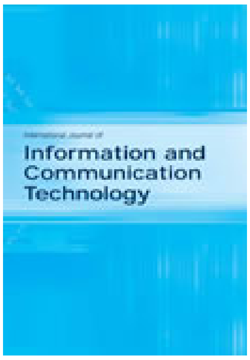


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Analysis and application of college students' academic emotions based on deep learning and psychological status

Yu Wang

Article History:

Received:	08 December 2024
Last revised:	18 December 2024
Accepted:	18 December 2024
Published online:	13 February 2025

Analysis and application of college students' academic emotions based on deep learning and psychological status

Yu Wang

School of Mechanical Engineering,
Henan Institute of Technology,
Xinxiang, 453003, China
Email: wangyu@hait.edu.cn

Abstract: With the rapid development of artificial intelligence technology, the application of deep learning in the field of education is gradually increasing. This article proposes a method for analysing college students' academic emotions that combines deep learning models with psychological state assessment, in order to help educators gain a more comprehensive understanding of students' emotional changes and optimise teaching effectiveness. By constructing an emotion classification model based on a bidirectional long short-term memory network (Bi LSTM), this paper conducts emotion recognition on academic related short text data of college students, and verifies the emotional state with a psychological scale to ensure the accuracy and effectiveness of emotion recognition. The experimental results show that the proposed method exhibits high accuracy and robustness in emotion classification tasks.

Keywords: deep learning; academic emotions; psychological condition; emotion analysis; college student.

Reference to this paper should be made as follows: Wang, Y. (2025) 'Analysis and application of college students' academic emotions based on deep learning and psychological status', *Int. J. Information and Communication Technology*, Vol. 26, No. 3, pp.125–139.

Biographical notes: Yu Wang studied at Yan'an University and received his Bachelor's degree in 2009 and Master's degree in 2012. Currently, he works in Henan Institute of Technology. His research focuses on deep learning, ideological and political education, as well as entrepreneurship and employment guidance.

1 Introduction

With the rapid development of information technology, the application of artificial intelligence (AI) and deep learning technology in the field of education is becoming increasingly widespread, bringing significant changes to the optimisation of teaching content, analysis of student learning behaviour, and improvement of educational equity. Among them, student emotions, as an important factor affecting learning outcomes, have gradually become a key area of educational research in recent years (Zhang and Aslan, 2021). Therefore, researching how to identify and analyse students' academic emotions through technological means, and transform these achievements into effective tools in teaching practice, is an important direction in current educational research and AI applications.

Traditional emotional research mostly relies on qualitative methods in psychology, such as interviews, observations, and questionnaire surveys. Although these methods can delve into the psychological mechanisms behind emotions, they have certain limitations in terms of data size and analysis efficiency, making it difficult to meet the needs of large-scale educational environments (Koolagudi and Rao, 2012). The emotion recognition method based on deep learning can efficiently process massive amounts of unstructured data, including text, speech, images, etc., with high accuracy and strong generalisation ability. Especially in the field of natural language processing (NLP) (Batbaatar et al., 2019), deep learning models such as bidirectional long short-term memory networks (Bi LSTM) (Zhao et al., 2019) have shown outstanding performance in emotion classification tasks due to their powerful modelling ability for time series data.

In the study of academic emotions, emotions are closely related to students' psychological conditions. Psychological factors such as academic pressure, learning motivation, and self-efficacy play an important regulatory role in students' emotional states. Modern college students face multiple challenges in their learning process, including heavy academic tasks, increased competitive pressure, and complex changes in their living environment. These factors lead to frequent fluctuations in academic emotions among college students, and some students may even fall into academic fatigue or psychological crisis due to negative emotions.

In recent years, the academic community has gradually deepened its research on the combination of emotion analysis and deep learning in the field of education. Especially in understanding the dynamic changes of college students' academic emotions and their complex relationship with psychological conditions, a large number of innovative works based on deep learning and emotional computing have emerged.

Emotion recognition is one of the core applications of emotion computing in educational technology. For example, Chakraverthi et al. (2022) proposed a hybrid model combining CNN and LSTM, which can extract emotional features from students' online discussions and significantly improve classification accuracy. Transformer and BERT, have been widely applied in text emotion recognition tasks. Chen et al. (2022) used a multi task learning approach to achieve joint modelling of emotion classification and emotion polarity detection, laying the foundation for multi-level analysis of emotions.

In addition, emotion recognition has gradually developed from a single mode to a multimodal direction. In addition to text, speech, facial expressions, and physiological signals (such as heart rate, skin conductance response, etc.) are also widely used for modelling. For example, Gupta et al. (2023) developed a deep learning framework that integrates speech and facial expressions, realising the application of multimodal emotion recognition in online education.

