

# Short Text Semantic Sentiment Analysis Based on Dual Channel Aspect Attention in Intelligent Systems

Yan Li, College of Computer and Control Engineering, Northeast Forestry University, China\*

## ABSTRACT

Traditional deep learning models for text sentiment analysis fail to fully harness the contextual semantic information of aspect nodes or use prior sentiment resources. This paper proposes a dual channel sentiment analysis model named M2BERT-BLSTM AA that is based on an enhanced Bidirectional Encoder Representations from Transformers(BERT)and Bidirectional Long short-term memory(BLSTM) model and incorporates a Dual Attention Mechanism. Firstly, an emotional resource database is constructed using existing emotional resources. Secondly, vectors are concatenated following mean and max pooling along the dimension of sentence length. These semantic features mitigate evaluation imbalance. Then the text and sentiment information are encoded separately, using distinct Attention Mechanism(Att-M) to extract contextual relationships and emotional features. The model's Aspect-Based multi-class sentiment prediction accuracies on the three Chinese datasets of Meituan ordering, restaurants, and laptops are 75.2%, 87.5%, and 75%, respectively, showing improved performance on sentiment classification.

## KEYWORDS

aspect attention, dual channel, M2BERT-BLSTM, sentiment classification, short text, social network

## INTRODUCTION

With the rapid development of social platforms and e-commerce websites, people are now free to express their opinions and emotions on topics of interest. This sentiment-driven information is widely disseminated in various forms, including text, images, videos, and audio. Analyzing these sentiment-imbued opinions can facilitate a better understanding of user behavior, discern users' product preferences, identify user needs, customize marketing strategies, enhance public sentiment analysis, and detect fake news (Sahoo & Gupta, 2021; Sarkissian & Tekli, 2021; Zhang et al., 2021). Accurately analyzing this vast amount of sentiment information has become a research hotspot in the field of natural language processing (NLP).

Text Sentiment Analysis, also referred to as Opinion Mining, is a process aided by computers to swiftly capture and organize the subjective evaluation information on the internet. It adeptly mines and assesses individuals' viewpoints, emotions, reviews, stances, and the sentimental

DOI: 10.4018/IJSWIS.336480

\*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

inclinations of texts about entities like products, services, organizations, events, and topics, followed by inductive reasoning and inference of this mined data (Do et al., 2022; Hajjar & Tekli, 2022; Wang & Yang, 2021).

Currently, researchers both domestically and internationally have shifted their focus from traditional research based on dictionaries and machine learning to sentiment classification methods grounded in deep learning (Barbosa et al., 2022; Daou et al., 2021). For instance, Abdi et al. (2019) acknowledged the strengths of CNN (convolutional neural network) and RNN, proposing an LSTM (long short-term memory) model that elevates the accuracy of sentiment classification in reviews by over 5% through multi-feature fusion. Bin et al. (2022) embarked on sentiment tendency analysis on Weibo using Baidu's ERNIE and a dual attention mechanism (Att-M). Bin et al. performs dynamic **Dynamic** feature representation of text through ERNIE semantic representation capability, which combines **combining** sentiment resources and attention mechanism to solve the problem of different meanings for the same word in traditional word vectors, but their approach lacks the use of linguistic features. Xingjie and Yunze (2020) introduced a method based on convolutional neural networks and attention models for text sentiment analysis and personalized recommendations. The dimensions in the user preference vector indicate the user's preference for the corresponding dimension of the product, but there is no effective fusion of ratings and text. Dandan et al. (2021) utilized the BERT model to pre-train balanced short texts, representing language models with feature vectors, subsequently executing Chinese short text categorization; but it does not incorporate location information such as emoticons and symbols.

These deep learning-based methodologies outperform those built on manually crafted features, yet their capability in capturing deeper semantic information remains somewhat deficient (Barr et al., 2022; Tekli et al., 2021). Moreover, there is limited research both domestically and internationally on aspect-based sentiment classification of Chinese short texts. The majority of these studies directly apply attention mechanisms, overlooking the importance of syntactic relationships. This results in an inability to fully utilize the contextual semantic information of aspect nodes and represent polysemy within contexts (Ismail et al., 2022; Sarivougioukas & Vagelatos, 2022). Concurrently, the untapped potential of prior sentiment resources within neural networks leads to subpar sentiment classification outcomes.

In most of the current sentiment analysis studies, especially for Chinese short texts, there is a prevalent reliance on direct attention mechanisms. This often ignores the importance of syntactic relationships, leading to several challenges. Such models might fail to fully harness the contextual semantic information of aspect nodes, represent word polysemy, or adequately use prior sentiment resources within the neural networks. As a result, the sentiment classification performance is suboptimal. To this end, this paper proposes a sentiment analysis model, named M2BERT-BLSTM AA, that is based on improved BERT and BLSTM models combined with a dual channel Attention Mechanism. Firstly, an emotional resource database containing sentiment words, negation words, and adverbs of degree is constructed using existing emotional resources. It can be utilized in the channel of emotion feature extraction in subsequent processes. Secondly, the BERT model is improved by converting the sequence of its hidden layer to vectors, which are then concatenated following mean and max pooling along the dimension of sentence length.

The model focuses on locally salient features while preserving global features. Then text semantic features and aspect sentiment features are extracted from different attention mechanism channels, and feature vectors concatenation is followed by classification. As a result, the interaction between textual semantic information and aspect-based sentiment information is enhanced, which helps to improve the effectiveness of classification. Introducing an attention mechanism, the model prioritizes key words in sentences and aspect-based sentiments, training and classifying the acquired deep word embedding. This achieves a more precise fine-grained classification of sentiments across different facets. The effectiveness of the sentiment resources and the application of this model have been validated through experiments.

The proposed model has important industrial significance, which can be integrated with various intelligent systems (Poria et al., 2023). In industrial and commercial applications, it is far-reaching in customer feedback, product recommendation, customer service evaluation, market research, market trend prediction, user experience improvement, public opinion monitoring, brand management, product quality control, customized marketing strategy, decision support, and personalized recommendation (Jiang et al., 2023). It also plays an important role in the financial investment analysis, analyzing patient cases in medical systems and so on.

## RELATED WORKS

### Deep Learning and BLSTM

The core of deep learning, **as suggested by Hinton and his colleagues**, is feature learning. It mainly acquires hierarchical feature information through layered networks, abstractly representing raw data at a higher level, which significantly aids in the classification of these abstract representations; relevantly, Zhang et al. (2023) proposes a dual-level architecture to recognize emotion. In recent years, deep learning has yielded numerous outstanding results in the field of NLP, such as adversarial networks for text-image emotion classification (Al-qerem et al., 2020; Chopra et al., 2022). BLSTM is one of the most widely applied deep learning architectures. In NLP, the handling of contextual relations is paramount. Sole reliance on LSTM for sentiment classification is subject to issues like vanishing gradient and exploding gradient, and it is prone to loss of context memory linkage. BLSTM is capable of rectifying this problem.

Xianying et al. (2019) introduced a Chinese sentiment analysis model, WEEF-BILSTM, based on BLSTM and word embedding for the problem of prior knowledge and lack of semantic understanding in the traditional sentiment classification. Xianda et al. (2020) used a hybrid CNN and BLSTM for text sentiment analysis, illustrating the superiority of the CNN and BLSTM model. However, the approach is limited to online reviews at the sentence level and is not fine-grained sentiment classification, and a large number of labeled samples are required. Cai et al. (2020) used statistical methods to classify netizens' emotional orientation and employed a bidirectional encoding combination prediction method with BERT and BLSTM for sentiment classification. Sun et al. (2019) proposed an ERNIE-based model that improves the masking strategy on the BERT model, uses massive Chinese corpus for training, and incorporates large-scale knowledge graph data, thereby enabling the ERNIE model to better capture the implicit relationships in Chinese text and enhancing the model's semantic representation capability; however, further enhancements can be made in unsupervised pre-training. Xukan and Zeyu (2021) achieved some success in sentiment classification of WeChat posts using UGC data with a multi-scale BLSTM-CNN approach, though at the cost of slower computation speed.

### Short Text Sentiment Analysis

Information disseminated by emerging social tools is brief and concise, representing a typical short text structure. Short texts encompass extremely rich information with vast commercial value. However, they exhibit a high degree of colloquialism and non-standard expressions and include dialects, abbreviations, and slang, resulting in a lack of semantic features. The sparsity and irregularity of short texts have limited traditional sentiment classification methods, making the accurate determination of their sentiment even more challenging. Phan et al. (2023) from Poland and Korea approached this from a linguistic perspective, utilizing Graph Convolutional Networks (GCNs) to analyze aspect-level sentiment. Priyadarshini and Cotton (2021) proposed a model based on LSTM-CNN-Grid Search for coarse-grained sentiment analysis, with this optimized approach effectively enhancing the accuracy rate to over 96%, but it did not perform fine-grained sentiment analysis.

Feng et al. (2019) introduced a Chinese Fast Text short text classification method that integrates Term Frequency-Inverse Document Frequency (TF-IDF) and Latent Dirichlet Allocation (LDA).

During the model's input phase, this method filters the dictionary processed by the n-gram model using TF-IDF, and then uses the LDA model for topic analysis of the corpus to supplement the feature dictionary. This ensures that the model, when calculating the average of input word sequence vectors, favors terms with high discriminative power, achieving a higher precision in Chinese short text classification.

Most of the aforementioned feature representation algorithms for short texts segment the short text, focusing on features at the character or word level. Due to the sparse nature of short texts, characters or words fail to convey the full semantics of the short texts, leading to feature representation vectors that might not adequately capture the semantics of the short text. Although many researchers have improved classification algorithms, the input to these algorithms remains the feature representation vector of the short text, and any errors in this vector propagate down to the classifier. Therefore, the feature representation of short texts is pivotal in enhancing the performance of short text classification. Based on the research mentioned above, this paper will introduce a Chinese short text classification algorithm based on an improved Transformer of BERT.

## Attention Mechanism

The Att-M was posited by Bahdanau et al. (2016) and implemented in the domain of machine translation, achieving better performance than conventional neural network models. It has been a highlight in the application of deep learning in NPL in recent years. Originating from cognitive psychology, it simulates human attention, quickly pinpointing essential position information in a vast array of information, eliminating irrelevant information, and more efficiently completing tasks. The attention mechanism, based on the importance of each unit, assigns different attention weights. When obtaining the overall semantic representation of a sentence, it uses attention weights to perform a weighted sum of the semantic representation of all words, resulting in a more precise sentence semantic representation. There are various classifications of attention mechanisms. From the perspective of the application, they can be divided into spatial attention and temporal attention. Based on the working method, they can be categorized into soft attention and hard attention. Several variants of attention, such as Self-Att-M, Multi-Dimensional Att-M, Hierarchical Att-M, and Key-Value Attention, have been widely applied across various domains.

Zhonglin et al. (2020) utilized the fewer parameter dependencies of the self-Att-M to integrate it with the on-LSTM model, better capturing text features for sentiment analysis. Due to the hierarchical structure within the model and the need for dynamic pre-training via ELMo, the training duration is longer than the typical models, and the semantic information segmentation needs to be further improved (Zhonglin et al., 2020). Dong and Peisheng (2023) integrated the attention mechanism with capsule networks to address the shortcomings of extracting semantic information for Chinese short text sentiment classification. However, the model still has deficiencies such as a large number of parameters and a long training time, which will be improved. Ting et al. (2020) introduced a hierarchical attention mechanism, using BLSTM as the underlying structure. The hidden layer output of BLSTM is used as the input vector for the attention mechanism. Weighted sums form sentence vectors, and the sum of these vectors forms document vectors, which can be used for sentence-level sentiment classification tasks, showing notably better results than RNNs. A regional convolutional neural network is used to acquire sentiment features while adding aspect-specific information and improving LSTM with an internal dynamic control chain. However, the model needs to be improved for the sentiment analysis of cross-domain vocabulary and online phrases. Huan et al. (2020) combined Self-Attention with LDA, enhancing the accuracy of short text sentiment classification, exploring the possibility of using graph convolutional networks for simultaneous modeling of contextual semantics, the type of syntactic dependency, and the word semantic information associated with that dependency. However, this method cannot solve the problem of multiple meanings of words, and the simpler semantic information also limits its application scope.

