is still room to optimize the model's ability to handle lengthy texts and reduce resource utilization, such as for entire book content. Future work plans to introduce lightweight technology to improve efficiency and explore the model's scalability for cross-language book classification tasks.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

Declaration of Artificial Intelligence (AI) Tools

The author declares that no artificial intelligence (AI) tools were used in the preparation, writing, editing, or revision of this manuscript.

References

Bahrani, P., Minaei-Bidgoli, B., Parvin, H., Mirzarezaee, M., and Keshavarz, A. (2024). A hybrid semantic recommender system enriched with an imputation method. *Multimedia Tools and Applications*, 83(6), 15985-16018. doi: 10.1007/s11042-023-15258-4

Benrachou, D. E., Glaser, S., and Elhenawy, M. (2023). Improving efficiency and generalisability of motion predictions with deep multi-agent learning and multi-head attention. *IEEE Transactions on Intelligent Transportation Systems*, 25(6), 5356-5373. doi: 10.1109/TITS.2023.3339640

Cui, F., and Gao, L. (2024). Book Classification and Recommendation for University Libraries using Grey Correlation and Bayesian Probability. *Scalable Computing: Practice and Experience*, 25(4), 2358-2370. doi: 10.12694/scpe.v25i4.2812

Ding, N., Qin, Y., Yang, G., Wei, F., Yang, Z., Su, Y., and Hu, S. (2023). Parameter-efficient fine-tuning of large-scale pre-trained language models. *Nature Machine Intelligence*, 5(3), 220-235. doi: 10.1038/s42256-023-00626-4

Dong, C. and Zhang, K. (2023). An improved cascade RCNN detection method for key components and defects of transmission lines. *IET Generation, Transmission & Distribution*, 17(19), 4277-4292. doi: 10.1049/gtd2.12948

Harisanty, D., Anna, N. E. V., Putri, T. E., and Firdaus, A. A. (2025). Is adopting artificial intelligence in libraries urgency or a buzzword? A systematic literature review. *Journal of Information Science*, 51(2), 511-522. doi: /10.1177/01655515221141034

Huang, Y. H. (2024). Exploring the implementation of artificial intelligence applications among academic libraries in Taiwan. *Library Hi Tech*, 42(3), 885-905. doi: 10.1108/LHT-03-2022-0159

Jiang, S., Lu, M. and Hu, K. (2023). Personalized federated learning based on multi-head attention algorithm. *International Journal of Machine Learning and Cybernetics*, 14(11), 3783-3798. doi: 10.1007/s13042-023-01864-z

Kumar, T., Mahrishi, M., and Sharma, G. (2023). Emotion recognition in Hindi text using multilingual BERT transformer. *Multimedia Tools and Applications*, 82(27), 42373-42394. doi: 10.1007/s11042-023-15150-1

Lee, P., Fyffe, S., Son, M., Jia, M., and Yao, Z. (2023). A paradigm shift from “human writing” to “machine generation” in personality test development: An application of state-of-the-art natural language processing. *Journal of Business and Psychology*, 38(1), 163-190. doi: 10.1007/s10869-022-09864-6

Liu, B., Guan, W., Yang, C., Fang, Z., and Lu, Z. (2023). Transformer and graph convolutional network for text classification. *International Journal of Computational Intelligence Systems*, 16(1), 1-15. doi: 10.1007/s44196-023-00337-z

Ma, X., Chen, T., and Wang, Y. (2025). Dynamic process monitoring based on dot product feature analysis for thermal power plants. *IEEE/CAA Journal of Automatica Sinica*, 12(3), 563-574. doi: 10.1109/JAS.2024.124908

Min, C., Zhang, R., and Liu, J. (2025). Bilingual generated text detection through semantic and statistical analysis. *Intelligent Data Analysis*, 29(5), 1248-1260. doi: 10.1177/1088467X241307192

Pal, S., Roy, A., Shivakumara, P., and Pal, U. (2023). Adapting a Swin Transformer for License Plate Number and Text Detection in Drone Images. *Artificial Intelligence and Applications*, 1(3), 145-154. doi: 10.47852/bonviewAIA3202549

Praveen, S. V. and Vajrobol, V. (2023). Understanding the perceptions of healthcare researchers regarding ChatGPT: a study based on bidirectional encoder representation from transformers (BERT) sentiment analysis and topic modeling. *Annals of Biomedical Engineering*, 51(8), 1654-1656. doi: 10.1007/s10439-023-03222-0

Shen, S., Liu, J., Lin, L., Huang, Y., Zhang, L., Liu, C., Feng, Y., and Wang, D. (2023). SsciBERT: A pre-trained language model for social science texts. *Scientometrics*, 128(2), 1241-1263. doi: 10.1007/s11192-022-04602-4

Shi, Y., Zhang, X., and Yu, N. (2023). PL-Transformer: a POS-aware and layer ensemble transformer for text classification. *Neural Computing and Applications*, 35(2), 1737-1751. doi: 10.1007/s00521-022-07872-4

Ten Lia, V. (2023). Topic modeling in automatic of news texts. *Terra Linguistica*, 52(2), 77-91. doi: 10.18721/JHSS.14207

Tulapurkar, H., Banerjee, B., and Buddhiraju, K. M. (2023). Multi-head attention with CNN and wavelet for classification of hyperspectral image. *Neural Computing and Applications*, 35(10), 7595-7609. doi: 10.1007/s00521-022-08056-w

Umer, M., Imtiaz, Z., Ahmad, M., and Nappi, M. (2023). Impact of convolutional neural network and FastText embedding on text classification. *Multimedia Tools and Applications*, 82(4), 5569-5585. doi: 10.1007/s11042-022-13459-x

Yixuan, L., Dongbo, W., and Jiawei, L. (2023). Aeroengine blade surface defect detection system based on improved faster RCNN. *International Journal of Intelligent Systems*, (1), 1992-1994. doi: 10.1155/2023/1992415

Zhang, H., Song, H., Li, S., Zhou, M., and Song, D. (2023). A survey of controllable text generation using transformer-based pre-trained language models. *ACM Computing Surveys*, 56(3), 1-37. doi: 10.48550/arXiv.2201.05337

Zhou, A., Ma, Y., and Ji, W. (2023). Multi-head attention-based two-stream EfficientNet for action recognition. *Multimedia Systems*, 29(2), 487-498. doi: 10.1007/s00530-022-00961-3

Bo Hong earned his master's degree in 2009 and has been working as a librarian since 2014. Over the years, he has served in several departments, including the foreign-language book stack, compact stack, document resources department, and ancient books special collections department. His current work focuses on promoting the intelligent transformation of library cataloging, collection management, and resource retrieval, and he actively conducts research on the application of intelligent information processing and information science in digital libraries.