The main applications of emotion analysis in education include teaching quality assessment, learning process optimisation, and personalised education design. MOOC platforms and intelligent tutoring systems are the main research objects. Anwar et al. (2023) studied online education scenario design based on emotion prediction models, which can dynamically adjust teaching content to match students' current emotional states. In addition, Abdullah et al. (2021) proposed an algorithm for optimising learning path recommendation through large-scale emotion analysis data, demonstrating the potential of emotion computing in large-scale personalised education.

Emotions are closely related to mental health, especially in high-pressure university learning environments. Sharma and Giannakos' (2020) study suggests that emotional states and mental health jointly affect students' academic performance. The emotion and mental health analysis model based on multimodal data has gradually become a research hotspot. For example, Karamimehr et al. (2023) constructed an emotion detection model that combines psychological states to dynamically adjust teaching strategies by monitoring students' learning progress and psychological states. In addition, real-time mental health prediction technology has also been introduced into educational settings to warn students of possible psychological problems, thereby improving the timeliness of academic support (Kim, 2012).

In recent years, the application of multimodal learning models in the field of emotion analysis has rapidly emerged. For example, Yadegaridehkordi et al. (2019) proposed a multimodal attention model that integrates text, speech, and image features into the same network for emotion recognition tasks. These technologies demonstrate strong potential in modelling student behaviour data.

Although the application of sentiment analysis in education has achieved significant results, there are still the following challenges. The acquisition and annotation of high-quality emotional data are crucial for model training, and emotional data in educational settings has characteristics such as high dimensionality, unstructured nature, and situational dependence, resulting in high annotation costs (Wu et al., 2016). Due to significant differences in emotional expression across different cultural backgrounds, the generalisation performance of existing models urgently needs to be improved (Zhou and Ye, 2023). The use of physiological data or facial recognition technology in educational contexts may lead to privacy controversies. Researchers have proposed solutions such as federated learning and encrypted model training, but further optimisation is still needed (Hasan et al., 2020).

Future research should further optimise the performance of multimodal emotion recognition models. Meanwhile, the combination of mental health and academic emotions can provide a new perspective for precision education. In addition, while optimising algorithms, it is also necessary to focus on privacy protection and ethical compliance in order to achieve widespread application of technology and increase social acceptance (Barron-Estrada et al., 2019).

This article proposes an emotion analysis framework based on deep learning and psychological assessment to address the complexity and diversity of college students' academic emotions. This framework is developed from the following aspects: Firstly, a deep learning technology is used to construct an emotion classification model for emotion recognition of academic related text data of college students, such as course feedback, classroom discussions, homework texts, etc. Secondly, combining with the analysis of psychological scales, the psychological status of students is evaluated to verify the accuracy and effectiveness of emotion recognition. Finally, applying the results of emotion analysis to educational settings, providing emotional driven teaching suggestions for teachers, and developing psychological warning mechanisms to help students better cope with academic pressure and psychological challenges.

Specifically, the Bi LSTM model constructed in this article can effectively capture emotional features in text data, overcoming the shortcomings of traditional methods that rely heavily on feature engineering. In addition, psychological scales such as the

Learning Emotion Scale and the Academic Burnout Scale were integrated into the study to compensate for the bias that may arise from relying solely on algorithms. By combining technological means with psychological methods, this article aims to provide a scientific and practical solution for academic emotion analysis.

The research in this article not only has important theoretical significance, but also has extensive practical value. At the theoretical level, this study expands the application scenarios of deep learning technology in the field of education, providing a new perspective for research on educational informatisation. At the practical level, this study provides educators with a new teaching aid tool that uses emotion analysis to help teachers identify students' emotional states and optimise teaching strategies. At the same time, the psychological warning mechanism proposed in this article can timely detect potential psychological problems in students, promote the personalised development of mental health education, and lay the foundation for achieving educational equity and comprehensive development of students.

2 Related theoretical analysis

2.1 Academic emotions

Emotions are usually directly related to academic tasks, academic achievements, and learning situations. Academic emotions reflect students' emotional responses to learning content, process, and outcomes, and are important influencing factors for learning motivation, cognitive engagement, and academic performance.

According to Pekrun et al.'s control value theory, academic emotions are influenced by a sense of control and a sense of value, where a sense of control is a student's sense of control over academic tasks, and a sense of value is a student's subjective evaluation of the importance of learning tasks. In addition, some scholars believe that calmness and satisfaction should be added to the dimension of positive academic emotions, while negative academic emotions should also include fatigue and depression.

In emotional calculation, academic emotions are often represented by multidimensional variables. Assuming that a student's academic emotional state can be expressed as:

$$E = [V, A, D] \tag{1}$$

where V represents valence, reflecting the positive or negative nature of emotions, with a value range of [–1, 1]. Positive values indicate positive emotions, while negative values indicate negative emotions. A represents the activation level (arousal), reflecting the intensity of emotions, with a value range of [–1, 1]. D represents duration, which is the duration of emotional maintenance, with a range of positive values.

