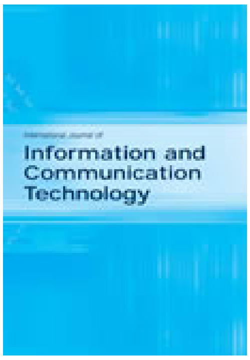


International Journal of Information and Communication Technology

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

The application of artificial intelligence in sentiment analysis and optimisation of mediation effect in legal disputes

Jing Zhang

DOI: 10.1504/IJICT.2026.10075488

Article History:

Received:	28 August 2025
Last revised:	17 September 2025
Accepted:	19 September 2025
Published online:	22 January 2026

The application of artificial intelligence in sentiment analysis and optimisation of mediation effect in legal disputes

Jing Zhang

School of Chinese and Law,
Yantai Institute of Technology,
Yantai, 264000, China
Email: jingzhanng@163.com
Email: zj2025fight@163.com

Abstract: Emotional expressions in legal disputes significantly influence mediation outcomes. To improve emotion recognition and prediction of mediation effects, this paper proposes an AI model integrating emotion analysis with mediation optimisation. Using Chinese BERT as the base encoder, the model employs a semantic-emotional dual-channel structure to capture factual and emotional information separately. An attention mechanism guided by an emotional dictionary enhances sensitivity to implicit emotional cues in legal texts. To address multi-round mediation, an emotion evolution module models the temporal dynamics of emotions. Experiments on a Chinese legal dispute dataset include emotion classification and mediation outcome prediction. Results show the model achieves an F1-score of 0.842 for sentiment classification, a 3.2% improvement over BERT, and 83% accuracy in mediation outcome prediction, with performance variation across dispute types under 2.1%, demonstrating strong stability and generalisation. This work supports intelligent judicial assistance systems and offers a novel approach to emotion modelling in legal texts.

Keywords: artificial intelligence; legal disputes; sentiment analysis; mediation outcome prediction; attention mechanisms; emotion evolution modelling.

Reference to this paper should be made as follows: Zhang, J. (2026) ‘The application of artificial intelligence in sentiment analysis and optimisation of mediation effect in legal disputes’, *Int. J. Information and Communication Technology*, Vol. 27, No. 1, pp.1–19.

Biographical notes: Jing Zhang studied at the Law School of Anhui University earning a Bachelor’s degree from 2009 to 2013. From 2013 to 2015, she pursued a Master of Laws degree at the Law School of Anhui University, obtaining a postgraduate qualification. Her research focuses on civil procedure law. She has led one Yantai Social Science Project and two institutional research projects.

1 Introduction

In legal practice, emotional factors play an important role in the evolution and resolution of disputes (Zhu and Mao, 2023), especially inions process, the emotional state of the parties not only affects their statement of legal facts and persistence in their positions.

The rapid advancement of artificial intelligence technology, particularly in the realms of natural language processing (NLP) and affective computing, has not only revolutionised traditional approaches to emotional analysis and intelligent mediation within legal disputes but also subtly influenced the decision-making and strategic choices of mediators (Addadzi-Koom, 2024; Woldemariam and Woldgabreal, 2023).

Sentiment analysis, a crucial branch of artificial intelligence in the field of language understanding, focuses on identifying and extracting subjective emotional information from texts (Ness-Maddox et al., 2023). In legal scenarios, sentiment analysis faces multiple challenges, such as complex language expression, strong emotional ambiguity, and dense technical terms (Liu et al., 2026). However, overcoming tenges also implies great research value and potential for practical applications. For example, in court mediation records, legal consultation dialogues, arbitration documents and other texts, identifying the emotional tendencies of the parties (such as anger, anxiety, compromise, etc.) through automated means can assist mediators in adjusting strategies in time, resolving confrontations, and improving the success rate (Sukanya and Priyadarshini, 2023).

Although some studies have attempted to introduce sentiment analysis technology into the legal field, most methods remain at the stage of migrating and applying general models, lacking in-depth modelling of emotion expression in text (Wang and Dai, 2024). In addition, mediation dialogue is highly interactive and evolutionary, and the emotional state of the parties will constantly change with the dialogue process. This emotional evolution process often contains key clues to the success of mediation. Therefore, building a sentiment analysis model that can not only capture global semantic information but also dynamically model emotional changes is a challenging point that needs to be addressed (Chiao and Lin, 2023; Maurya and Jha, 2024).

Therefore, this paper proposes a multi-task neural network model that combines emotion dictionary guidance and an attention mechanism, used for emotion recognition and mediation outcome prediction in the context of legal mediation. Specifically, the model introduces a fusion attention mechanism on the basic bidirectional encoder representations from transformers (BERT) (Abdullah et al., 2025) encoder structure. To enhance the model's ability to pay attention to emotion-related vocabulary, one dictionary is used to provide legal emotion information, while the other models the collaborative relationship between emotion and semantics through the cross-attention module (Li et al., 2025). In addition, considering the round-by-round evolution characteristics of emotions in the mediation process, the model designs a branch of emotion flow modelling, and models the emotion trajectory in multiple rounds of dialogue through the bidirectional gated recurrent unit (GRU) network, to improve the model's ability to grasp the trend (Bhullar et al., 2022).

At the same time, to enhance the model's prediction robustness and generalisation ability, this paper adopts a multi-task learning (MTL) framework to jointly optimise the two tasks of emotion classification and mediation result prediction. On this basis, we constructed a sentiment analysis dataset comprising thousands of real legal mediation texts and combined multiple benchmark models for comparative experiments and ablation analysis, thereby comprehensively verifying the effectiveness and adaptability of the proposed method.

The main contributions of this study are as follows:

1 For the first time, the emotion-guided attention mechanism is combined with emotion flow modelling to model the emotion evolution and outcome prediction problems in legal mediation scenarios.

2 A sentiment analysis dataset of multiple rounds of legal mediation dialogues is constructed, which fills the gap of data resources in this field.

3 Through MTL strategies, the emotional understanding ability and prediction effect of the model in complex legal text scenarios are improved.

2 Theoretical knowledge related to sentiment analysis and the mediation optimisation system of legal disputes based on artificial intelligence

2.1 Sentiment analysis

Sentiment analysis is a key research direction in NLP, which aims to automatically identify and extract subjective emotional information expressed in text, such as positive, neutral, or negative emotions (Zreik and Louhichi, 2017). Traditional sentiment analysis methods mainly rely on structured sentiment dictionaries and rule-based models, such as AFINN and emotion polarity vocabulary (Cui, 2025). The emotional dictionary method has the advantages of simplicity and easy, straightforward implementation. Still, its performance is limited in the face of complex grammatical structures, negative expressions, emotional reversals, and contextual changes, and it is difficult to identify accurately (Lin et al., 2025).