Currently, applying attention mechanisms to Chinese short text sentiment analysis is still a novel attempt, and there is an urgent need to explore different attention models and optimize them. This paper proposes a model combining an improved Att-M with BERT-BLSTM for aspect-based sentiment classification of short texts.

## **M2BERT-BLSTM AA SENTIMENT ANALYSIS MODEL**

### **M2BERT-BLSTM AA Introduction**

This paper proposes M2BERT-BLSTM AA that is based on an enhanced BERT and BLSTM model and incorporates a Dual Channel Attention Mechanism. Firstly, an emotional information collection containing sentiment words, negation words, and adverbs of degree is constructed using existing emotional resources. Secondly, the sequence of the hidden layers of the BERT model is transformed into text character vectors, which are then concatenated following mean and max pooling along the dimension of sentence length. These concatenated word semantic features reflect the global feature information while highlighting salient feature information, which are input into the BLSTM for text sentiment analysis to mitigate evaluation imbalance. Subsequently, the text and sentiment information are encoded separately by the network of BLSTM and the fully connected network, employing distinct attention mechanisms to extract semantic features and emotional features from both sources.

### **Sentiment Information Extraction**

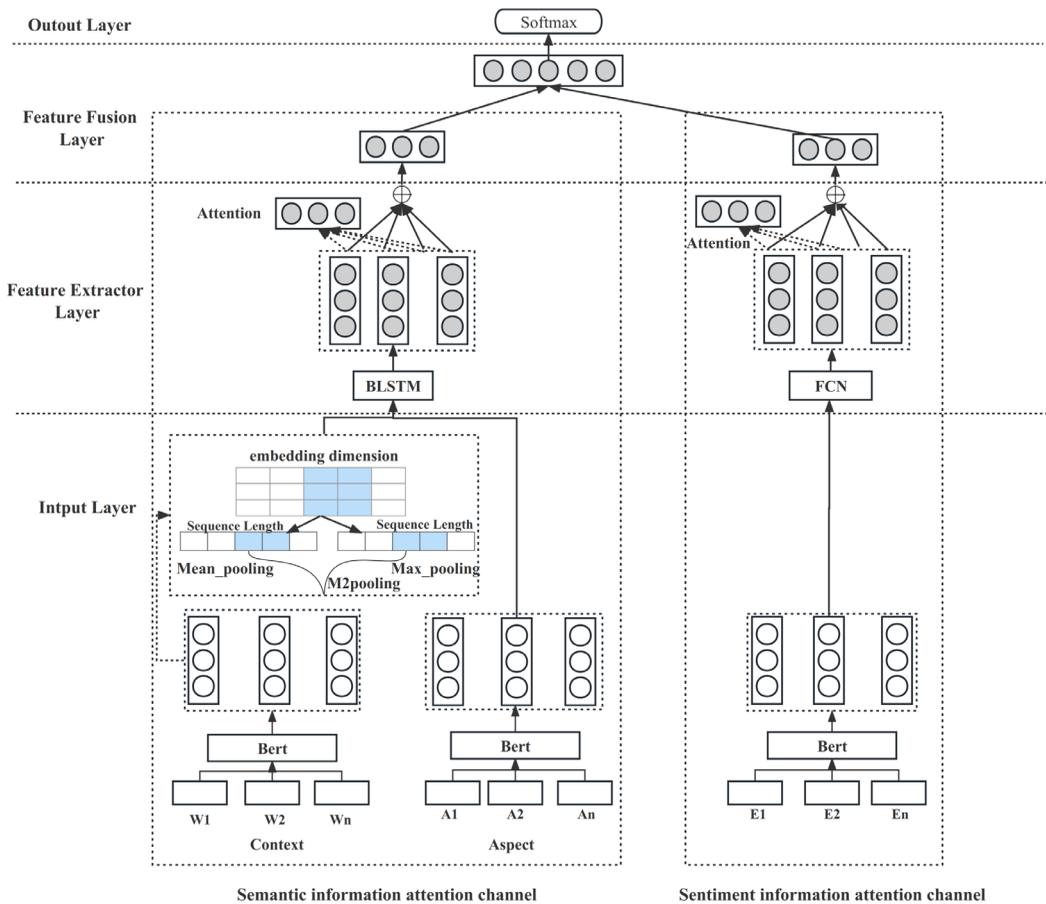
Short texts in social website comments contain a vast amount of sentiment information, such as sentiment words or combinations of modifiers with sentiment words. When a sentiment word is modified by a modifier, the sentiment polarity of the entire sentence might change, such as polarity reversal, intensification, or weakening. Modifiers generally include negation words, conjunctions, and adverbs (Runzhong & Ye, 2020). This paper considers both sentiment words and modifiers to build a sentiment resource library. By setting extraction rules, sentiment information contained in the text is extracted, resulting in a collection of sentiment information corresponding to each text. The extraction rules are as follows:

1. Rule 1: Traverse the text. If the current word is a sentiment word, add it directly to the sentiment information set.
2. Rule 2: Traverse the text. If the current word is an adverb or a negation word and the next word is a sentiment word, combine them and add them to the sentiment information set. Remove the sentiment word from the sentiment information set.
3. Rule 3: Traverse the text. If the current word is an adverb, followed immediately by a negation word and a sentiment word, combine the three and add them to the sentiment information set. Similarly, if the current word is a negation word, followed immediately by an adverb and a sentiment word, combine the three and add them to the sentiment information set.

### **M2BERT-BLSTM AA Model Construction**

The M2BERT-BLSTM AA model proposed in this paper has a vertical structure comprising four components: input layer, feature extraction layer, feature fusion layer, and sentiment output layer. The model's horizontal structure contains two information processing channels. The left channel, based on the BLSTM network and attention mechanism, is responsible for extracting contextual relationship features and aspect features of the review short text. The right channel, based on the fully connected network and attention mechanism, extracts the emotion features contained in the text. The improved BERT pre-training model captures the dynamic feature representation at the word level of the text. The overall model architecture is shown in Figure 1.

Figure 1. Structure of M2BERT-BLSTM AA model



### Input Layer

This model’s input layer includes Chinese comment short texts and the sentiment information they contain. For each text, in the semantic information channel, it is segmented at the character level and prefixed with the classification token [CLS], represented as  $W = \{[CLS], W_1, W_2, \dots, W_n\}$ . This serves as the input for the text context sequence. Moreover, for a detailed analysis of different aspects of the text, such as “delicious food but the wait is too long” with aspects like “food” and “wait”, aspect encoding is introduced to generate aspect vectors. In the sentiment information channel, based on sentiment information extraction rules, sentiment information contained in the short text is extracted, denoted as  $E = \{[CLS], E_1, E_2, \dots, E_n\}$ , and used as the input for the sentiment information channel.  $W_i$ ,  $A_i$ , and  $E_i$  represent the  $i$ -th character in the text, aspect information, and sentiment information, respectively. These are then input into the BERT pre-training model to obtain the feature vector matrices; for instance,  $W = w_1 \oplus w_2 \oplus \dots \oplus w_n$  and  $E = e_1 \oplus e_2 \oplus \dots \oplus e_n$ , where  $\oplus$  denotes the concatenation operation, and  $n$  represents the length of the information.

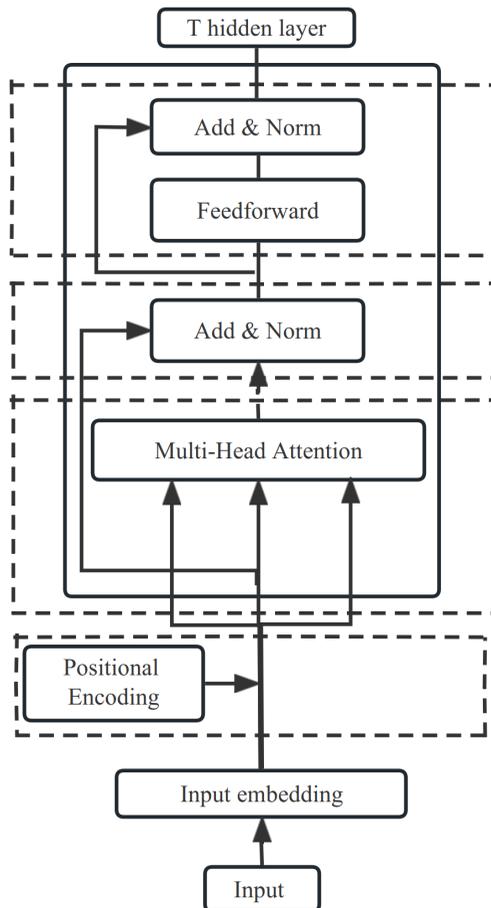
The model of BERT was introduced by Google’s Devlin et al. (2019) and others as a pre-trained language comprehension model based on multi-layer bidirectional Transformer encoders. For every input character or word, it is composed by adding token embeddings, segment embeddings, and positions embeddings. This offers a more robust word vector representation capability than other

traditional models, capturing more contextual information and reducing the model’s semantic understanding bias. BERT utilizes the random masking technique, adopting the Masked Language Model (MLM) and Next Sentence Prediction (NSP) as unsupervised targets, reflecting semantic spatial relationships and enhancing polysemy and homonym analysis capabilities.

In this study, the model leverages BERT for character embedding to obtain initial word embedding, which is then input into the neural network to enhance the fusion between feature vectors and contextual information, preventing the loss of crucial contextual semantic information. The Transformer model encoder is shown in Figure 2.

The input to the Encoder is a sentence’s character embedding representation, combined with the position embeddings of each character in the sentence. This is processed through the Multi-Head Attention layer, allowing the Encoder to reference preceding and following context when encoding each character. The output from the Encoder then passes through an Add & Norm layer. The term “Add” means adding the input and output of the layer that performs Multi-Head Attention, and “Norm” function normalizes the output by setting a fixed mean and standard deviation to 0 and 1, respectively. A fully connected feed forward neural network is then fed the normalized vector list. Similarly, the output from the feed forward layer undergoes an associated Add & Norm layer treatment before outputting a list of normalized word vectors. The most critical module in the Encoder is the

Figure 2. Structure of transformer encode



Multi-Head Attention. Its core goal is to compute the relationship of each word to every other word in a sentence and then adjust each word's weight accordingly.

From the hidden layers of the original BERT model, researchers only need to extract the vector (dimension is  $Thidden \in R[\text{batch size, sequence length, embedding dimension}]$ ) and input it into the downstream BLSTM for calculations. First, in the semantic channel's context feature pre-training, Maxpooling, Meanpooling, and M2pooling (a combination of average and max pooling) are computed along the sequence length dimension. Meanpooling takes the mean along the text length dimension, and Maxpooling only takes the maximum along the sentence length dimension. Mean pooling treats the majority of overall information features equally; max pooling results in sparser features but retains polarized features. The two are cascaded in order to capture different aspects of the information. Concatenating them emphasizes polarized features while retaining most general information features, enhancing the model's representational capabilities and creating richer and higher level feature representation. The goal is to reduce computational complexity while retaining and increasing sensitivity to important information. It is effective in improving model performance and generalization. In the experiments, problems such as fitting occur when the sample size is small, which can be solved by adding features to solve complex models with small samples using M2pooling.

$$Mean_{pooling} = \sum_{M=0}^n mean[Thidden(1)]_n \quad (1)$$

$$Max_{pooling} = \sum_{M=0}^n lookup\{\max[Thidden(1)]_n\} \quad (2)$$

$$M2_{pooling} = concatenate(Mean_{pooling} \& Max_{pooling}) \quad (3)$$

$N$  represents the embedding dimension.  $Thidden(1)$  indicates the position of the sequence length dimension in  $Thidden$  is 1.  $M2pooling \in R\{\text{batch size, embedding dimension} * 2\}$ . The vector  $g_t$  obtained after pooling serves as the input to BLSTM, represented by the formula as follows:

$$g_t = S_1(W_g * M_t + b_g) \quad (4)$$

Where  $t = \{1, 2, \dots, Y\}$ ,  $W_g \in R^{bg * Y}$ ,  $b_g$  is the bias of  $g_t$ , and  $M_t$  is the feature vector after concatenating maximum and mean values.  $S_1$  is the activation function. After generating context vectors through the pre-trained model and adding aspect encoding vectors, the sum of the generated character vector and the aspect encoding vector produces a more semantically rich vector representation. This layer can pay heightened attention to specific aspect features during training, thus effectively determining the sentimentality of various aspects.