According to the control value theory, the formation of academic emotions can be expressed as a function:

$$E = f(C, V) \tag{2}$$

where C represents a sense of control, measuring students' level of control over academic tasks, with a range of [0, 1]. V represents a sense of value, measuring students' subjective importance of academic tasks, with a range of [0, 1].

The interaction between sense of control and sense of value significantly affects the valence and activation of emotions. For example, when students perceive a high sense of control and high value, they are prone to positive emotions (such as excitement). When the sense of control or value is low, negative emotions such as anxiety or helplessness may arise.

Academic emotions are associated with academic performance through the following formula:

$$P = g(E, M, R) \tag{3}$$

where E represents the emotional state. M represents students' learning motivation. R represents students' learning resources, including time, energy, and external support.

Academic emotions can indirectly affect academic performance by influencing cognitive processes such as attention, memory, and problem-solving abilities. Positive emotions can help improve learning motivation and attention, while negative emotions may lead to task avoidance and decreased efficiency.

2.2 Deep learning theory

Deep learning relies on the hierarchical feature learning ability of neural networks for large-scale data. Emotion analysis involves the extraction and classification of emotional features. Two core models play an important role in emotion analysis.

LSTM models time series data through memory units and gate mechanisms, and is a core tool for processing sequential inputs such as text and speech in sentiment analysis. Its core lies in avoiding the problem of gradient vanishing by dynamically adjusting the information flow through forget gates, input gates, and output gates

The basic structure of LSTM includes input gates, forget gates, and output gates, whose mechanisms ensure that the model can selectively remember or discard information.

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{4}$$

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{5}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{6}$$

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \tag{7}$$

$$h_t = o_t * \tanh\left(C_t\right) \tag{8}$$

where i_t, o_t, and f_t respectively represent the input gate, output gate, and forget gate; σ represents the sigmoid activation function, weight sets W_i, W_o, W_f represent the weights of the input gate, output gate, and forget gate, W_c is the weight of the cell state, and h_t h_t is the new state. The LSTM network structure is shown in Figure 1.

The core component of artificial neural networks is artificial neurons, which generally use a multi input single output mode for information processing, simulating the structure of biological neurons. Neural networks are constructed using a large number of artificial neurons. These neurons are usually arranged in an ordered manner in a directed acyclic graph, forming what is known as a feedforward neural network. By distributing neurons at different levels and allowing only neurons between adjacent levels to connect to each other, this can further organise and simplify the network structure. Figure 2 shows a schematic diagram of a five layer feedforward neural network consisting of three input units and five output units, demonstrating the hierarchical and structured characteristics of the feedforward neural network. This hierarchical and structured design not only enhances the ability of neural networks to handle complex data patterns, but also provides convenience for network training and optimisation by simplifying and clarifying the functions of each layer.

Figure 1 LSTM network architecture (see online version for colours)

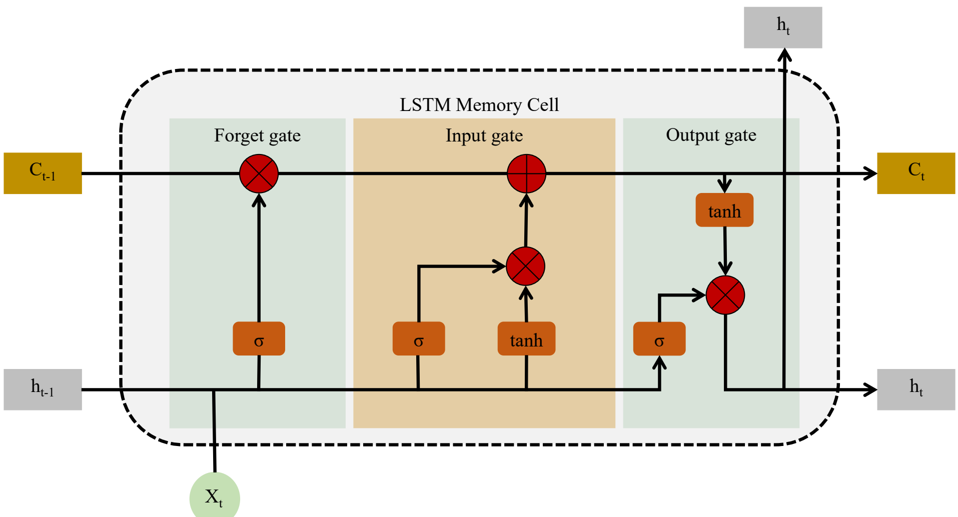
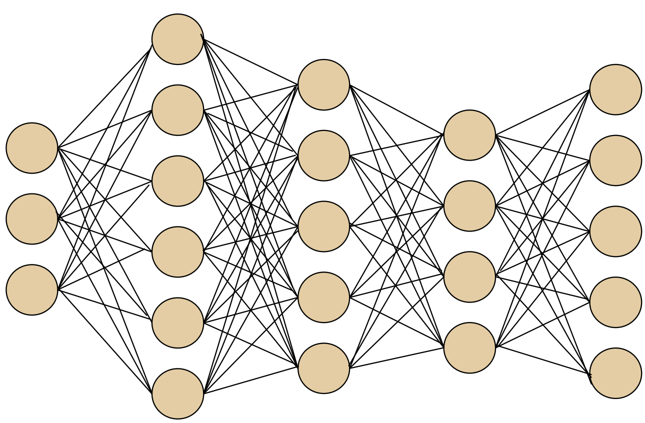