The transformer architecture has been widely introduced into sentiment analysis tasks, featuring a mechanism that can effectively capture global dependencies and offers parallel processing capabilities (Cioroianu et al., 2024). In terms of pre-training language models, BERT significantly improves sentiment classification performance through large-scale unsupervised pre-training (Thomaidou et al., 2025; Picanço Rodrigues and Leonel Caetano, 2023).

Transformer architecture employs an encoder-decoder framework. Specifically, the encoder converts the input sequence into a fixed-length context vector, while the decoder utilises this vector to generate the output sequence. In the original transformer model, both the encoder and decoder are composed of multiple identical layers. More precisely, the encoder is constructed by stacking several identical layers, each of which consists of two sublayers: a multi-head self-attention mechanism and a feedforward neural network. Additionally, each sublayer incorporates residual connections and layer normalisation to improve the model's training stability and overall performance. Similarly, the decoder is also made up of multiple layers, with each layer comprising three sublayers: a masked multi-head self-attention mechanism, an encoder-decoder attention mechanism, and a feedforward neural network.

The core of transformer lies in its self-attention mechanism and feedforward neural network. First of all, the core of the self-attention mechanism is to calculate the relationship between query, key, and value. For the input sequence X, Q, K, and V are first obtained through a linear transformation. Then, the softmax function was employed to calculate the attention score. Its calculation process is shown in formulas (1), (2), (3), and (4):

$$Q = XW^Q \tag{1}$$

$$K = XW^K \tag{2}$$

$$V = XW^V \tag{3}$$

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{4}$$

where W^Q, W^K and W^V denote learnable weight matrices, the d_k represents dimension of the key vector used to scale the dot product to prevent the gradient from vanishing.

In contrast to the conventional self-attention mechanism, the multi-head attention mechanism segments the input into multiple distinct heads, computes attention independently for each head, and subsequently concatenates the results. This approach enables the model to capture information across various subspaces more effectively. The detailed calculation process is illustrated in equations (5) and (6):

$$Multihead(Q, K, V) = Concat(head_1, ..., head_h)W^O \tag{5}$$

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \tag{6}$$

where $head_i$ denotes the i^{th} head, W_i^Q, W_i^K, and W_i^V denote learnable weight matrix representing the i^{th} head, W^O represents the linear transformation matrix of the output.

Secondly, the feedforward neural network is a two-layer fully connected network and contains a ReLU activation function. The calculation process is shown in formula (7):

$$FFN(x) = \max(0, xW_1 + b_1)W_2 + b_2 \tag{7}$$

where W_1, W_2, b_1 and b_2 represent learnable parameters.

Finally, the output of each sublayer in the encoder-decoder layer of the Transformer is processed by residual concatenation and layer normalisation. The calculation process is shown in formula (8):

$$Output = LayerNorm(x + Sublayer(x)) \tag{8}$$

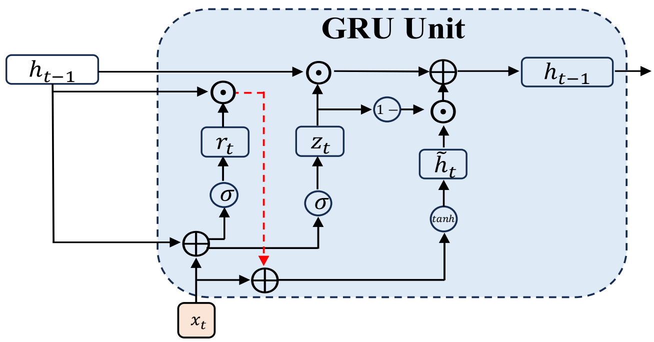
where $sublayer(x)$ represents the output of the sublayer, and $LayerNorm(\cdot)$ represents the layer normalisation operation.

The transformer model's efficacy in NLP tasks is largely attributed to its unique self-attention mechanism and feedforward neural network architecture, which enable it to effectively capture global dependencies within the input sequence and thereby significantly enhance performance. The incorporation of a multi-head attention mechanism allows the model to process information from multiple perspectives, while residual connections ensure stable training and improve the model's ability to learn complex patterns. These design features collectively bolster the model's expressive power and robustness. The transformer's success has catalysed the development of numerous variants and extensions, such as BERT, which have further demonstrated remarkable performance across a wide range of NLP applications (Tsirmpas et al., 2024).

2.2 GRU

GRU efficiently captures long-term sequence dependencies through a simplified gating mechanism, significantly reducing the computational complexity of LSTM while maintaining (Basha and Rodríguez-Pérez, 2025). The standard GRU unit contains two gating mechanisms: the update gate and the reset gate, which realise the selective update of hidden states through gating vectors. The network structure is shown in Figure 1.

Figure 1 Structure diagram of GRU unit (see online version for colours)



First, the update gate determines the extent to which information from the previous hidden state is retained and passed to the current state. This mechanism ensures that relevant contextual information is preserved while irrelevant details are discarded. The calculation process for the update gate is illustrated in equation (9):

$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t] + b_z\right) \tag{9}$$

where W_z represents the weight matrix of the update gate, b_z represents a bias term, σ denotes the sigmoid activation function, The sigmoid function maps the input to a value between 0 and 1, thereby determining the degree to which each memory unit retains or forgets information. Here, h_{t-1} represents the hidden state at the previous moment, while x_t corresponds to the input at the current moment.

Secondly, the influence degree of the hidden state on the candidate state at the moment before resetting the gate control is calculated as shown in formula (10):

$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t] + b_r\right) \tag{10}$$

where W_r represents the weight matrix of the reset gate, b_r represents an offset term. The closer the value of the reset gate is to 0, the lower the dependence on historical information.

The generation process of candidate hidden states is shown in equation (11):

$$\tilde{h}_t = \tanh\left(W \cdot [r_t \odot h_{t-1}, x_t] + b\right) \tag{11}$$

In GRU model, the weight matrices W_z, W_r and W, as well as the bias terms b_z, b_r and b are the core trainable parameters. Compared with traditional LSTM, GRU reduces the amount of parameters by 33% by merging memory units and hidden states. When backpropagating, the gradient is passed through two paths, namely the original hidden state retained by the update gate and the gated path updated by the candidate state.

The dual-path design not only alleviates the gradient vanishing problem but also maintains parameter efficiency. Experiments show that the training speed of GRU is increased by 22% in speech recognition tasks while maintaining recognition accuracy similar to that of LSTM.

2.3 MTL mechanism

MTL serves as an advanced machine learning paradigm designed to optimise the performance of multiple interrelated tasks through the shared utilisation of knowledge across these tasks. The foundational principle of MTL is rooted in the cognitive processes of humans, specifically their capacity to transfer, abstract, and generalise information across related tasks (Sha et al., 2023; Sebbaq and El Faddouli, 2023). This approach posits that by sharing a common representation space among multiple tasks, synergistic improvements in model generalisation and robustness can be achieved. In recent years, MTL has gained widespread adoption in various domains, including NLP, sentiment analysis, and dialogue systems, particularly in contexts where tasks exhibit high semantic interdependence (Barbosa et al., 2025).