### Feature Extraction Layer

Regarding the semantic features in review short texts, it is essential to consider not only the internal contextual semantic dependencies but also the semantic relevance of the context to aspect words. Furthermore, it is crucial to focus on parts more pertinent to sentiment classification and assign higher weights to significant sections. This study combines the BLSTM network with the Att-M to encode the context information and aspect word information of short texts, aiming to capture the context relationship features specific to text aspects.

The BLSTM sentiment analysis algorithm, derived from LSTM, comprises a forward LSTM and a backward LSTM. The BLSTM model treats the preceding and following sequences of each training sequence as two separate recurrent networks, then connects them to the same output layer. For each point in the sequence, as BLSTM can simultaneously access the respective forward and backward

sequence information, it can individually learn each word's left and right context information in the sequence, ensuring the integrity of the sequence information. Figure 3 shows its network structure.

The feed forward calculation of the LSTM neural network is as follows:

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

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (6)$$

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

$$\tilde{c}_t = \tanh(W_c \cdot [x_t, h_{t-1}] + b_c) \quad (8)$$

$$c_t = f_t \times c_{t-1} + i_t \times \tilde{c}_t \quad (9)$$

$$h_t = o_t \cdot \tanh(c_t) \quad (10)$$

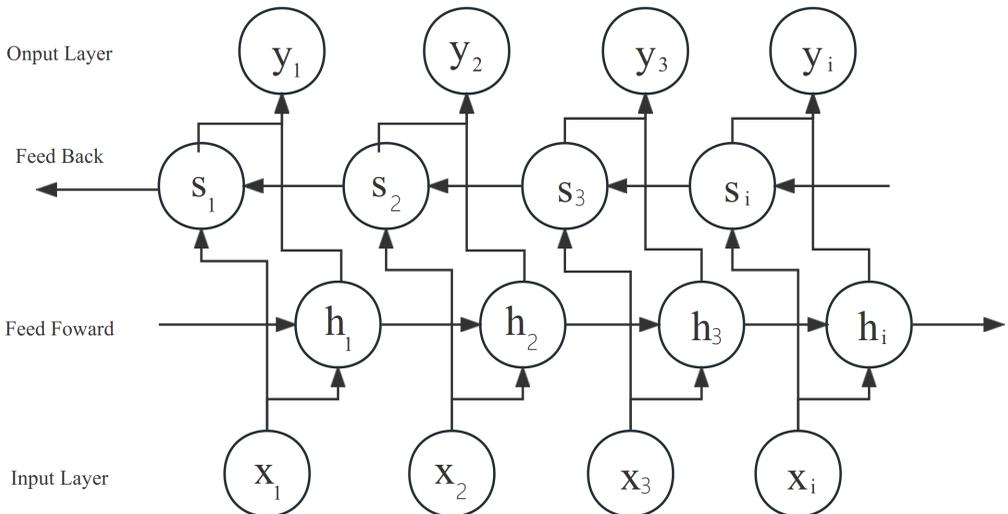
Where,  $x_t$  is the current input,  $f_t$  represents the forget gate,  $i_t$  represents the input gate; the output gate  $o_t$ ,  $W_i$ ,  $W_f$ , and  $W_o$  respectively represent the weight matrices between  $x_t$  and each gate.  $W_c$  is the weight matrix.  $b_f$ ,  $b_i$ ,  $b_o$  are the bias matrices corresponding to each gate, and  $b_c$  is the bias matrix.  $c_t$  is the current output state of the memory cell,  $c_{t-1}$  is the state of the previous moment, and  $\tilde{c}_t$  is the intermediate output state.  $h_t$  is the hidden layer output, and  $h_{t-1}$  is the hidden layer output of the previous moment.  $\sigma$  and  $\tanh$  are activation functions.

BLSTM takes the semantic channel character vectors of the input layer of the M2BERT-BLSTM AA model as input, and according to the calculation principle of the LSTM hidden layer, the output at time t is:

$$y_t = \text{concat}(\vec{h}_t + \vec{s}_t) \quad (11)$$

Where,  $x_1, x_2, x_3, \dots, x_i$  represents the semantic expression of the input data,  $h_1, h_2, h_3, \dots, h_i$  represents the hidden layer in the forward LSTM,  $s_1, s_2, s_3, \dots, s_i$  represents the hidden layer in the

Figure 3. Network structure of BLSTM



backward LSTM network, and  $y_1, y_2, y_3, \dots, y_i$  represents the semantic information after merging forward and backward. During forward calculation, the hidden layer's  $h_t$  is related to  $h_{t-1}$ , and during backward calculation, the hidden layer's  $s_t$  is related to  $s_{t-1}$ .

To extract more critical information from review short texts, the output results of the BLSTM are input into the Att-M to allocate weights to the feature information of each input at each moment, thereby obtaining semantic features. The attention calculation method is as follows:

$$z_t = \tanh(W_w y_t^c + b_w) \quad (12)$$

$$a_t = \frac{\exp(z_t^T z_w)}{\sum \exp(z_t^T z_w)} \quad (13)$$

$$v^c = \sum_t a_t y_t^c \quad (14)$$

Where,  $y_t^c$  is the output of the semantic channel BLSTM,  $z_t$  is the hidden state of  $y_t^c$ ,  $W_w$  and  $b_w$  are the adjustable weights and biases of the attention mechanism, respectively.  $z_w$  is the weight parameter of the Softmax classifier,  $a_t$  represents the importance information of the input state  $y_t^c$ , and  $v^c$  is the feature vector after calculating the attention model.

For the sentiment information channel, the sentiment information contained in review short texts only has functions such as sentiment representation, strengthening of sentiment polarity, weakening, and inversion, unlike text sequences that have strong semantic dependencies. Moreover, review texts may contain multiple sentiment words, and different sentiment words may influence the sentiment tendency of the review text to varying degrees. Therefore, a method combining a fully connected network with the Att-M is applied to encode sentiment information to capture the most prominent sentiment features. The attention calculation method is as follows:

$$u_t = \tanh(W_w y_t^e + b_w) \quad (15)$$

$$\beta_t = \frac{\exp(u_t^T u_w)}{\sum \exp(u_t^T u_w)} \quad (16)$$

$$v^e = \sum_t \beta_t y_t^e \quad (17)$$

Where  $W_w$  and  $b_w$  are adjustable weights and biases of the Att-M.  $u_w$  is the weight parameter of the Softmax classifier.  $\beta_t$  represents the weight of the input state  $y_t^e$ , and  $v^e$  is the sentiment vector after calculating the attention model in the sentiment channel.

### Feature Fusion Layer

The primary task of this layer is to merge the semantic feature vectors and sentiment feature vectors within the two channels to construct the short text's overall feature vector. To simplify the model's computation, row connection is adopted for feature fusion, concatenating  $v^c$  and  $v^e$  to obtain the final vector  $v^*$ .

$$v^* = [v^c, v^e] \quad (18)$$

### Sentiment Output Layer

Pass the feature vector from the feature fusion layer  $v^*$  through a softmax classifier to obtain the model's final prediction for the category to which the short text belongs. The calculation is as follows:

$$q = \text{softmax}(wv^* + b) \quad (19)$$

Where,  $w$  is the weight parameter matrix,  $b$  is the bias matrix, and  $q$  is the predicted sentiment probability distribution.

### Model Training

During the model training process, this paper uses cross-entropy as the loss function and updates parameters using end-to-end backpropagation.

$$L = - \sum_{i=1}^D \sum_{j=1}^C Y_i^j \log \hat{Y}_i^j \quad (20)$$

Where,  $D$  is the training dataset,  $C$  is the number of sentiment label categories,  $Y$  is the actual sentiment category, and  $\hat{Y}$  is the predicted sentiment category.

## EXPERIMENT AND ANALYSIS OF RESULTS

### Experimental Setup and Dataset

The programming language for this experiment is Python 3.7, the deep learning framework is TensorFlow 2.7.0. and NumPy 1.16. The experiment dataset includes the Meituan Review Dataset from Zhejiang Lab, a Restaurant Review Chinese Dataset and a laptop Chinese Dataset that have been adapted and processed from the combined and organized SemEval 2014 Task4, SemEval 2015, and SemEval 2016 datasets.

Specifically, the Meituan Review Dataset contains a total of 13,000 review data entries, including restaurant service quality, such as waiter's attitude, service speed, and so on; food quality, such as food taste, freshness, and quality; delivery service evaluation, such as punctuality, delivery staff's attitude, and so on. The reviews cover positive, negative, and neutral sentiments. The Restaurant Review Chinese Dataset contains a total of 9,575 review data entries, including restaurant environment, food, and service. The laptop Review Chinese Dataset contains a total of 2,598 review data entries, including appearance, price, and configuration. The text in the three datasets have a high degree of textual diversity due to being from different restaurants or manufacturers and different reviewers.

After deduplication, the three datasets are filtered using regular expressions. Characters irrelevant to sentiment expression such as "#", "@", and "/" as well as abbreviations and irregular punctuation are removed or modified. Additionally, stopwords are removed, and reviews that are too short, too long, objective event descriptions, or unclear in meaning are discarded to make the text structured and rigorous and improve robustness. The datasets are then analyzed with the LTP (Language Technology Platform) tool and manually annotated. Reviews are labeled into three categories: positive sentiment is marked as 1, neutral as 0, and negative sentiment as -1. In the end, the Meituan Review Dataset consists of 12,812 unique review entries and their sentiment orientations, including 4,875 positive, 5,235 neutral, and 2,702 negative. To ensure the accuracy and validity of the experiment, the dataset is randomly shuffled and divided into training, validation, and testing sets at a ratio of 8:1:1. The Restaurant Chinese Dataset has 6,330 review samples and their sentiment orientations, including 3,741

positive, 1,633 neutral, and 956 negative. The laptop Chinese Dataset has 2,566 review samples and their sentiment orientations, including 1,335 positive, 233 neutral, and 998 negative.

The sentiment lexicon is sourced from the Sentiment Ontology Database of Dalian University of Technology. The adverb of degree and negative words come from the HowNet Chinese Word Database and the NTUSD of National Taiwan University. The three are merged and deduplicated. During word segmentation, this paper uses the sentiment lexicon as a custom word segmentation dictionary, allowing the sentiment information in the text to exist as a complete linguistic unit. The sentiment lexicon is shown in Table 1.

## Parameter Settings

This paper involves numerous parameter settings. The BERT layer is based on GOOGLE’s pre-trained Chinese model “BERT-Base, Chinese,” which uses a 12-layer Transformer with a hidden size of 768. The optimizer used is Adam. The maximum text length is set to 200, and the maximum length of the sentiment information set is 40. The values of other parameters related to the BLSTM layer, etc. are shown in Table 2:

Among them, *Batch\_size* represents the data batch processing amount, *Dropout* is the random neuron deactivation rate to prevent overfitting, *LearningRate* is the initial learning rate for the Adam optimizer. *Hidden\_units* refer to the number of hidden units in BLSTM, and *Num\_layers* is the number of layers in BLSTM. In the Attention layer, the *attention\_size* parameter is set to 64, meaning each character or word vector is compressed to 64. The output vector dimension of the Attention layer is 768, and *hidden\_layers* is set to 12. *Epoch* uses an early-stop strategy. The number of aspect attention modules is 5.