Figure 2 A feedforward neural network with three input layers and five output layers (see online version for colours)



CNN extracts local features through convolution operations and reduces model complexity by utilising parameter sharing and sparse connections, significantly improving the learning efficiency of large-scale data. Convolutional layers are the foundation of CNN, which use convolutional kernels (or filters) to perform linear transformations on input data and extract local features. The convolution operation formula is:

$$y_{i,j} = \sum_{m,n} x_{i+m,j+n} \cdot \omega_{m,n} + b \tag{9}$$

where $x_{i+m,j+n}$ is the pixel value of the input feature map, $\omega_{m,n}$ is the convolution kernel weight, b is the bias term, and $y_{i,j}$ is the convolution output, which is the value on the feature map.

To introduce nonlinear characteristics, the convolution operation will be activated by an activation function, such as ReLU:

$$f(x) = \max(0, x) \tag{10}$$

The pooling layer reduces the size of feature maps through downsampling operations, reduces computational complexity, and enhances the robustness of features.

$$y_{i,j} = \max\left(x_{i+k,j+k}\right), k \in \textit{pooling window} \tag{11}$$

CNN flattens the feature map into a one-dimensional vector through a fully connected layer and passes it to the fully connected layer for final classification or regression tasks:

$$\widehat{y} = \sigma(W \cdot x + b) \tag{12}$$

where x is the input feature vector, W is the weight matrix, b is the bias vector, and σ is the activation function.

3 Method

This article constructs an emotion classification model based on a Bi LSTM to recognise emotions in academic related short text data of college students, and verifies the accuracy and effectiveness of the recognition results with a psychological scale.

3.1 Collection and preprocessing of student academic emotional feedback text data

The evaluation and feedback texts of students on the classroom, teachers, etc. contain rich academic emotional information, which can serve as an important source for studying students' academic emotional states. Negative academic emotions can seriously affect students' physical and mental health. The main objective of this section is to identify academic emotions in student feedback texts, laying the foundation for further guiding teachers' teaching and students' learning.

1 Dataset source.

 This dataset collects students' emotional reactions and academic experiences in different learning scenarios, such as exam pressure, academic achievement, and adaptability to course difficulty, through questionnaire surveys and in-depth interviews. The questionnaire includes open-ended questions (such as 'Describe your feelings during final review') and closed ended questions (such as 'Describe your academic status in three words'). We crawl discussion content about learning and exams from university social media platforms such as Zhihu, Weibo, and Tieba. Keyword filtering includes 'learning pressure', 'exam anxiety', 'homework procrastination', etc.

2 Data processing.

   Firstly, preprocess the text in the collected dataset to remove invalid characters (such as HTML tags and emoticons), and then convert it to a unified Simplified Chinese or English format. Delete sentences unrelated to academic studies, such as advertising content or generalised emotional expressions, and remove duplicate samples.

In response to the complex and varying length of text content in the comment section of online learning platforms, this study first standardises the annotation granularity. For student feedback texts included in online learning platforms, they usually contain multiple emotions or viewpoints, and the topic categories of comments (i.e. the objects of comments) may also vary. In terms of selecting the granularity range for text annotation, the original text data is first divided into sentences and first level topic segmentation. If the same text involves evaluations from three first level dimensions: platform, course, and teacher, the original text will be divided into three sentences, with each first level dimension topic belonging to a separate sentence text. In this way, the original text becomes three paragraphs. For example, regarding the feedback text sample "I think the teacher taught very well, the course content was very rich, but sometimes there were crashes during class. It would be better to improve this point." This sample involves three evaluation objects: the teacher, the course, and the platform. Therefore, when annotating, the original example becomes three clauses, namely:

1 I think the teacher taught very well.

2 The course content is very rich.

3 Sometimes the class crashes, which would be better if we could improve it.

Firstly, based on the themes of the first level dimension, divide it into several first level dimension themes. Subsequently, based on the theme division of the second level dimension under the first level dimension, the second level dimension themes under the first level dimension were further refined and clarified to ensure the accuracy and reliability of the results, thereby helping teachers focus on the refined theme categories and clarify the direction of teaching optimisation in subsequent data analysis.