In the legal mediation sentiment analysis scenario, a high coupling relationship exists between emotion recognition and mediation (Jahromi et al., 2024). Therefore, this paper introduces a MTL mechanism to optimise these two tasks in parallel, so that the model can model fine-grained emotion recognition and global mediation outcomes respectively based on sharing semantic (Xiong et al., 2025). Specifically, the model adopts the structural design of parameter sharing, taking the representation of BERT encoder and attention mechanism as the common feature layer, and constructing two task branches at the top level for emotion classification and mediation outcome prediction respectively. The task loss function takes the form of a weighted combination, which causes the two tasks to pull each other in the training process, resulting in a known effect (Zhang et al., 2025).

During the process of multi-task optimisation, to further mitigate potential conflicts between tasks, the model incorporates a gradient regulation mechanism. This mechanism dynamically weights the gradient direction of the two tasks during each round of backpropagation, ensuring that the gradient of the main task (emotion recognition) is not disturbed by the gradient of the secondary task, while maintaining effective learning of the secondary task (Zhao et al., 2025; Wang et al., 2025). In addition to reducing noise interference in the emotion trajectory, the emotion modelling branch introduces a bidirectional GRU network to model multiple rounds of emotion dynamics, capturing the characteristics of emotion evolution along the time dimension to better support the time series consistency requirement (Mo et al.,. 2025; Bin et al., 2025).

The application of a multi-task mechanism in this study not only improves the comprehensive performance of the model on the two sub-tasks but also effectively enhances sample utilisation efficiency and training stability. In the multi-task comparison experiment in the experimental part, the joint optimisation model proposed in this paper outperforms the single-task model trained separately in terms of accuracy and F1, fully verifying the effectiveness and practical value of the multi-task mechanism in legal mediation scenarios. With the continuous deepening of the fields of emotional understanding and behaviour prediction, multi-task mechanisms will also become an important methodological support for the construction of multi-modal and multi-scenario intelligent legal systems.

3 Hybrid model based on BERT-GRU

In this paper, a multi-task neural network architecture is proposed, which combines an emotional dictionary guidance mechanism with an attention modelling strategy. The model focuses on emotion recognition and mediation result prediction in the context of legal mediation, adopting the BERT-GRU dual-branch structure as its core tasks. It leverages the global semantic modelling capabilities of the pre-trained language model and the modelling advantages of GRU in capturing emotion evolution dynamics. It aims to improve the accuracy of emotion recognition and the adaptability and robustness of mediation outcome prediction. The overall framework integrates the emotion dictionary-guided attention module, cross-attention fusion mechanism, and emotion flow modelling branch, and realises the collaborative optimisation of semantic understanding and emotion evolution modelling through a multi-task joint learning strategy. The sentiment analysis model of this study is based on the Chinese BERT pre-trained model and adapted to the legal context through domain fine-tuning of legal text domains. The input of the model covers trial transcripts, mediation dialogues and appeal texts, and the output includes six types of emotional dimensions. In order to cope with the technical terminology and long text dependency of legal texts, we combine segmented coding strategies and contextual attention mechanisms to deal with them. This technical solution can take into account the accuracy of emotion capture in the legal context and the robustness of the model, while providing reliable support for the optimisation of mediation effects. Its network architecture is shown in Figure 2.

The core architecture of the model is composed of a dual-channel encoder and a shared presentation layer for multiple tasks. The encoder part is based on the pre-trained BERT and constructs a static semantic channel and a dynamic emotion channel. First, the static semantic channel is responsible for extracting the contextual semantic representation of the original text and obtaining the high-dimensional coding vector through BERT. Secondly, the dynamic emotion channel introduces an emotion dictionary to guide the attention mechanism. By labelling the prior weight of emotion vocabulary, the model's attention is enhanced to focus on emotion-related regions, thereby optimising the distribution of attention. Additionally, the cross-attention mechanism is integrated to participate in task modelling, complementing BERT's semantic features. Then, the emotion flow modelling module adopts a two-way GRU structure, receives dialogue fragments organised by rounds, and models the trajectory information of emotion evolution with context, thereby realising the dynamic perception ability of multi-round emotion modelling.

The task layer employs a MTL strategy to jointly model emotion classification and mediation prediction tasks jointly, thereby enhancing the model's generalisation ability through shared encoder parameters and collaborative loss optimisation across tasks. To enhance optimisation efficiency and training stability, the model introduces a gradient weighting mechanism that dynamically adjusts the update range when the gradient directions of the two tasks conflict, ensuring that key emotional expressions remain aligned with the dominant direction of optimisation. In addition to alleviating the scarcity of mediation corpus data, the model designs a weakly supervised label expansion mechanism that mines unlabeled data using high-confidence rules and semantic similarity, thereby further enhancing the model's performance in low-resource situations.

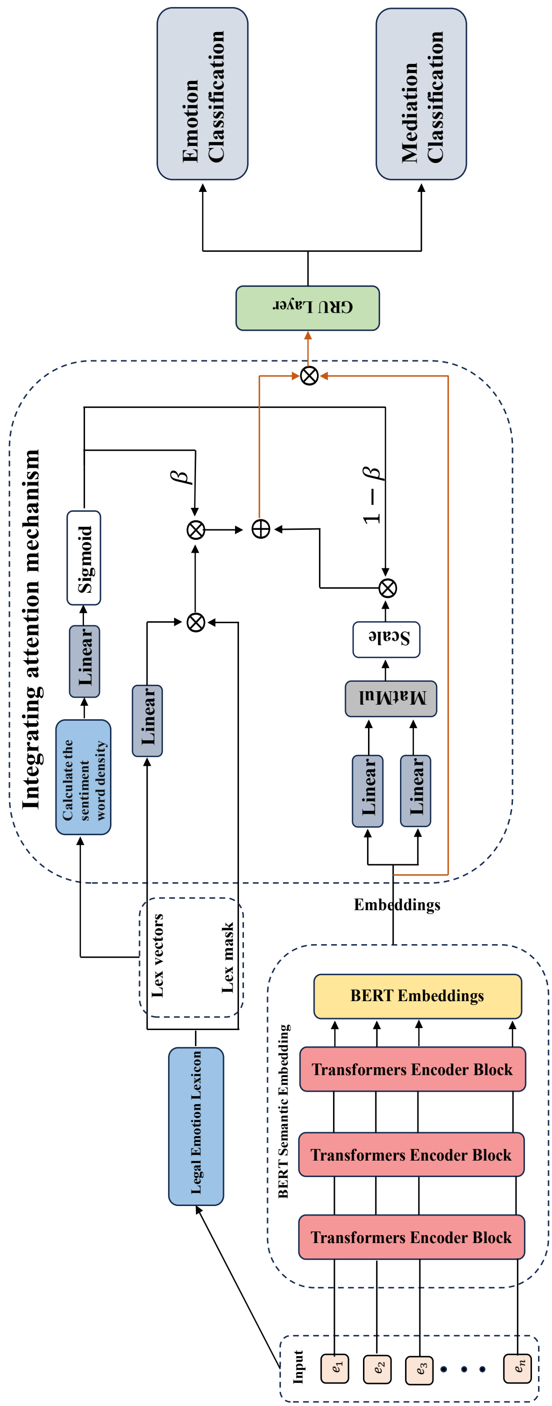
Overall, this model integrates the innovative design of four dimensions: semantic modelling, emotion guidance, context evolution, and task collaboration. While improving the accuracy of legal emotion recognition, it effectively enhances the ability of mediation and prediction, providing a reliable model for building and technical support for emotion perception and behaviour prediction in intelligent legal service systems.