Table 1. Sentiment resource library

Sentiment Information	Number	Examples
Sentiment Words	27,466	Outstanding, Diligent, Like, Weak
Negative Words	59	No, Never, Not as good as, Doesn’t have, Pointless
Degree Adverbs	219	Slightly, Overly, Extremely, Increasingly, Completely, A bit

Table 2. Model parameter setting parameter settingparameter setting

Parameter	Value
<i>Batch_size</i>	32
<i>Character embedding size</i>	200
<i>Embedding_Size</i>	768
<i>Multi-head Attention</i>	12
<i>Dropout</i>	0.5
<i>LearningRate</i>	1e-6
<i>hidden_units</i>	768
<i>num_layers</i>	4
<i>Act-function</i>	Tanh
<i>epoch</i>	Early stop

## Evaluation Metrics

This paper calculates the accuracy (A), precision (P), recall (R), and F1 score for the classification results of positive, neutral, and negative sentiments separately. Different aspects use independent models for learning, and the average values of the three categories are used as the overall performance evaluation indicators. Accuracy indicates the proportion of correctly classified samples in the dataset, reflecting the classifier's ability to correctly identify classification samples; recall reflects the proportion of correctly detected samples by the classifier in this class of samples in the dataset; the F1 score is the harmonic mean of accuracy and recall. The formulas are as follows:

$$A = \frac{Tp + Tn}{Tp + Tn + Fp + Fn} \quad (21)$$

$$P = \frac{Tp}{Tp + Fp} \quad (22)$$

$$R = \frac{Tp}{Tp + Fn} \quad (23)$$

$$F_1 = \frac{2 * P * R}{P + R} \quad (24)$$

Where TP stands for True Positives, where both the predicted and actual values are true; FP represents False Positives, where the predicted value is true but the actual value is false; FN denotes False Negatives, where the predicted value is false but the actual value is true; TN refers to True Negatives, where both the predicted and actual values are false. A 5-fold cross-validation approach could be used.

## Result Analysis

To explore the accuracy, precision, etc. of the M2BERT-BLSTM AA model, this paper conducted 24 experiments, calculating Precision, F1 score, Recall, and Accuracy on the Meituan, Restaurant, and laptop test datasets. In Figure 4, Figure 5, and Figure 6, the experimental results for the three datasets are shown respectively. Train-acc and train-loss indicate the accuracy and loss values of the training set during training, while val-acc and val-loss represent the accuracy and loss values of the validation set during training, respectively.

To demonstrate the efficiency and accuracy of the M2BERT-BLSTM AA model proposed in this paper, a comparative experiment was conducted with the GRU, LSTM, Bilstm, Attn-Lstm, and Graph-Bilstm sentiment analysis models, which currently have good results. To ensure the principle of a single variable in the comparative experiment, the same type of pre-trained word vector is used for all.

The accuracy, precision, F1, and recall trends for the six models on the Meituan review dataset are shown in Figures 7 through 10. The ROC-AUC curve on the Meituan dataset are shown in Figure 11.

A higher F1 score indicates better model performance.

The changes in accuracy, precision, F1, and recall for the six models on the Meituan review dataset are shown in Table 3.

From Figures 7 through 10 and Table 3, the experimental results show that, for the aspect-level fine-grained 3-category sentiment classification problem, the model proposed in this paper achieves commendable results compared to the other three learning methods, namely GRU, LSTM, BiLSTM, Attn-Lstm, and Graph-Bilstm, on the Meituan review dataset. Specifically, the accuracy rate reached 75.2%, which is a 5.9% increase compared to the GRU model, 6.1% higher than the LSTM model, 5.6% higher than the Graph-Bilstm model, and 5.2% better than the BiLSTM and Attn-Lstm. The precision rate achieved 68.6%, showing improvements of 0.1%, 2.1%, 4.1%, 0.6%, and 2.6% compared

Figure 4. Sentiment analysis results for Meituan Review dataset

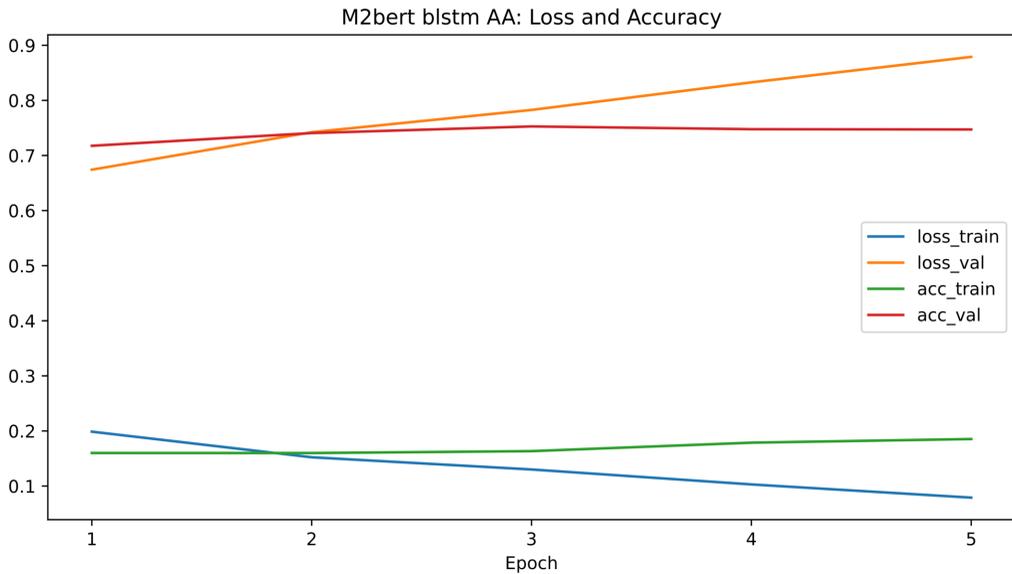
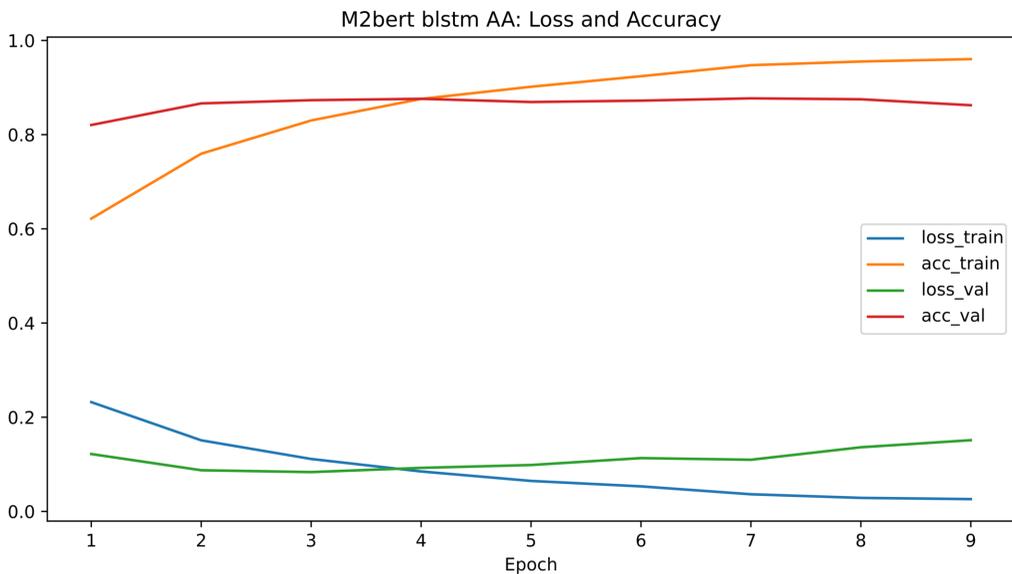


Figure 5. Sentiment analysis results for Restaurant Review dataset



to the GRU, LSTM, Bilstm, Attn-Lstm, and Graph-Bilstm models, respectively. Furthermore, recall is 64%, which is 13% better than the best-performing Attn-Lstm model, among the GRU, LSTM, Bilstm, Attn-Lstm, and Graph-Bilstm models. The F1 score was 66.5%, marking respective increments of 18.1%, 16.5%, 16.2%, 13%, and 15.3% when compared to the GRU, LSTM, Bilstm, Attn-Lstm, and Graph-Bilstm models. This can be attributed to the M2BERT-BLSTM AA model, which is a modification of the BiLSTM model with the integration of attention mechanism. This neural network model can focus on the feature information of the aspect, recognizing that different aspects of the

Figure 6. Sentiment analysis results for Laptop Review dataset

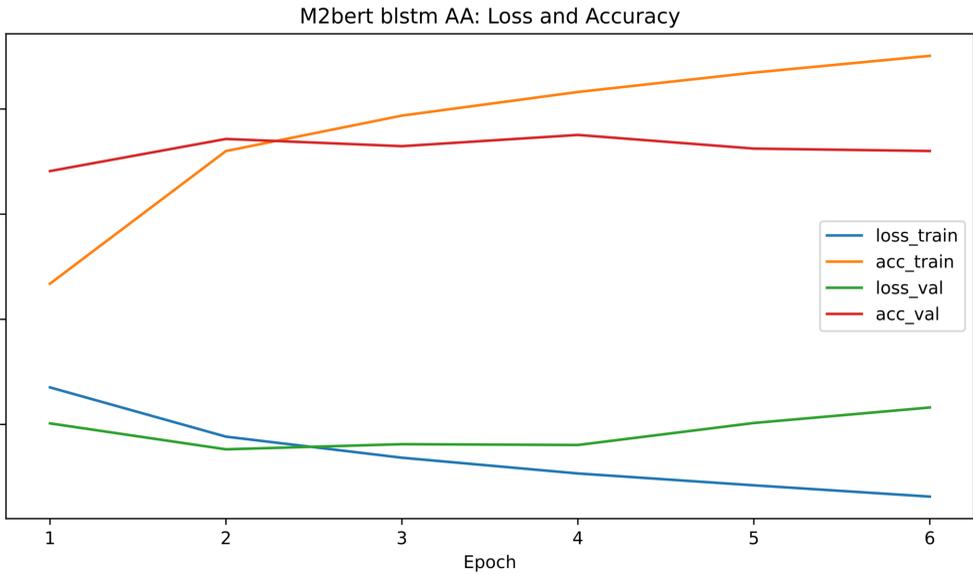
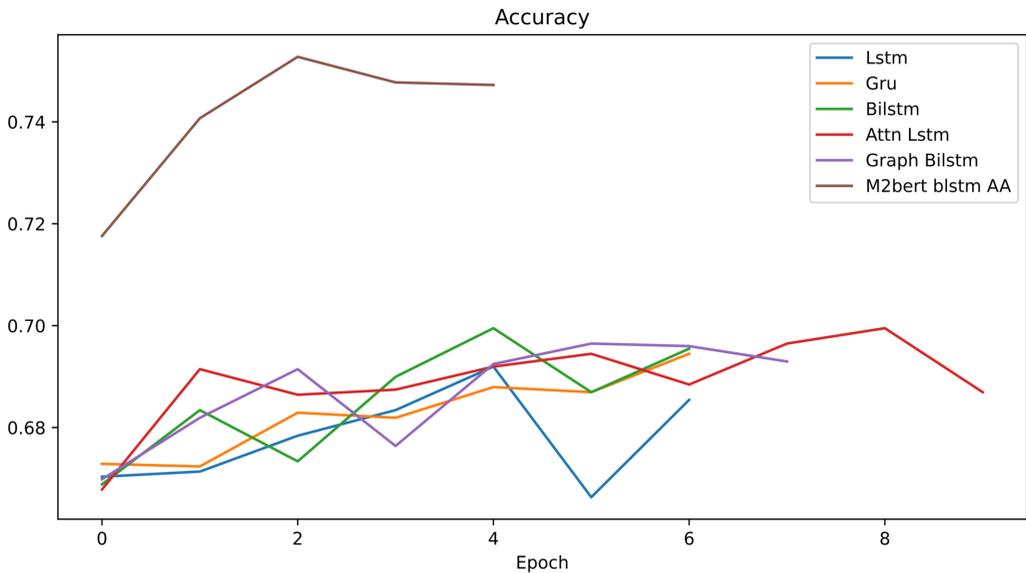


Figure 7. Sentiment analysis accuracy for Meituan Review dataset



same sentence may have different sentiments. The attention mechanism can concentrate intensely on the sentiment features of a particular aspect, accurately determining the sentiment polarity of different aspects within the same sentence. **A five-fold cross-validation approach could be employed for a more robust.** Five-fold cross-validation approach can be used for a more robust result.