Table 1 Annotated examples

ID	*Test*	*Dimension*	*Result*
001	I think the teacher spoke very well	Teaching method	Positive
002	The course content is very comprehensive	Course content	Positive
003	Sometimes classes may crash	Media Technology	Negative

In the annotation, according to psychological theory, five types of academic related emotions are defined: positive emotions (such as satisfaction), negative emotions, neutral emotions (such as calmness and indifference), mixed emotions (such as both anxiety and anticipation), and other emotions (emotions that cannot be classified). Use an emotional dictionary to automatically annotate text, construct a keyword list, and extract academic related emotional words such as 'stress', 'tension', 'achievement', etc. To improve the generalisation ability of the model, data augmentation techniques are used to increase the sample size, such as synonym replacement, word order adjustment, and random deletion. The final samples are stored in the form of Table 1.

3.2 Emotion classification model based on Bi LSTM

The bi long short term memory network (Bi LSTM) model is a deep learning method that can capture bidirectional contextual information from sequence data, making it highly suitable for short text sentiment classification tasks. In the analysis of academic emotions among college students, it is possible to effectively handle short text corpora and achieve more accurate recognition results.

Bidirectional LSTM captures contextual information in two directions and processes both positive and negative inputs separately. The forward LSTM is:

$$\vec{h}_t = LSTM\left(E_t, \vec{h}_{t-1}\right) \tag{13}$$

The backward LSTM is:

$$\overleftarrow{h}_t = LSTM\left(E_t, \overleftarrow{h}_{t+1}\right) \tag{14}$$

The bidirectional output is:

$$h_t = \left[\vec{h}_t; \overleftarrow{h}_t\right] \tag{15}$$

where [ ] represents the connection operation.

In the process of emotion recognition, attention mechanism is used to calculate the relevant weights and strengthen the model's attention to emotion related words. The calculation formula is as follows:

$$\alpha_t = \frac{\exp\left(u_t^{\mathrm{T}} v\right)}{\sum_{k=1}^{T} \exp\left(u_t^{\mathrm{T}} v\right)} \tag{16}$$

where u_t^{T} represents the hidden state of time step t in the input sequence, usually output by LSTM or Bi LSTM, and is the feature representation of the current time step. V represents a learnable context vector used to capture the weight calculation of important features in a sequence. Exp is an exponential function used to calculate the non-negativity of weights.

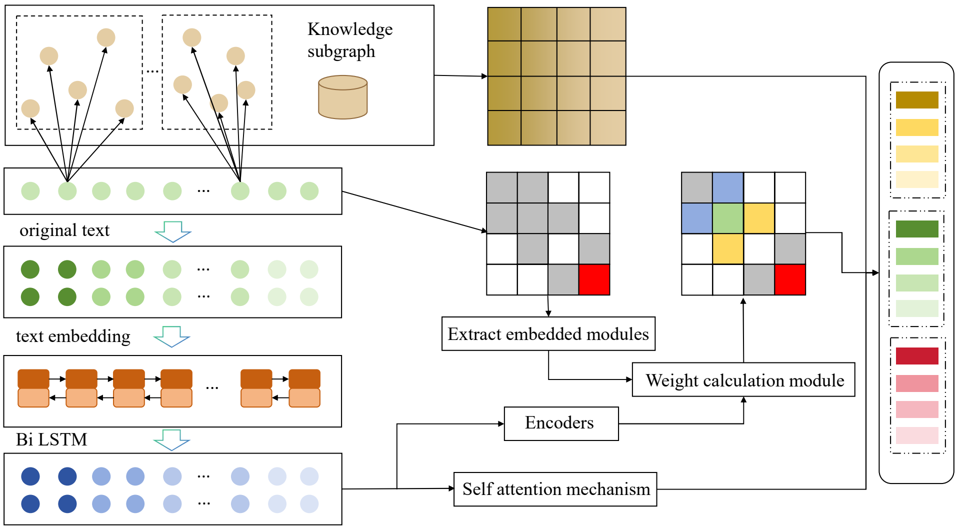
Input the final output or attention weighted result of LSTM into a fully connected layer for generating classification results. Finally, use the Softmax activation function to output the probability of each emotion category.

$$P\left(y \mid x\right) = Softmax\left(W_o h + b_o\right) \tag{17}$$

where W_o is the weight and b_o is the bias term.

The model introduces an external concept knowledge graph and constructs a weighted 'syntax semantics knowledge'. The overall network structure of the model is shown in Figure 3.

Figure 3 Model architecture (see online version for colours)



4 Experiment

4.1 Data set

This article collects short text data related to academic emotions among college students at a certain university, including comments on online learning platforms, reflection logs on learning tasks, and open-ended questionnaire responses. Emotional tags have been annotated by experts and verified by crowdsourcing platforms, with an accuracy rate of 95%.