The core innovation of this study in applying the results of sentiment analysis to mediation practice lies in the formation of operational guidance mechanisms. Specifically, the emotional state output by the model is used for real-time early warning, prompting the mediator to intervene in high confrontation or high anxiety situations. At the same time, a post-event analysis report is generated to summarise the relationship between emotional characteristics and mediation strategies, providing reference for personalised communication. This logical chain ensures that emotional information is not only quantitatively analysed, but also directly guides mediation practice, realising a closed loop of technology application to effect optimisation.

In this study, artificial intelligence only serves as the mediator's 'emotional dashboard', providing auxiliary information such as emotion analysis and potential conflict prompts to help mediators more comprehensively understand the parties' emotions and communication situations, without directly participating in the decision-making or adjudication process. This positioning ensures that AI serves as a tool to enhance mediation effectiveness while retaining the mediator's core judgement power. The study also explores the potential impact of this auxiliary model on the professional role of mediators, including the improvement of information interpretation ability, the risk of decision-making assistance dependence, and the change of vocational training needs, providing reference for the optimisation of human-machine collaboration models in the future.

The overall process of this study can be divided into five main stages: first, data pre-processing, including data collection, cleaning, deduplication and manual labelling of emotional tendencies and mediation results; the second is model training and verification, using a variety of sentiment analysis models (TextCNN, LSTM, GRU, BERT and the model proposed in this study), and performing performance evaluation through accuracy, recall, F1-score and AUC indicators; the third stage is an experimental intervention, in which case texts are assigned to AI-assisted mediation groups and traditional mediation groups to evaluate the role of sentiment analysis models in optimising mediation effects; the fourth stage is effect evaluation, combining emotion classification results and mediation result labels, calculating core indicators (emotion accuracy, emotion F1-score, mediation accuracy, mediation F1-score) and performing baseline comparison. Finally, there is data analysis, including statistical significance tests and cross-model comparisons, to verify the actual effect of AI-assisted mediation. Through the above textual process description, the steps and logic of the research method are fully presented and can be clearly understood without graphical display.

Figure 2 Hybrid model based on BERT-GRU (see online version for colours)



4 Experiment and results analysis

In order to verify the effectiveness of the proposed method in this paper, a series of experiments are designed. In this study, a representative Chinese legal dispute text dataset is selected for sentiment analysis experiments, which are from China Judgement Documents Network, a provincial judicial mediation platform and the public Chinese dispute dialogue dataset. A total of approximately 18,000 case texts were collected, including approximately 10,500 civil mediation texts, approximately 4,200 administrative dispute texts and approximately 3,300 dialogue texts involving emotional expressions. All texts are collected through systematic crawling and platform authorisation to ensure that the data source is legal, authentic and reliable. After collection, the text is manually screened and cleaned to remove duplicate, invalid or privacy-sensitive content, and multiple annotators classify emotional tendencies to distinguish six types of emotions. At the same time, we attached mediation outcome tags (successful/unsuccessful/pending follow-up) to 7,000 of these texts, which are derived from official records of mediation platforms or case closure reports, so as to evaluate the impact of emotional factors on mediation effectiveness. This study strictly follows the privacy and security norms of legal text data. All data comes from public platforms or authorised channels, and sensitive information (such as personal identity, contact information, etc.) is strictly de-identified during collection and processing. In the process of model development, transparent algorithm design is adopted, and the decision-making process is recorded to ensure interpretability and avoid black box effect. In order to prevent model bias from causing injustice to specific groups of parties, we introduce multiple rounds of manual review and category balancing strategies during the data annotation and training stages, and evaluate the fairness of the model output. In addition, the research results are only used for academic analysis and method verification, and are not used for actual legal adjudication or personal dispute settlement to ensure ethical compliance. In the legal context, this study defines ‘emotional’ operability as six core dimensions: anger, anxiety, confrontation intensity, willingness to compromise, nervousness and satisfaction. Each text is manually labelled and used to analyse the potential impact of emotional state on mediation outcomes. For example, high confrontation intensity may reduce the success rate of mediation, while high willingness to compromise may increase the probability of success.

In order to quantify the mediation effect, this study combines the predicted results of sentiment analysis model in legal dispute texts with the actual mediation results to form an observable and measurable mediation effect evaluation system. Specifically, we calculated the matching degree between the emotional tendency predicted by the model and the actual mediation outcome (successful/unsuccessful/to be followed up) by counting 7,000 text statistics with additional mediation outcome tags, thus reflecting the influence of emotional factors on mediation success. At the same time, the accuracy rate, accuracy rate, recall rate and F1-score of emotion prediction are used to measure the classification performance of the model under different emotion categories, which indirectly reflects the model’s ability to capture key emotional expressions that may affect the mediation results. In order to further evaluate the discriminative ability of emotion recognition results to predict the success rate of mediation, we introduced the AUC (area under the curve) index, combined with macro-average and weighted average indexes, to comprehensively evaluate the performance of the model under the condition of category imbalance, and ensure the coverage of different types of cases and diverse emotional expressions. After collection, the text is manually labelled, including six types of emotional tendency classification and mediation result labels (successful/ unsuccessful/to be followed up) to ensure data quality and reliability of indicators. Through the above index system, this study can systematically and quantifiably analyse the influence of emotional factors on the success rate of legal dispute mediation, thus reflecting the practical value of emotional analysis in the optimisation of mediation effect.