Figure 8. Precision comparison results for sentiment analysis on Meituan Review dataset

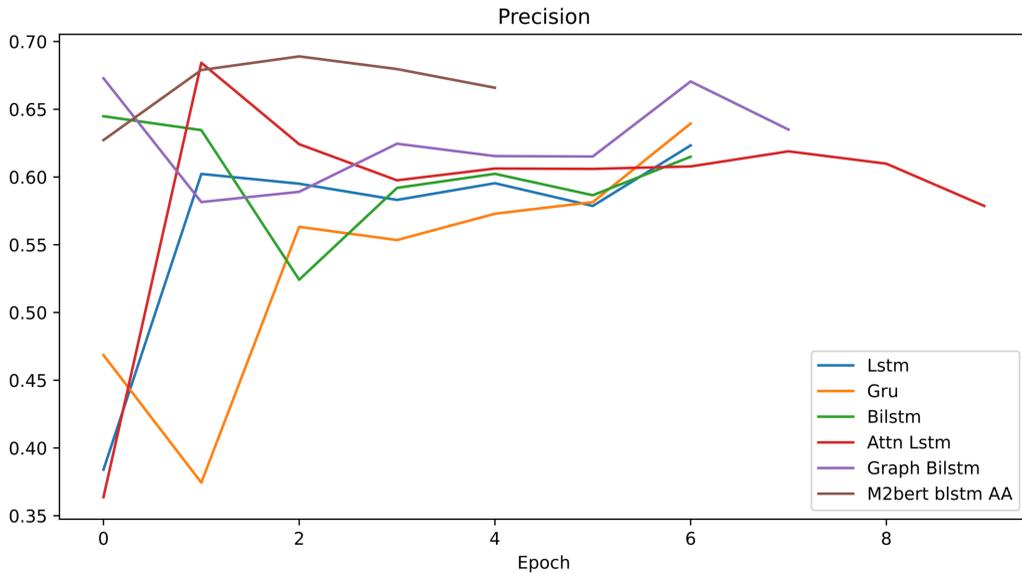
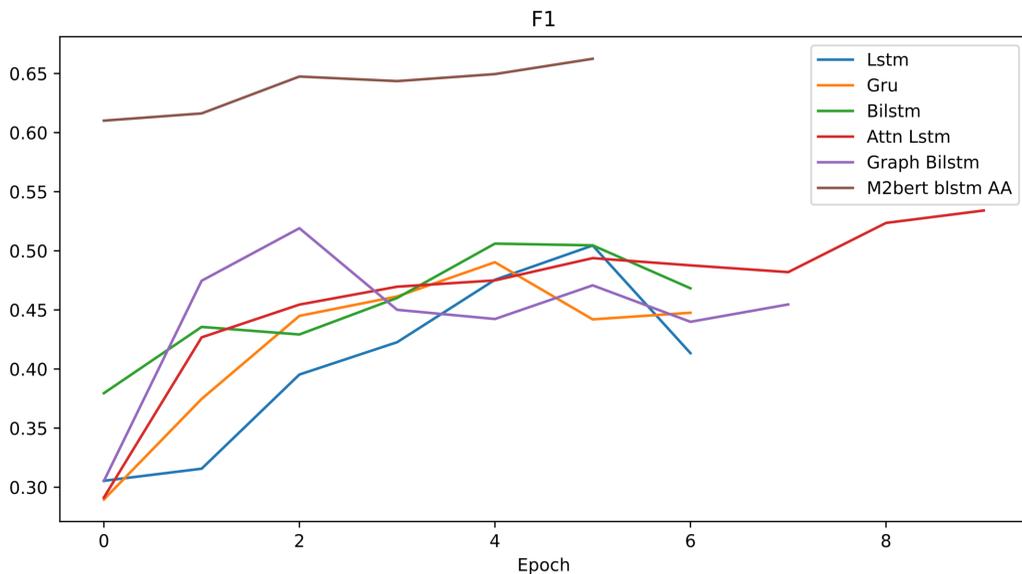


Figure 9. F1 Score comparison for sentiment analysis on Meituan Review dataset



The accuracy, precision, F1 score, and recall trends of the proposed M2BERT-BLSTM AA model on the restaurant dataset are shown in Figures 12 through 15. The ROC-AUC curve on the restaurant dataset is shown in Figure 16.

The variations in accuracy, precision, F1, and recall for the restaurant review dataset across the six models are displayed in Table 4.

The dataset covers five aspects: price, quality, service, ambiance, and others. The overall classification results can be seen from Figures 12 through 15. The LSTM model, without any added

Figure 10. Recall comparison results for sentiment analysis on Meituan Review dataset

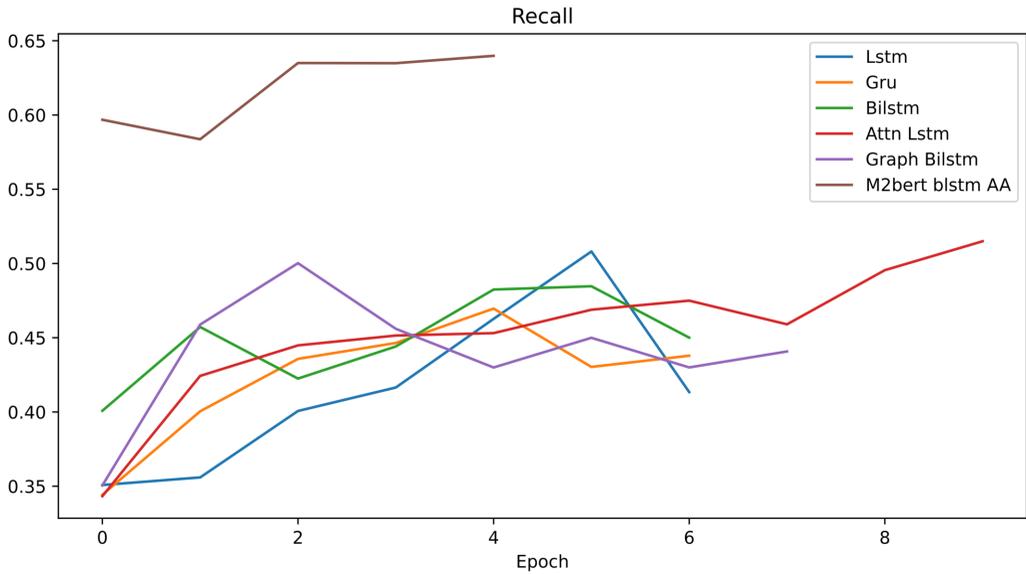
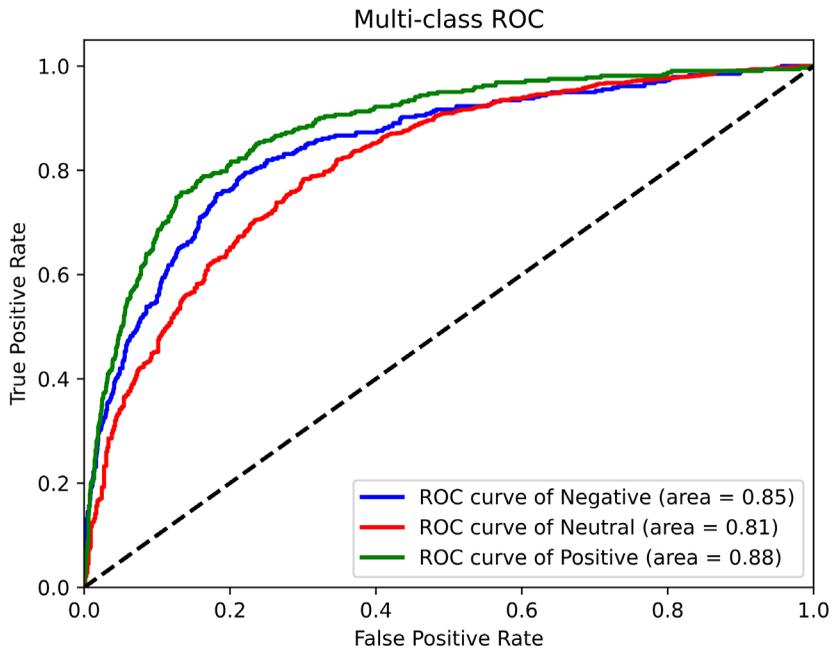


Figure 11. ROC Curve vs. AUC Area on Meituan Review dataset

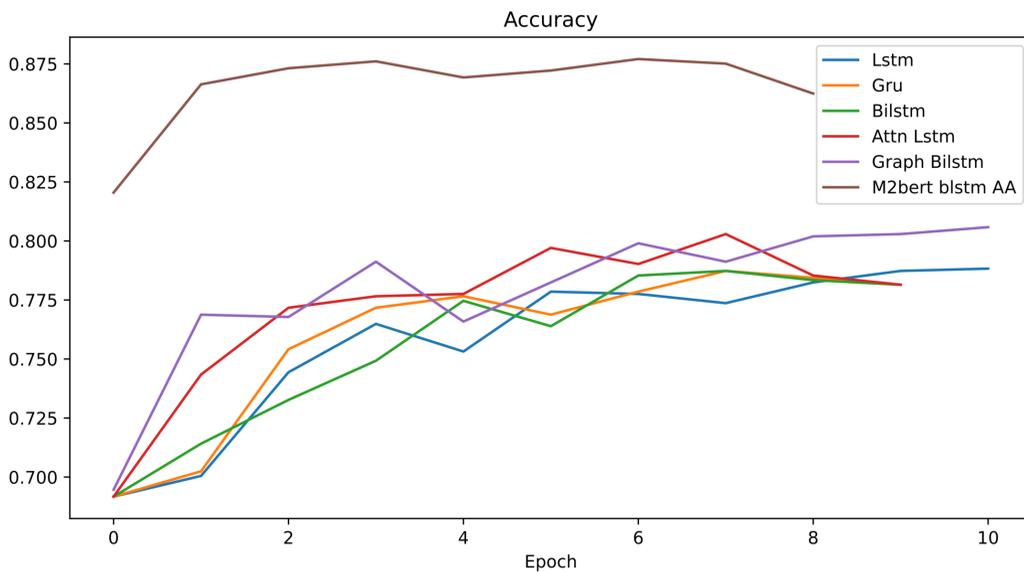


features, has lower values for accuracy, precision, and average F1 score across different datasets compared to the BiLSTM models. Specifically, the accuracy is lower by 0.9% and 0.2%. This indicates that the BiLSTM deep learning method is superior in feature extraction and learning, capturing sentiment information better than the LSTM method. Meanwhile, the M2BERT-BLSTM AA model

Table 3. Results of models on Meituan Review dataset

Model	Accuracy	Precision	F1	Recall
GRU	0.693	0.685	0.484	0.465
LSTM	0.691	0.665	0.5	0.505
BiLSTM	0.7	0.645	0.503	0.475
Attn Lstm	0.7	0.68	0.535	0.51
Graph Bilstm	0.696	0.66	0.512	0.498
M2BERT-BLSTM AA	0.752	0.686	0.665	0.64

Figure 12. Accuracy comparison results of sentiment classification on the Restaurant Review dataset



proposed in this paper performs well across different themed datasets. As evident from Figures 12 through 15 and Table 4, the M2BERT-BLSTM AA demonstrates that it is best aspect-based sentiment analysis on the restaurant dataset. After 10 iterations, the accuracy rate reaches 87.5%, improving by 9.2%, 9.2%, 9%, 7.5%, and 8.5%, over the GRU, LSTM, Bilstm, Attn-Lstm, and Graph-Bilstm models respectively. Precision hits 81%, marking an increase of 2.4%, 11%, 2%, and 1% over the GRU, LSTM, Bilstm, and Graph-Bilstm models respectively. The F1 score for aspect sentiment analysis on the restaurant review dataset reaches 79.7%, surpassing the best-performing Graph-Bilstm model by 11.7% and outpacing the GRU, LSTM, Bilstm, Attn-Lstm, and Graph-Bilstm models by 15.2%, 14.8%, 14.7%, and 13.7% respectively. The recall stands at 80%, outperforming the best of the GRU, LSTM, Bilstm, Attn-Lstm, and Graph-Bilstm models by 13.3%. This suggests that incorporating aspect information or employing attention mechanisms can significantly enhance classification outcomes, effectively processing texts characterized by varying sentiment polarities or unclear expressions.