4.2 Experimental setup

This study mainly relies on hardware devices such as four RTX, 2080/RTX 2080 Ti, and eight Tesla v100. This article implements a technical model using PyTorch and Adam optimiser. Randomly shuffle the training samples before starting each batch of training. In terms of training time, an early stopping strategy was adopted (Yao et al., 2007). Specifically, during each iteration, only a portion of the training data is selected for iteration, and the iteration is stopped after a certain number of iterations. This strategy can effectively prevent overfitting of the model, thereby improving its accuracy.

4.3 Evaluation

SVM is a classic classification model suitable for classifying high-dimensional data. In the experiment, TF-IDF features were used as input, and a linear kernel was selected as the kernel function. RF is an ensemble learning method based on decision trees, which improves classification stability and accuracy through a voting mechanism. Unidirectional long short-term memory networks can model long-term dependencies in sequences, but can only capture unidirectional contextual information. GRU is a simplified version of LSTM with fewer parameters, achieving a balance between computational efficiency and performance (Shewalkar et al., 2019).

The accuracy and F1 score of each model were experimentally measured. Accuracy is defined as the proportion of correctly predicted samples to the total sample size.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{18}$$

F1 score is used for comprehensive evaluation of model performance:

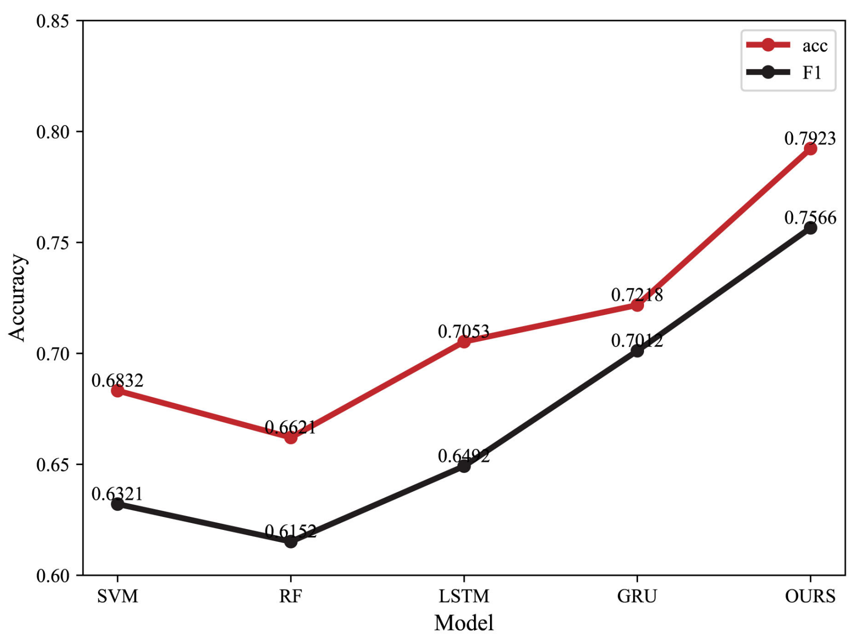
$$F1 = 2 \times \frac{\frac{TP \cdot TP}{(TP + FP)(TP + FN)}}{\frac{TP}{TP + FP} + \frac{TP}{TP + FN}} \tag{19}$$

where *TP* refers to correctly identified positive samples, *TN* to correctly identified negative samples, *FP* to negative samples incorrectly classified as positive, and *FN* to positive samples incorrectly classified as negative.

4.4 Experimental results and analysis

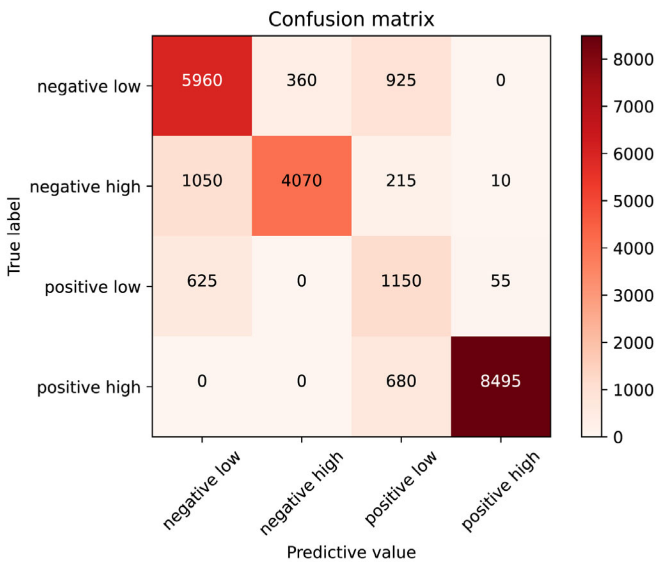
As shown in Figure 4, ours model proposed in this paper has significantly improved accuracy in academic emotion classification, with a specific data of 79.23% and an F1 value of 75.66%. On the student feedback text dataset, the accuracy of recognising academic emotions improved by 7.83% compared to LSTM. Compared to previous studies, there has been a significant improvement.