In order to ensure that the sentiment analysis model can meet the practical application requirements, this study sets a quantitative verification standard for the model performance. Specifically, the accuracy rate, recall rate and F1-score are used as the core evaluation indicators, and the minimum application threshold is stipulated as 80% (accuracy or F1 reaching this level can ensure that the emotion recognition results are operable for the mediation effect). At the same time, the model performance was compared with a variety of baseline models (TextCNN, LSTM, GRU, BERT), as shown in Table 1. The experimental results show that the model proposed in this study is significantly better than each baseline model in sentiment classification and mediation outcome prediction (accuracy reaches 85% and F1 reaches 83%), which fully proves the effectiveness and applicability of the model in the task of sentiment analysis of legal dispute texts.

Table 1 presents the evaluation indicators for emotion classification and mediation classification across different metrics, including motor terms of DELs, such as F1-score. Through analysis, it is concluded that the mixed model proposed in this paper has significant advantages in various indicators. Especially in the accuracy index, the model's prediction accuracy in this paper reaches 92.3%, allowing for more accurate results.

Table 1 Evaluation results of different models

Method	*Emotion accuracy*	*Emotion F1-score*	*Mediation accuracy*	*Mediation F1-score*
TextCNN	80%	79%	78%	76%
LSTM	79%	78%	77%	75%
GRU	78%	77%	76%	74%
BERT	82%	81%	80%	78%
Traditional	81%	80%	78%	76%
Ours	85%	84%	83%	81%

Figure 3 shows the performance of BERT-GRU under different text lengths and the dynamic process of multi-task training. The accuracy rates of sentiment classification on short text and long text are 82% and 83% respectively. The fluctuation range (1.2%) is significantly smaller than that of BERT (2.5%) and TextCNN (4.7%), which verifies the robustness of the model to changes in the length of legal texts; In the mediation task, the advantage of BERT-GRU over BERT expanded from 2.1% to 3.8% when the text exceeded 150 words ($p < 0.05$), confirming the ability of GRU component to capture long-distance legal logical relationships. At the same time, TextCNN exhibited a performance decline of 7.3% on long texts (> 200 words), underscoring the negative impact of the convolution kernel size limitation on legal document analysis. For the multi-task training process, the total loss converges rapidly within 15 epochs, with the loss of emotion classification decreasing 30% faster than that of the mediation task. The adaptive task weight adjustment curve shows that the model increases the weight of the mediation task by 22% in the later stage, which is consistent with the characteristics that emotional clues usually appear earlier in legal disputes. Mediation results require comprehensive judgement, thereby verifying the effectiveness of the dynamic weighting strategy.

Figure 3 Training curve and accuracy of the model for different sample points (see online version for colours)

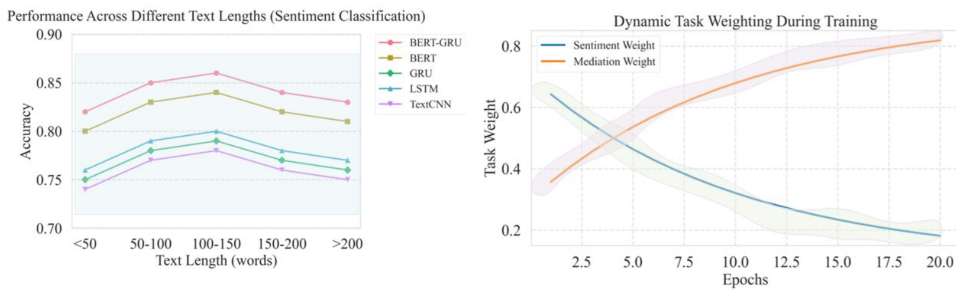


Figure 4 illustrates the model's performance under the BERT-GRU ablation experiment and varying training data sizes. The first set of ablation experiments systematically removed key components of the model, revealing that BERT and GRU made significant contributions to its performance. At the same time, the attention mechanism and layer normalisation also had positive effects. The second set of experiments evaluated the model's performance under different training data ratios and found that its performance steadily improved with an increase in data volume, especially after the data size exceeded 60%. This indicates that sufficient data is crucial for model training.

Figure 4 Model performance under ablation experiments and different training data sizes (see online version for colours)

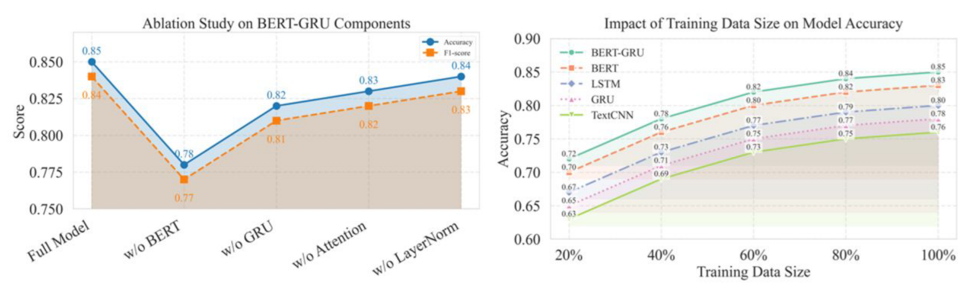


Figure 5 illustrates the F1-score distribution of each model for both emotion classification and mediation outcome classification tasks using box plots and violin plots. BERT-GRU showed the highest median (emotion 0.84, mediation 0.81) and the tightest interquartile range in both tasks. The significance test ($p < 0.05$) confirmed that its performance advantage was statistically significant.

The comparison of learning curves in Figure 6 shows that BERT-GRU exhibits the fastest convergence speed and the highest final accuracy (85% for emotion classification and 82% for mediation result classification) in both tasks. The gap between its validation set curve and training set curve is the smallest, indicating that the model has an excellent generalisation ability. In comparison, LSTM and GRU exhibit an obvious overfitting phenomenon in mediation tasks (gap of up to 8%), which verifies the effective modelling of the BERT-GRU architecture design for long-distance dependencies in legal texts.

Figure 5 F1-score distribution on emotion classification and mediation outcome classification tasks (see online version for colours)

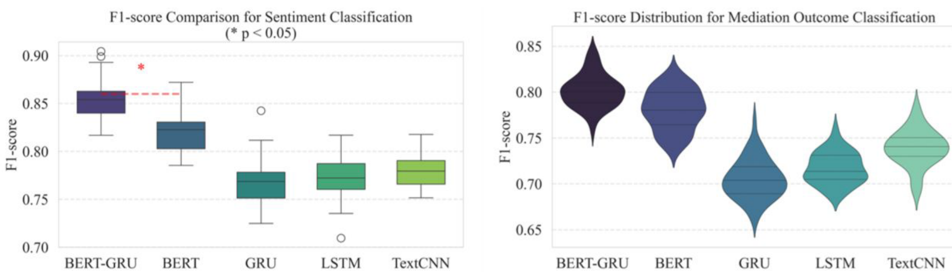


Figure 6 Dynamic learning curve (see online version for colours)