The accuracy, precision, F1 score, and recall trends of the proposed M2BERT-BLSTM AA model on the laptop dataset are shown in Figures 17 through 20.

The variations in accuracy, precision, F1, and recall for the restaurant review dataset across the six models are displayed in Table 5.

Figure 13. Results of sentiment classification precision comparison on the Restaurant Reviews dataset

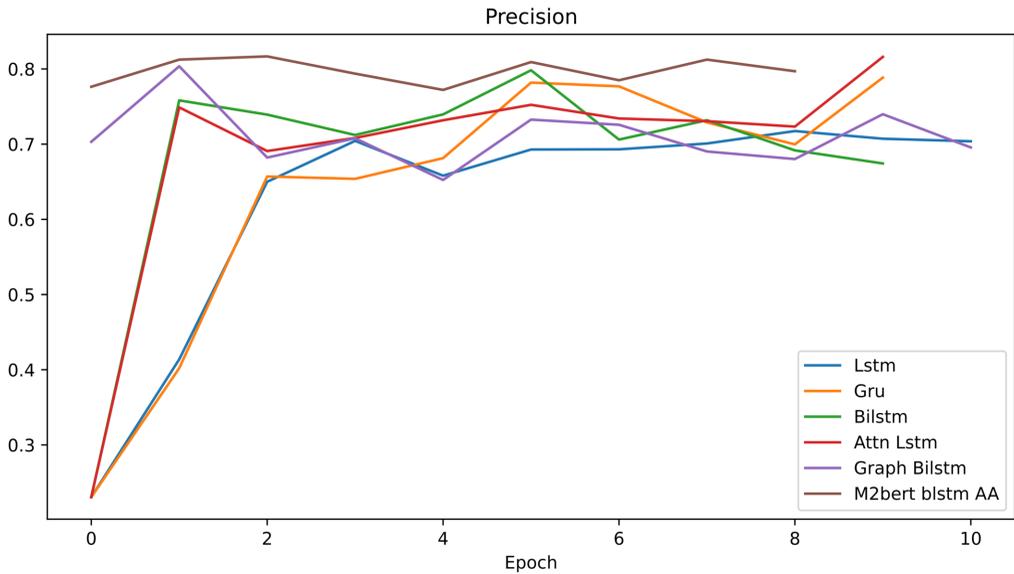
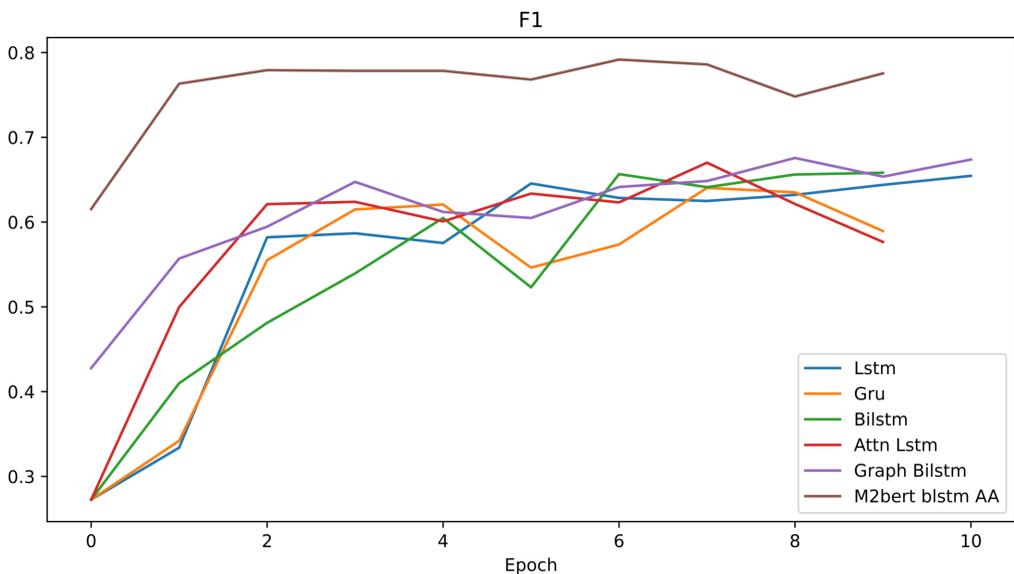


Figure 14. F1 Score results of analyzing the sentiment on the dataset containing reviews of restaurants



The dataset covers five aspects: price, quality, appearance, configuration, and others. The overall classification results can be seen from Figures 17 through 20. The M2BERT-BLSTM AA model has highest values for accuracy, precision, and average F1 score across different datasets compared to the GRU, LSTM, Bilstm, Attn-Lstm, and Graph-Bilstm models. This indicates that the proposed model is superior in feature extraction and learning, capturing sentiment information better than the other methods. As evident from Figures 17 through 20 and Table 5, the M2BERT-BLSTM AA demonstrates its best aspect-based sentiment analysis on the laptop dataset. The accuracy rate reaches

Figure 15. Recall comparison results of sentiment analysis on the Restaurant Review dataset

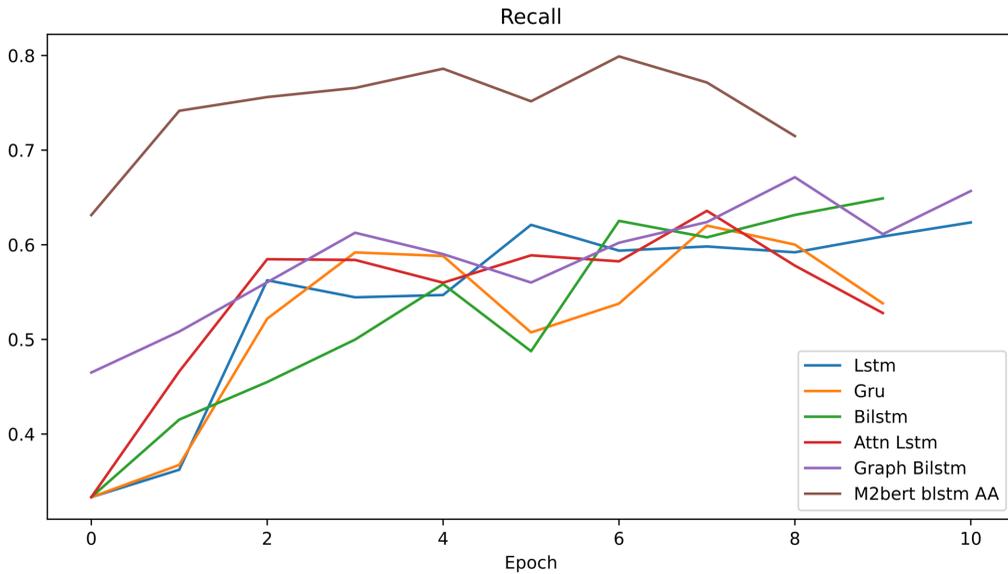
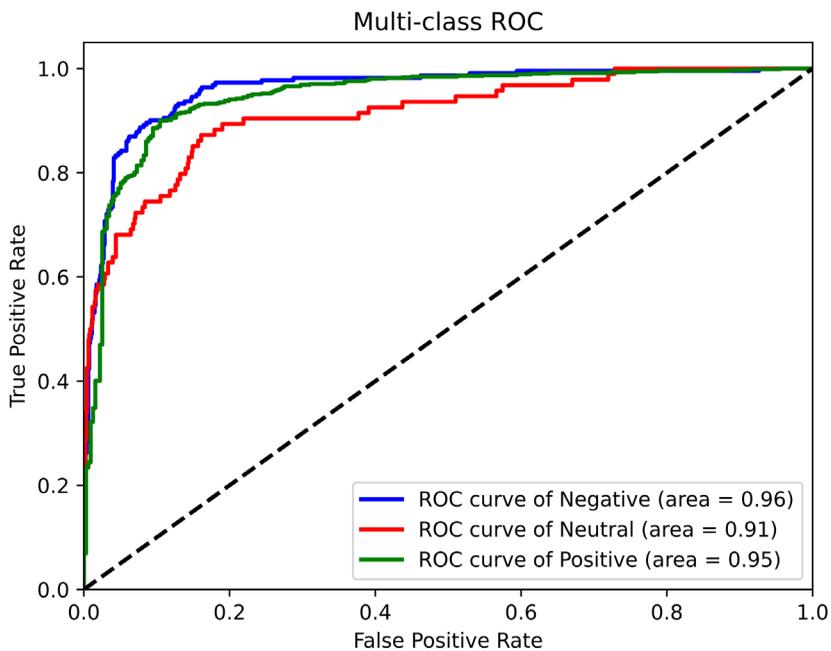


Figure 16. ROC Curve vs. AUC Area on Restaurant Review dataset

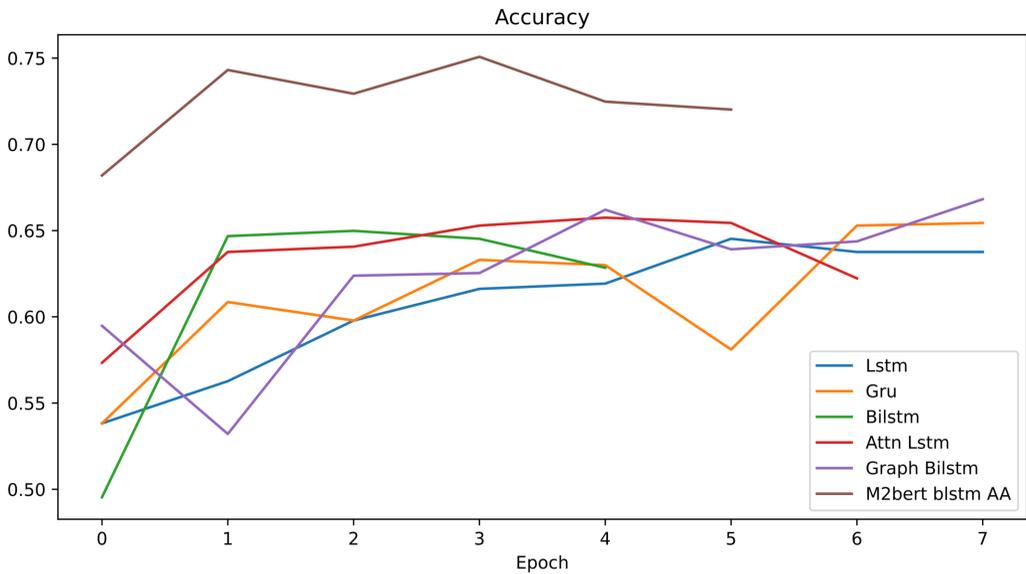


75%, improving by 9.5%, 10.7%, 10%, 9%, and 8% over the GRU, LSTM, Bilstm, Attn-Lstm, and Graph-Bilstm models respectively. Precision hits 71.2%, marking an increase of 8.2%, 6.2%, 2.2%, 10.2%, and 9.2% over the GRU, LSTM, Bilstm, Attn-Lstm, and Graph-Bilstm models respectively. The F1 score reaches 67.7%, surpassing the best-performing Graph-Bilstm model by 8.3% and outpacing

Table 4. Results of the Models on the Restaurant Review Dataset

Model	Accuracy	Precision	F1	Recall
GRU	0.783	0.786	0.645	0.625
LSTM	0.783	0.7	0.649	0.616
BiLSTM	0.785	0.79	0.65	0.65
Attn Lstm	0.8	0.81	0.66	0.634
Graph Bilstm	0.79	0.8	0.68	0.667
M2BERT-BLSTM AA	0.875	0.81	0.797	0.8

Figure 17. Accuracy comparison results of sentiment classification on the Laptop Review dataset



the GRU, LSTM, Bilstm, and Attn-Lstm models by 9.7%, 11.5%, 10.7%, and 8.7% respectively. The recall stands at 69.2%, on par with the best of the GRU, LSTM, Bilstm, Attn-Lstm, and Graph-Bilstm models. This validates the M2BERT-BLSTM AA model.