Figure 4 Model accuracy and F1 value (see online version for colours)



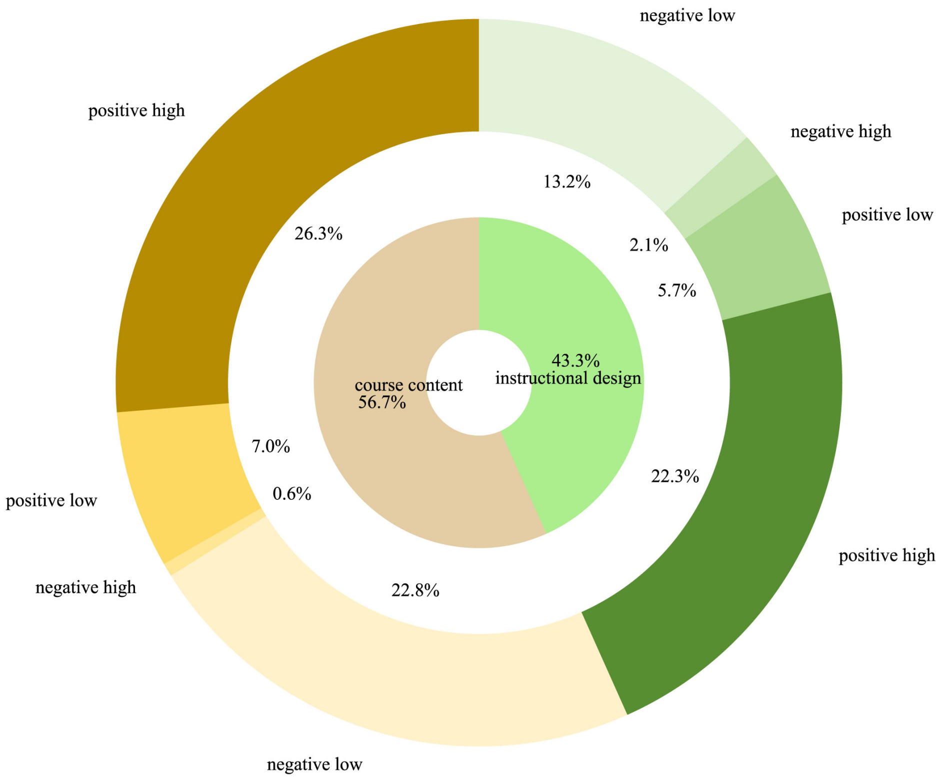
In the confusion matrix of academic emotion recognition shown in Figure 5, labels are classified into negative low, negative high, positive low, and positive high academic emotions. In the experimental dataset, the distribution of emotion categories is mainly concentrated in two dimensions: positive high and negative low, which dominate the validation set. From the confusion matrix, it can be found that the data with a positive high dimension has better prediction results, with a total of 9314 data predicted correctly and an accuracy rate of 93.14%. Meanwhile, the prediction accuracy of the negative low dimension also reached 83.62%. Due to the small proportion of negative high-dimensional data in the dataset, only a small amount of data exists in the confusion matrix in the test set, and the accuracy of classification is relatively low.

Figure 5 Model confusion matrix (see online version for colours)



As shown in Figure 6, in the dimension of course theme, there are two levels of theme course content and instructional design, with course content related data accounting for 56.7% and instructional design related data accounting for 43.3%. In the content of the secondary dimension theme courses, 26.3% of them actively awaken academic emotions; The proportion of positive low arousal academic emotions is 7.0%; Negative high arousal academic emotions account for 0.6%; The proportion of negative low arousal academic emotions is 22.8%. In the secondary dimension theme teaching design, actively awakening academic emotions accounts for 22.3%; The proportion of positive low arousal academic emotions is 5.7%; Negative high arousal academic emotions account for 2.1%. From the above data distribution, it can be concluded that in the course dimension, the proportion of high negative and low arousal dimensions of academic emotions is mainly concentrated in the course content dimension, with a ratio of about 2:1 to the teaching design dimension. Therefore, teachers can focus their teaching optimisation on intervention strategies in the course content dimension.

Figure 6 Distribution of academic emotions (see online version for colours)



5 Conclusions

By introducing attention mechanisms, the model can focus on key emotional information in the text, thereby significantly improving classification performance. The experimental results show that compared with traditional machine learning methods and other deep learning models, our method has significant advantages in accuracy, precision, recall, and F1 score, proving its effectiveness in academic emotion analysis and the accuracy and practical significance of the model's output emotion categories.

The research results not only provide technical support for real-time assessment of college students' emotional states, but also provide new ideas for personalised education intervention and mental health management. By applying the results of emotion classification to optimise teaching strategies and design emotional education platforms, this work has strong practical value.