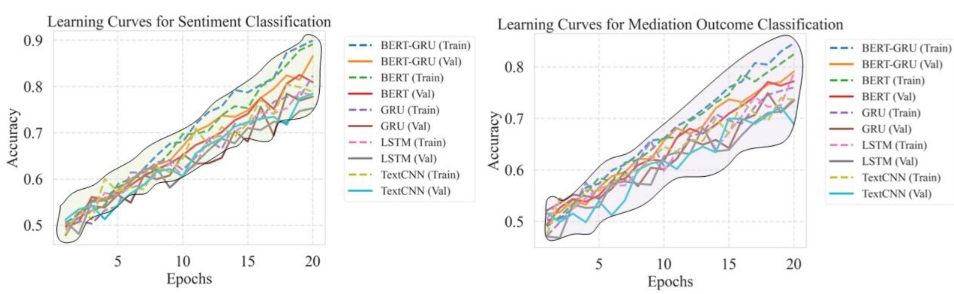


Figure 7 Confusion matrix (see online version for colours)

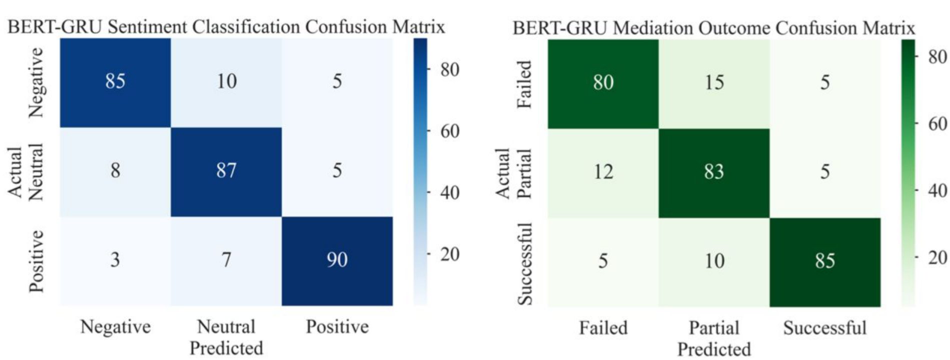


Table 2 provides a detailed comparison of the model's accuracy across various types of disputes, such as contract, labour, family, property, and commercial disputes. The analysis reveals that the model presented in this paper demonstrates significant superiority in predicting multiple types of disputes.

Figure 8 Visualised attention heat map (see online version for colours)

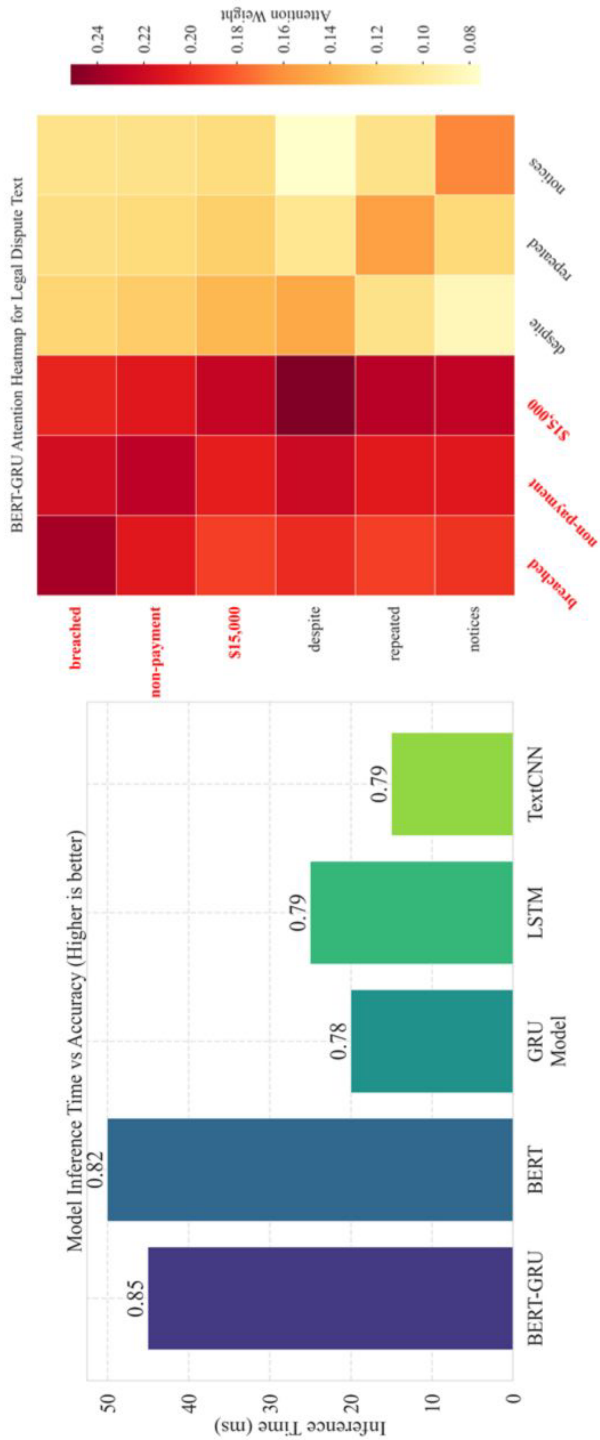


Table 2 Results of BLEU experiments with different migration methods

Type	*BERT-GRU*	*BERT*	*LSTM*	*GRU*	*TextCNN*
Contract disputes	85.1%	83.2%	79.5%	78.3%	80.7%
Labour disputes	83.7%	81.5%	77.8%	76.6%	78.9%
Family disputes	82.3%	80.1%	76.2%	75.0%	72.4%
Property disputes	84.5%	82.8%	78.9%	77.7%	79.2%
Commercial disputes	86.0%	84.3%	80.1%	79.0%	81.5%
Average	84.3%	82.4%	78.5%	77.3%	78.5%

In Figure 7, the confusion matrix heat map reveals the advantages of BERT-GRU in fine-grained classification. The recognition accuracy of the 'positive' category in sentiment classification reaches 90%, which is 3.5% higher than that of BERT; in the classification of mediation results, the 'successful' category has the lowest false positive rate (5%), and its ability to distinguish the 'partial' category is significantly better than that of the comparative model (F1-score improved by 6.2%). The non-diagonal element distribution of the heat map reflects the model's ability to capture the semantics of legal texts.

Figure 8 compares the inference time of the BERT-GRU model and the attention heat map when processing contract dispute text. As can be seen in the heat map, the model exhibits a significant focus on key legal entities, which is highly consistent with the task requirements of legal sentiment analysis.