M2BERT-BLSTM AA model is robust in handling noisy, ambiguous, or incomplete data, which is attributed to the following: firstly, pre-processing and cleaning of data can denoise, deal with missing values, and improve the quality of the input data. Secondly, the M2BERT model has the adaptive learning capability that automatically learns the features of input data and adapts to different noise patterns. Learning from large amounts of pre-trained text data, it can handle ambiguous text. There is a strong representation ability to extract useful features from limited text and robustness to incomplete data. Thirdly, BLSTM captures long-term dependencies in sequential data to effectively deal with noisy data and also obtains contextual information in text to help the model understand the sentiment in ambiguous contexts and infer missing information. Finally, the dual attention mechanism focuses on the features that are most relevant for emotion classification, and choosing the right parameters such as dropout also improved robustness of the proposed model.

Figure 18. Results of sentiment classification precision comparison on the Laptop Reviews dataset

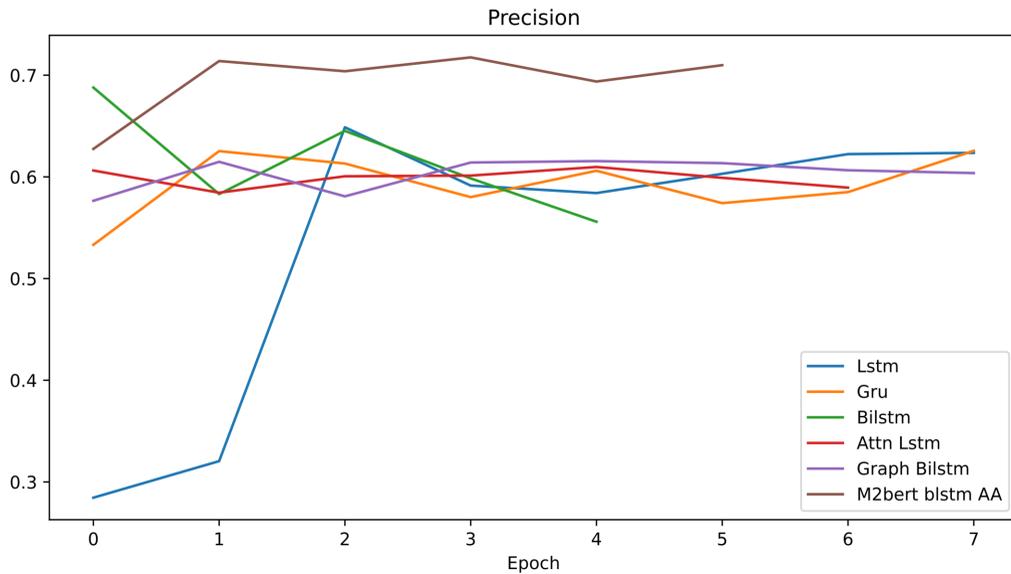
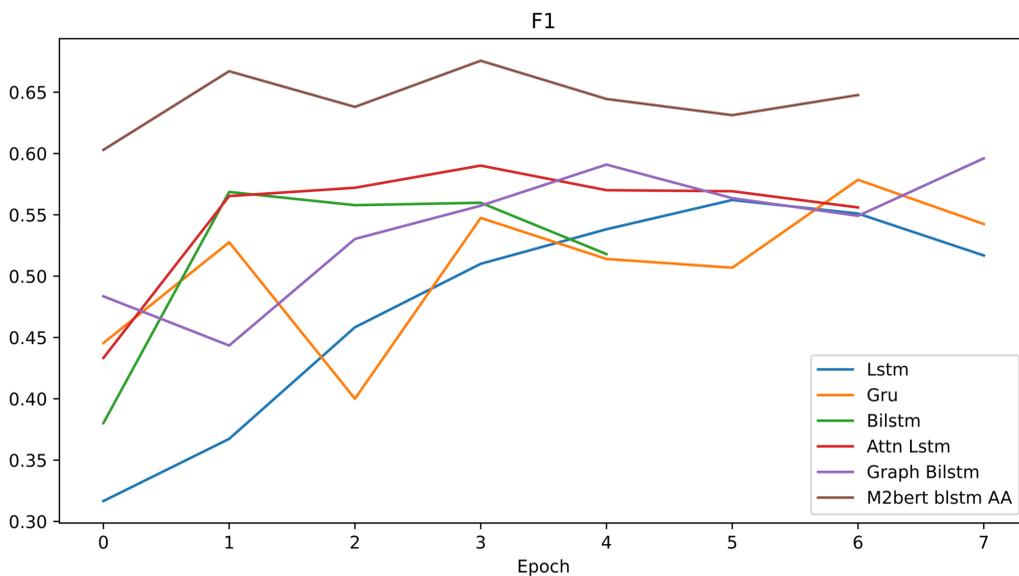


Figure 19. F1 Score results of analyzing the sentiment on the dataset containing reviews of laptop



### Overhead Analysis

The M2BERT-BLSTM-ATT model contains many complex components. Due to the BERT, BLSTM, and Att-M based models, there are multiple transformer layers, hidden states in both forward and backward directions, which increases the non-linearity of the model and the number of parameters. The attention mechanism dynamically focuses on the different positions of the input sequences, which also further induces additional parameters and computational overhead. The training time

Figure 20. Recall comparison results of sentiment analysis on the Laptop Review dataset

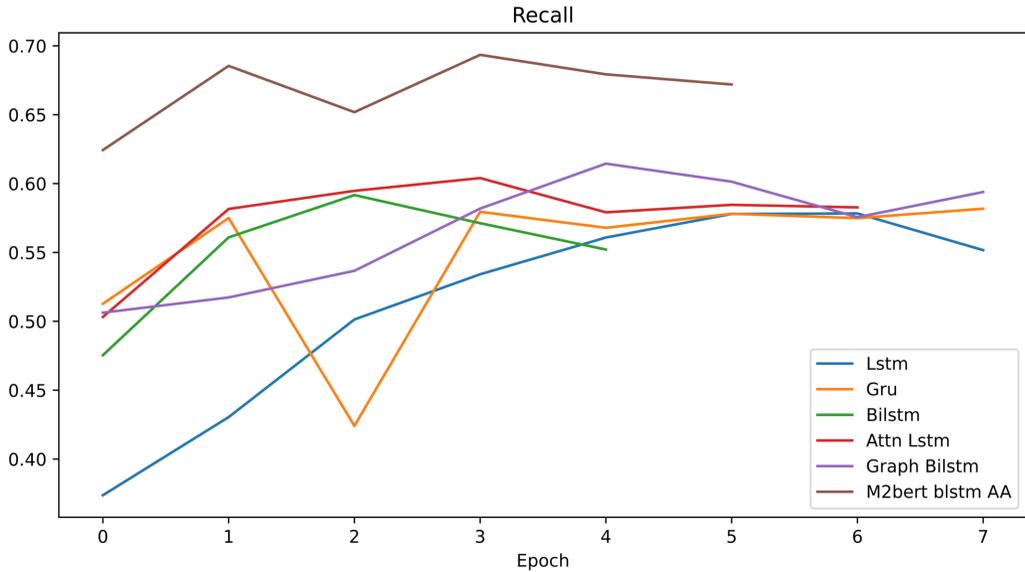


Table 5. Results of the models on the Restaurant Review dataset

Model	Accuracy	Precision	F1	Recall
GRU	0.655	0.63	0.58	0.579
LSTM	0.643	0.65	0.562	0.58
BiLSTM	0.65	0.69	0.57	0.59
Attn Lstm	0.66	0.61	0.59	0.692
Graph Bilstm	0.67	0.62	0.594	0.615
M2BERT-BLSTM AA	0.75	0.712	0.677	0.692

of the M2BERT-BLSTM-ATT model is relatively long on large-scale datasets. The computational overhead is also high when inference is performed in application scenarios that require real-time inference. The proposed model runs in 446s, training duration 2230s on the Meituan ordering dataset, 1786s on the restaurant dataset, and 1140s on the laptop dataset, which is less time compared to the traditional model that does not use the attention mechanism. In order to simplify the complexity and computation of the model, a simple vector concatenation is used; for example, M2pooling is a simple concatenation of Maxpooling and Meanpooling, and in the dual channel attention mechanism, the contextual feature vectors and sentiment feature vectors are also connected along the dimensional direction to form the final feature vector. The author will explore the optimization and compression of the model in the future.

The M2BERT-BLSTM-ATT model is integrated in the application system. Deploying the model as an API service by converting the model into an executable file or using container technology. An

Aspect-level user review sentiment analysis system or Merchandise service recommendation system that require good load capacity and fast responses can be formed.

The interpretability of the model helps the user to understand the decision making process of the model. For example, the M2BERT-BLSTM AA model can be applied in a recommender system, forming a product interpretable recommender system or restaurant interpretable recommender system. It is possible to form an interpretation template according the recommendation results. For example, when a user chooses a restaurant, he/she may pay attention to the restaurant in terms of cost-effectiveness, environment, and taste, etc. When there is a relatively good performance in these aspects, the system can recommend the restaurant to the user. If the predicted rating is low, it can explain why this restaurant is not suitable for a particular user.

When M2BERT-BLSTM AA is applied in social networks and intelligent systems, mass access to user-generated content involves a number of ethical considerations and privacy issues. First, the collection, storage, and processing of personal information is subject to regulations and data minimization principles to prevent disclosure or misuse of data that misleads or manipulates user sentiment. Second, models may be influenced by data bias, leading to biased judgments about certain groups or specific sentiments. This can trigger discriminatory comments or decisions, and such biases need to be minimized through diverse and balanced data. The proposed model provides transparency and interpretability to enable users to understand the decision-making process and build their trust in the model.

In conclusion, based on the experimental results, the M2BERT-BLSTM AA model shows improved values in accuracy, precision, F1, and recall on both the Meituan review dataset and the restaurant dataset compared to traditional deep learning models. Not only does M2BERT-BLSTM AA incorporate an attention mechanism, but using multiple aspect attention modules also extracts hidden features of specific aspects, enhancing sentiment classification. The model also uses both semantic and sentiment channels, and, in the semantic channel's context feature training, mean and max pooling are computed along the sequence length dimension and concatenated, retaining most of the overall information while emphasizing polarized features. This model can separately model sentences and aspects while considering inter-sentence relationships, addressing the issue where base models cannot differentiate between different aspects within a single sentence, causing misclassifications when similar sentiment polarities appear across different aspects. The model proposed in this paper achieves commendable results in aspect sentiment analysis across different themes, validating the effectiveness of the introduced methodology.