Although this study has achieved good results, there are still some directions worth further research. This research only used text data as input, and in the future, multimodal data such as speech and facial expression images can be combined to further improve the accuracy and robustness of emotion recognition. The current dataset size and sources may limit the generality of the model. In the future, it can be expanded to more diverse academic scenarios and student groups from different cultural backgrounds, thereby enhancing the adaptability of the model. Academic emotions have the characteristic of dynamic changes. In the future, time series analysis or reinforcement learning methods can be introduced to model the trend of emotional changes and predict the development trajectory of students' emotional states.

References

Abdullah, S.M.S.A., Ameen, S.Y.A., Sadeeq, M.A. et al. (2021) 'Multimodal emotion recognition using deep learning', *Journal of Applied Science and Technology Trends*, Vol. 2, No. 1, pp.73–79.

Anwar, A., Rehman, I.U., Nasralla, M.M. et al. (2023) 'Emotions matter: a systematic review and meta-analysis of the detection and classification of students' emotions in stem during online learning', *Education Sciences*, Vol. 13, No. 9, p.914.

Barron-Estrada, M.L., Zatarain-Cabada, R. and Bustillos, R.O. (2019) 'Emotion recognition for education using sentiment analysis', *Res. Comput. Sci.*, Vol. 148, No. 5, pp.71–80.

Batbaatar, E., Li, M. and Ryu, K.H. (2019) 'Semantic-emotion neural network for emotion recognition from text', *IEEE Access*, Vol. 7, No. 2, pp.111866–111878.

Chakraverthi, B., Ng, S-C., Ezilarasan, M. et al. (2022) 'EEG-based emotion recognition using hybrid CNN and LSTM classification', *Frontiers in Computational Neuroscience*, Vol. 16, No. 9, p.1019776.

Chen, X., Zou, D., Xie, H. et al. (2022) 'Two decades of artificial intelligence in education', *Educational Technology & Society*, Vol. 25, No. 1, pp.28–47.

Gupta, S., Kumar, P. and Tekchandani, R.K. (2023) 'Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models', *Multimedia Tools and Applications*, Vol. 82, No. 8, pp.11365–11394.

Hasan, M.A., Noor, N.F.M., Rahman, S.S.B.A. et al. (2020) 'The transition from intelligent to affective tutoring system: a review and open issues', *IEEE Access*, Vol. 8, No. 3, pp.204612–204638.

Karamimehr, Z., Sepehri, M.M., Sibdari, S. et al. (2023) 'Personalised emotion-aware e-learning systems with interventions', *International Journal of Smart Technology and Learning*, Vol. 3, Nos. 3–4, pp.187–211.

Kim, C. (2012) 'The role of affective and motivational factors in designing personalized learning environments', *Educational Technology Research and Development*, Vol. 60, No. 1, pp.563–584.

Koolagudi, S.G. and Rao, K.S. (2012) 'Emotion recognition from speech: a review', *International Journal of Speech Technology*, Vol. 15, No. 2, pp.99–117.

Sharma, K. and Giannakos, M. (2020) 'Multimodal data capabilities for learning: what can multimodal data tell us about learning?', *British Journal of Educational Technology*, Vol. 51, No. 5, pp.1450–1484.

Shewalkar, A., Nyavanandi, D. and Ludwig, S.A. (2019) 'Performance evaluation of deep neural networks applied to speech recognition: RNN, LSTM and GRU', *Journal of Artificial Intelligence and Soft Computing Research*, Vol. 9, No. 4, pp.235–245.

Wu, C.H., Huang, Y.M. and Hwang, J.P. (2016) 'Review of affective computing in education/learning: trends and challenges', *British Journal of Educational Technology*, Vol. 47, No. 6, pp.1304–1323.

Yadegaridehkordi, E., Noor, N.F.B.M., Ayub, M.N.B. et al. (2019) 'Affective computing in education: A systematic review and future research', *Computers & Education*, Vol. 142, No. 5, p.103649.

Yao, Y., Rosasco, L. and Caponnetto, A. (2007) 'On early stopping in gradient descent learning', *Constructive Approximation*, Vol. 26, No. 2, pp.289–315.

Zhang, K. and Aslan, A.B. (2021) 'AI technologies for education: recent research & future directions', *Computers and Education: Artificial Intelligence*, Vol. 2, No. 1, p.100025.

Zhao, J., Mao, X. and Chen, L. (2019) 'Speech emotion recognition using deep 1D & 2D CNN LSTM networks', *Biomedical Signal Processing and Control*, Vol. 47, No. 1, pp.312–323.

Zhou, J. and Ye, J-M. (2023) 'Sentiment analysis in education research: a review of journal publications', *Interactive Learning Environments*, Vol. 31, No. 3, pp.1252–1264.