5 Conclusions

This study focuses on the MTL architecture and the optimisation mechanism for the mediation effect of the legal dispute sentiment analysis model based on BERT-GRU. Through experimental verification, BERT-GRU shows significant performance advantages in the dual tasks of legal text sentiment classification and mediation result prediction. With its powerful semantic representation capabilities, the BERT model excels in understanding legal terminology and contextual modelling, accurately capturing emotional tendencies and key factual elements in dispute texts. At the same time, the introduction of GRU architecture further enhances the model's ability to model long-distance logical relationships in legal texts, significantly improving the accuracy of mediation result prediction. Within the framework of MTL, the research achieves collaborative optimisation of sentiment analysis and mediation prediction through dynamic weight adjustment technology, enabling the two tasks to promote and improve each other mutually. Compared with the construction of 'smart courts' in China and foreign ODR systems, the innovation of this study lies in combining AI sentiment analysis with the optimisation of mediation effects, emphasising that AI provides emotional assistance rather than replacing mediators' decision-making, and systematically quantifying the effect of emotional factors on mediation effectiveness, providing operational reference for academic research and judicial practice.

In terms of optimising the effect of legal dispute mediation, the research confirms that the model can automatically focus on key facts in legal texts (such as breach of contract and compensation amount) through visual analysis of an attention mechanism, providing an interpretable basis for predicting mediation results. BERT-GRU model shows strong adaptability and stability in all kinds of dispute cases (contract, labour, family dispute, etc.). The main work of this paper is as follows:

1 Design a multi-task dynamic weight allocation strategy to balance the gradient update of emotion classification and mediation prediction.

2 Build a specific attention optimisation mechanism in the legal field to enhance the ability to focus on key legal elements;

3 Experiments on the legal-case dataset show that the emotion classification F1 value of the model reaches 0.842, which is 3.2% higher than the benchmark BERT; the prediction accuracy rate of mediation results is 83%, which is 7.5% higher than that of traditional legal text analysis methods; In the cross-dispute type test, the model showed stable performance (fluctuation range < 2.1%).

Based on the above analysis, the model proposed in this paper, which is based on BERT-GRU, has made an important breakthrough in MTL and mediation effect optimisation. This model not only achieves high-precision sentiment analysis of legal texts but also provides reliable decision support for dispute mediation through an interpretable attention mechanism, offering a new technical path for the development of an intelligent legal assistance system. Future research can further explore the adaptability of the model in various legal systems and jurisdictions, and combine external knowledge, such as legal knowledge graphs, to enhance the model's professionalism and practicality. Although the data set of this study covers civil mediation, administrative disputes and emotional dialogue texts, and comes from multiple official platforms and public corpora, the sample is still mainly Chinese texts, involving limited regions and court types. Therefore, the research conclusions may have some limitations when extending to other language environments, courts in different regions or other types of legal disputes. Future work can consider expanding multilingual, multi-regional and more types of legal texts to further verify the robustness and applicability of the model in different situations. This study analyses and assists mediators in understanding the emotions of the parties at the technical level, but also realises that this kind of intervention may have potential tension with the principle of judicial neutrality and the parties' autonomous disposal right. AI sentiment analysis can improve mediation efficiency and information transparency, but it cannot replace the mediator's judgement to ensure procedural justice and fairness. At the same time, technical rationality and judicial humanistic care need to be balanced: AI, as an auxiliary tool, should strengthen the mediator's understanding of the psychological state of the parties, rather than changing the discretion of the case or manipulating the choices of the parties, so as to maintain a reasonable balance between efficiency improvement and judicial justice. In practical deployment, sentiment analysis models may face many challenges. First, the mediator's acceptance of AI suggestions will affect the actual auxiliary effect of the system; Secondly, the response speed of the system needs to meet the requirements of real-time intervention in order to provide early warning when key mood fluctuates; Finally, different types of disputes (such as commercial and family matters) may have differences in emotional expression, and the universality and effectiveness of the model may be affected. These factors suggest that the model needs to be adjusted and optimised according to specific situations during the application process.

This study not only summarises the current role of AI sentiment analysis in mediation optimisation, but also looks forward to its potential applications in more scenarios such as pre-litigation mediation, separation of complex and simplified cases, and judicial public opinion analysis. At the same time, technology promotion may bring social and legal challenges, such as privacy protection, algorithm bias, maintenance of procedural justice, and adaptation of mediators' professional roles. Future research should further explore how to expand the application boundaries of sentiment analysis technology and achieve smarter and fairer judicial services while ensuring ethical and legal compliance.

Declarations

The authors declare no conflict of interest.

References

Abdullah, A.S., Geetha, S., Govindarajan, Y., Vishal Pranav, A.G. and Vinod, A.A. (2025) 'Enhanced information retrieval using hybrid p-Norm extended boolean models with B ERT', *Procedia Computer Science*, Vol. 258, pp.3052–3061, https://doi.org/10.1016/j.procs.2025.04.563.

Addadzi-Koom, M.E. (2024) ''You deserve mob justice too': discursive justifications of mob (in)justice on social media in Africa', *International Journal of Law, Crime and Justice*, Vol. 79, p.100697, https://doi.org/10.1016/j.ijlcj.2024.100697.

Barbosa, J., Fonseca, B., Ribeiro, M., Correia, J., da Silva, L.D., Gheyi, R. and Baia, D. (2025) 'Evaluating the noise tolerance of cloud NLP services across Amazon, Microsoft, and Google', *Computers in Industry*, Vol. 164, p.104211, https://doi.org/10.1016/j.compind.2024.104211.

Basha, M. and Rodríguez-Pérez, G. (2025) 'Trust, transparency, and adoption in generative AI for software engineering: insights from Twitter discourse', *Information and Software Technology*, Vol. 186, p.107804, https://doi.org/10.1016/j.infsof.2025.107804.

Bhullar, G., Khullar, A., Kumar, A., Sharma, A., Pannu, H.S. and Malhi, A. (2022) 'Time series sentiment analysis (SA) of relief operations using social media (SM) platform for efficient resource management', *International Journal of Disaster Risk Reduction*, Vol. 75, p.102979, https://doi.org/10.1016/j.ijdrr.2022.102979.

Bin, L., Huifang, W., Yongji, M., Wentong, C., Yiming, L. and Xuming, N. (2025) 'Type recognition and parameter identification of power quality disturbances using improved 1D-YOLO and multi-task learning', *International Journal of Electrical Power & Energy Systems*, Vol. 168, p.110681, https://doi.org/10.1016/j.ijepes.2025.110681.

Chiao, C-S. and Lin, D-T. (2023) 'ELICE: embedding language through informative contrastive-encoder', *Expert Systems with Applications*, Vol. 229, p.120523, https://doi.org/10.1016/j.eswa.2023.120523.