## CONCLUSION

This paper analyzes the current research status of natural language sentiment analysis technology and the limitations of traditional deep learning models in feature word selection and text classification. A dual-channel deep learning sentiment analysis model, M2BERT-BLSTM AA, based on the constructed sentiment resource library, improved BERT character vector preprocessing method, BLSTM, and aspect attention mechanism, is proposed. This model uses character vectors to extract more textual information than word vectors, concatenating based on mean and max pooling along the sentence length dimension. It extracts features from review texts and embedded sentiment information, obtaining the final feature vector representation. By integrating sentiment resources with the attention mechanism, the model captures rich feature information, focusing on the polarity of crucial sentiment words in specific aspects and contextual word information. This addresses the challenge where the same word in different contexts represents different meanings in traditional word vectors. Experimental results indicate that the proposed model achieves good classification results in aspect-level sentiment multi-classification on both the Meituan review, restaurant, and laptop datasets, with accuracy rates reaching 75.2%, 87.5%, and 75%, respectively, outperforming the best traditional deep learning model, BiLSTM, by 5.2% and the latest model, Graph-Bilstm, by 8.5 and 8.9%. Precision hit 68.6%,

81%, and 71.2%, all higher than the comparative models. F1 levels of 66.5%, 79.7%, and 67.7% are achieved, which are 13% higher than the traditional attention model Attn-lstm and 11.7% and 8.3% higher than Graph-Bilstm.

Overall, the M2BERT-BLSTM AA model, leveraging a dual-channel attention mechanism, effectively enhances the ability to capture text sentiment semantics, subsequently boosting the performance of fine-grained aspect-based sentiment analysis.

The proposed model is currently trained and used mainly in the Chinese datasets. Different linguistic contexts or dataset sizes may lead to the performance degradation. The model will be trained using multi-language datasets or expanded datasets in the future to improve the generalization ability and adaptability. The paper has yet to consider the authenticity of reviews and the limitations of review data range. Future training of sentiment classification models will consider integrating fake review scenarios and further incorporating features such as emoticons, shallow semantic information, and semantics to explore further improvements in sentiment classification outcomes. Also, exploring the interpretability of the proposed model is a direction for further research.

We have no known conflict of interest to disclose.

Correspondence concerning this article should be addressed to College of Computer and Control Engineering, Northeast Forestry University, Harbin, Heilongjiang 150040, China.

Corresponding Author: Yan Li, Email: woshigou2011@163.com

The authors thank the editor and anonymous reviewers for their contributions toward improving the quality of this paper.

## REFERENCES

- Abdi, A., Shamsuddin, S. M., Hasan, S., & Piran, J. (2019). Deep learning based sentiment classification of evaluative text based on multi-feature fusion. *Information Processing & Management*, 56(4), 1245–1259. doi:10.1016/j.ipm.2019.02.018
- Al-qerem, A., Alauthman, M., Almomani, A., & Gupta, B. B. (2020). IoT transaction processing through cooperative concurrency control on fog–cloud computing environment. *Soft Computing*, 24(8), 5695–5711. doi:10.1007/s00500-019-04220-y
- Bahdanau, D., Chorowski, J., Serdyuk, D., Brakel, P., & Bengio, Y. (2016). *End-to-end attention-based large vocabulary speech recognition*. IEEE. doi:10.1109/ICASSP.2016.7472618
- Barbosa, A., Bittencourt, I. I., Siqueira, S. W., Dermeval, D., & Cruz, N. J. (2022). A context-independent ontological linked data alignment approach to instance matching. [IJSWIS]. *International Journal on Semantic Web and Information Systems*, 18(1), 1–29. doi:10.4018/IJSWIS.295977
- Barr, J. R., Shaw, P., Abu-Khzam, F. N., Thatcher, T., & Hocking, T. D. (2022). Graph embedding: A methodological survey. In *2022 fourth international conference on transdisciplinary AI (TransAI)* (pp. 142–148). IEEE. doi:10.1109/TransAI54797.2022.00031
- Bin, S., Xin, Y., Lihua, Z., Guangyi, X., & Yanchao, L. (2022). Sentiment analysis of Weibo based on ERNIE and dual attention mechanism. *Yunnan Daxue Xuebao. Ziran Kexue Ban*, 44(03), 480–489.
- Cai, R., Qin, B., Chen, Y., Zhang, L., Yang, R., Chen, S., & Wang, W. (2020). Sentiment analysis about investors and consumers in energy market based on BERT-BiLSTM. *IEEE Access : Practical Innovations, Open Solutions*, 8(8), 171408–171415. doi:10.1109/ACCESS.2020.3024750
- Chopra, M., Singh, S. K., Sharma, A., & Gill, S. S. (2022). A comparative study of generative adversarial networks for text-to-image synthesis. [IJSSCI]. *International Journal of Software Science and Computational Intelligence*, 14(1), 1–12. doi:10.4018/IJSSCI.300364
- Dandan, D., Jiashan, T., Yong, W., & Kehai, Y. (2021). Chinese short text classification algorithm based on BERT mode. *Computer Engineering*, 47(1), 79–86.
- Daou, G., Maroun, C. B., & Hammoud, B. (2021). Advanced iterative multi-frequency algorithm used by radar remote-sensing systems for oil-spill thickness estimation. In *2021 international conference on electrical, computer and energy technologies (ICECET)* (pp. 1-6). IEEE. doi:10.1109/ICECET52533.2021.9698742
- Devlin, J., & Chang, M. W., & Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Minneapolis, Minnesota: Association for Computational Linguistics.
- Do, P. T., Phan, T., Le, H., & Gupta, B. B. (2022). Building a knowledge graph by using cross-lingual transfer method and distributed MinIE algorithm on apache spark. *Neural Computing & Applications*, 34(11), 1–17. doi:10.1007/s00521-020-05495-1
- Dong, W., & Peisheng, L. (2023). Sentiment analysis of Chinese short texts integrating capsule networks. [Natural Science]. *Journal of Chongqing University of Technology*, 37(05), 178–184.
- Feng, Y., Bohao, Q., & Hongyan, X. (2019). Chinese FastText short text classification method based on TF-IDF and LDA. *Journal of Applied Sciences*, 37(3), 378–388.
- Hajjar, A., & Tekli, J. (2022). Unsupervised extractive text summarization using frequency-based sentence clustering. In *European conference on advances in databases and information systems* (pp. 245–255). Cham: Springer International Publishing. doi:10.1007/978-3-031-15743-1\_23
- Huan, C., Bo, H., Yimin, Z., Lei, Y., & Yuxin, Y. (2020). Short text sentiment classification method combining LDA and self-attention. *Computer Engineering and Applications*, 56(18), 165–170.
- Ismail, S., Shishtawy, T. E., & Alsammak, A. K. (2022). A new alignment word-space approach for measuring semantic similarity for Arabic text. [IJSWIS]. *International Journal on Semantic Web and Information Systems*, 18(1), 1–18. doi:10.4018/IJSWIS.297036

- Jiang, L., Zou, Z. W., Liao, J., & Li, Y. (2023). A study on the application of sentiment-support words on aspect-based sentiment analysis. *International Journal of Pattern Recognition and Artificial Intelligence*, 37(7), 2357004. doi:10.1142/S0218001423570045
- Phan, H. T., Nguyen, N. T., & Hwang, D. (2023). Aspect-level sentiment analysis: A survey of graph convolutional network methods. *Information Fusion*, 91, 149–172. doi:10.1016/j.inffus.2022.10.004
- Poria, S., Hazarika, D., Majumder, N., & Mihalcea, R. (2023). Beneath the tip of the iceberg: Current challenges and new directions in sentiment analysis research. *IEEE Transactions on Affective Computing*, 14(1), 108–132. doi:10.1109/TAFFC.2020.3038167
- Priyadarshini, I., & Cotton, C. (2021). A novel LSTM-CNN-grid search-based deep neural network for sentiment analysis. *The Journal of Supercomputing*, 77(12), 13911–13932. doi:10.1007/s11227-021-03838-w PMID:33967391
- Runzhong, X., & Ye, L. (2020). Text sentiment classification model based on BERT and dual-channel attention. *Journal of Data Acquisition and Processing*, 35(04), 642–652. doi:10.16337/j.1004-9037.2020.04.005
- Sahoo, S. R., & Gupta, B. B. (2021). Multiple features based approach for automatic fake news detection on social networks using deep learning. *Applied Soft Computing*, 100, 106983. doi:10.1016/j.asoc.2020.106983
- Sarivougioukas, J., & Vagelatos, A. (2022). Fused contextual data with threading technology to accelerate processing in home UbiHealth. [IJSSCI]. *International Journal of Software Science and Computational Intelligence*, 14(1), 1–14. doi:10.4018/IJSSCI.285590
- Sarkissian, S., & Tekli, J. (2021). Unsupervised topical organization of documents using corpus-based text analysis. In *Proceedings of the 13th international conference on management of digital ecosystems* (pp. 87–94). ACM. doi:10.1145/3444757.3485078
- Sun, Y., Wang, S., Li, Y., Feng, S., Tian, H., Wu, H., & Wang, H. (2019). ERNIE 2. 0: A continual pre-training framework for language understanding. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(5), 8968–8975. ACM.
- Tekli, J., Tekli, G., & Chbeir, R. (2021). Almost linear semantic XML keyword search. In *Proceedings of the 13th international conference on management of digital ecosystems* (pp. 129–138). doi:10.1145/3444757.3485079
- Ting, S., Zhanwei, C., & Haifeng, Y. (2020). Aspect sentiment analysis based on hierarchical attention networks. *Big Data*, 6(5), 10.
- Wang, T., & Yang, W. Z. (2021). Review of text sentiment analysis methods. *Computer Engineering and Applications*, 57(12), 11–24.
- Xianda, G., Risa, N., & Shaoze, C. (2020). Consumer online review sentiment analysis based on CNN-BiLSTM. *Systems Engineering Theory and Practice*, 40(3), 11.
- Xianying, H., Guangfeng, L., Xiaoyang, L., & Anzhi, Y. (2019). Deep sentiment classification model based on word2vec and bidirectional LSTM. *Computer Application Research*, 36(12), 6.
- Xingjie, F., & Yunze, Z. (2020). Deep learning e-commerce recommendation model based on dynamic aspect attention user reviews. *Computer Applications and Software*, 37(03), 38–44+71.
- Xukan, X., & Zeyu, Z. (2021). WeChat sentiment classification model and application research based on multi-scale BiLSTM-CNN. *Information Science*.
- Zhang, C., Xiangyuan, M., Yang, Z., & Renzhong, G. (2021). Analysis of public opinion evolution based on user sentiment changes during the COVID-19 pandemic. *Journal of Geo-Information Science*, 23(2), 10. doi:10.12082/dqxkx.2021.200248
- Zhang, Q., Guo, Z., Zhu, Y., Vijayakumar, P., Castiglione, A., & Gupta, B. B. (2023). A deep learning-based fast fake news detection model for cyber-physical social services. *Pattern Recognition Letters*, 168, 31–38. doi:10.1016/j.patrec.2023.02.026
- Zhonglin, Z., Linchuan, L., Xiangqi, Z., & Haiyun, M. (2020). Aspect sentiment analysis based on ON-LSTM and self-attention mechanism. *Small and Micro Computer Systems*, 41(9), 6.

*Yan Li, lecturer, Graduated from Northeast Forestry University in 2015. Worked in Northeast Forestry University. Her research interests include natural language processing, Software Engineering, social media analysis, artificial intelligence.*