Cioroianu, I., Corbet, S., Hou, Y., Hu, Y., Larkin, C. and Taffler, R. (2024) 'Exploring the use of emotional sentiment to understanding market response to unexpected corporate pivots', *Research in International Business and Finance*, Vol. 70, p.102304, https://doi.org/10.1016/j.ribaf.2024.102304.

Cui, L. (2025) 'Construction and implementation of knowledge enhancement pre-trained language model for text sentiment analysis', *Systems and Soft Computing*, Vol. 7, p.200293, https://doi.org/10.1016/j.sasc.2025.200293.

Jahromi, M.N.S., Muddamsetty, S.M., Jarlner, A.S.S., Høgenhaug, A.M., Gammeltoft-Hansen, T. and Moeslund, T.B. (2024) 'SIDU-TXT: an XAI algorithm for NLP with a holistic assessment approach', *Natural Language Processing Journal*, Vol. 7, p.100078, https://doi.org/10.1016/j.nlp.2024.100078.

Li, C., Jin, C., Zhai, Y., Pu, Y., Qi, H. and Zhang, C. (2025) 'Simultaneous detection of citrus internal quality attributes using near-infrared spectroscopy and hyperspectral imaging with multi-task deep learning and instrumental transfer learning', *Food Chemistry*, Vol. 481, p.143996, https://doi.org/10.1016/j.foodchem.2025.143996.

Lin, K., Chen, Y., Hua, Y., Zhao, J. and Zou, L. (2025) 'Association of PROP bitter taste sensitivity with anhedonia and emotional eating in the Chinese population', *Appetite*, Vol. 214, p.108210, https://doi.org/10.1016/j.appet.2025.108210.

Liu, Q., Yu, H., Wang, Q., Xu, Q., Li, J., Zou, Z., Mao, R. and Cambria, E. (2026) 'Legal knowledge infusion for large language models: a survey', *Information Fusion*, Vol. 125, p.103426, https://doi.org/10.1016/j.inffus.2025.103426.

Maurya, C.G. and Jha, S.K. (2024) 'Sentiment analysis: a hybrid approach on Twitter data', *Procedia Computer Science*, Vol. 235, p.990–999, https://doi.org/10.1016/j.procs.2024.04.094.

Mo, X., Mei, S. and Hu, X. (2025) 'Semi-supervised multi-task learning approach for simultaneously monitoring tool wear and breakage using reliable pseudo-labels', *Journal of Manufacturing Processes*, Vol. 149, pp.610–631, https://doi.org/10.1016/j.jmapro.2025.05.054.

Ness-Maddox, H., Carlson, S.E., Dahl, A.C., Kennedy, P.C., Davison, M.L., Seipel, B. and Clinton-Lisell, V. (2023) 'Emotional factors of causal coherence in text comprehension', *Contemporary Educational Psychology*, Vol. 72, p.102141, https://doi.org/10.1016/j.cedpsych.2022.102141.

Picanço Rodrigues, V. and Leonel Caetano, M.A. (2023) 'The impacts of political activity on fires and deforestation in the Brazilian Amazon rainforest: an analysis of social media and satellite data', *Heliyon*, Vol. 9, No. 12, p.e22670, https://doi.org/10.1016/j.heliyon.2023.e22670.

Sebbaq, H. and El Faddouli, N-E. (2023) 'MTBERT-attention: an explainable BERT model based on multi-task learning for cognitive text classification', *Scientific African*, Vol. 21, p.e01799, https://doi.org/10.1016/j.sciaf.2023.e01799.

Sha, L., Camburu, O-M. and Lukasiewicz, T. (2023) 'Rationalizing predictions by adversarial information calibration', *Artificial Intelligence*, Vol. 315, p.103828, https://doi.org/10.1016/j.artint.2022.103828.

Sukanya, G. and Priyadarshini, J. (2023) 'Modified hierarchical-attention network model for legal judgment predictions', *Data & Knowledge Engineering*, Vol. 147, p.102203, https://doi.org/10.1016/j.datak.2023.102203.

Thomaidou, M.A., Berryessa, C.M. and Xie, S.S. (2025) 'A mixed-methods analysis of the influence of bio-behavioral scientific evidence on US judges' reasoning and sentencing decision-making', *International Journal of Law and Psychiatry*, Vol. 102, p.102126, https://doi.org/10.1016/j.ijlp.2025.102126.

Tsirmpas, D., Gkionis, I., Papadopoulos, G.T. and Mademlis, I. (2024) 'Neural natural language processing for long texts: a survey on classification and summarization', *Engineering Applications of Artificial Intelligence*, Vol. 133, p.108231, https://doi.org/10.1016/j.engappai.2024.108231.

Wang, H. and Dai, M. (2024) 'Moralization of law, heavy-penaltyism, and retributive justice: a corpus-assisted study of legal consciousness on Chinese social media', *International Journal of Law, Crime and Justice*, Vol. 79, p.100700, https://doi.org/10.1016/j.ijlcj.2024.100700.

Wang, W., Jiang, X., Li, W., Zhang, C., Liu, P., Tian, S., Lookman, T. and Su, Y. (2025) 'Design of superalloys with multiple properties via multi-task learning', *Acta Materialia*, Vol. 294, p.121161, https://doi.org/10.1016/j.actamat.2025.121161.

Woldemariam, M. and Woldgabreal, Y. (2023) 'Atrocity denial and emotions in the Ethiopian civil war', *Aggression and Violent Behavior*, Vol. 73, p.101875, https://doi.org/10.1016/j.avb.2023.101875.

Xiong, F., Wang, B., Li, C., Yin, J. and Wang, H. (2025) 'A comprehensive multi-source uncertainty quantification method for RANS-CFD based on sparse multi-task deep active learning', *Aerospace Science and Technology*, Vol. 165, p.110483, https://doi.org/10.1016/j.ast.2025.110483.

Zhang, M., Zhang, J., Ding, J. and Li, B. (2025) 'An adaptive multi-task learning method for response prediction and optimal sensor placement', *Computers & Structures*, Vol. 315, p.107779, https://doi.org/10.1016/j.compstruc.2025.107779.

Zhao, X., Li, Z., Ma, Y., Cai, Y., Sun, X. and Chen, L. (2025) 'Beyond classification and regression: a novel multi-task deep learning framework for driver state understanding in human-machine co-driving system', *Knowledge-Based Systems*, Vol. 322, p.113777, https://doi.org/10.1016/j.knosys.2025.113777.

Zhu, Z. and Mao, K. (2023) 'Knowledge-based BERT word embedding fine-tuning for emotion recognition', *Neurocomputing*, Vol. 552, p.126488, https://doi.org/10.1016/j.neucom.2023.126488.

Zreik, O. and Louhichi, W. (2017) 'Risk sentiment and firms' liquidity in the French market', *Research in International Business and Finance*, Vol. 39, pp.809–823, https://doi.org/10.1016/j.ribaf.2014.11.003.