International Journal of Computer Science and Information Technology
ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 5, Number 1, Year 2025
DOI: https://doi.org/10.62051/ijcsit.v5n1.09
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Research on Sentiment Analysis of Weibo Comments Based on BERT-BILSTM-Attention Deep Learning

Xu Xiong

School of Cyber Science and Engineering, Southeast University, China
xxu94200@gmail.com

ABSTRACT

Sentiment analysis of COVID-19-related content on Weibo is of significant importance for studying public sentiment during the pandemic and economic recovery. Due to the lack of well-annotated Chinese Weibo COVID-19 data (such as the Weibo NCOV dataset), as well as the emotional complexity and ambiguity of Chinese Weibo texts, this paper proposes an innovative sentiment analysis model for Chinese Weibo COVID-19 data, namely BERT-BiLSTM-Attention. The model first encodes Weibo comment data using BERT to enhance the semantic feature representation of the text and improve its contextual understanding. Next, BiLSTM is used to enrich the contextual information of the Weibo text, helping to extract important and effective information from the text sequences. Finally, an Attention mechanism is employed to quickly capture the most relevant information. Experimental results show that the model is effective in sentiment analysis of Weibo COVID-19 data, achieving an accuracy of 88.2%. It can be concluded that the proposed model significantly improves the performance of Weibo text classification and demonstrates strong generalizability, making it suitable for sentiment analysis in various fields.

KEYWORDS

Weibo NCOV Data; Sentiment Analysis; BERT-BiLSTM-Attention Model

1. INTRODUCTION

In today's era of rapidly disseminated digital information, Weibo has become a prominent social platform for sharing opinions and expressing emotions. The wealth of emotional information embedded in Weibo comments holds critical significance across various fields, such as public opinion monitoring, brand management, and social psychology research. Consequently, sentiment analysis of Weibo comments has emerged as a research hotspot in natural language processing (NLP) and machine learning (ML).

The rapid advancement of deep learning techniques has provided new opportunities for sentiment analysis. In recent years, various deep learning architectures have demonstrated strong capabilities in text sentiment recognition tasks. Against this backdrop, this study aims to explore in depth the methods and practices of using the innovative BERT-BiLSTM model architecture for sentiment analysis of Weibo comments. The research comprehensively covers the entire process, from data collection and preprocessing to model construction, training, evaluation, and optimization, aiming to provide new ideas and methodologies for research and application in this field.

For data collection, comments related to specific themes (e.g., trending societal events, product releases) on the Weibo platform were targeted to construct a dedicated Weibo comment dataset, named "nCoV." Using a carefully designed data collection procedure, the official Weibo API was leveraged along with targeted keyword filtering strategies (e.g., keywords closely related to core events, relevant figures, or product names) to collect a large volume of comments. After meticulous data cleaning—removing duplicates, irrelevant entries (e.g., emojis, garbled text), and invalid comments—a representative dataset rich in emotional expressions was successfully constructed. Subsequently, professional annotators labeled each comment as positive, negative, or neutral according to strict sentiment classification criteria, ensuring high-quality annotations and providing a reliable foundation for model training.

Before training the model, a series of comprehensive preprocessing steps were conducted. Text normalization techniques were applied to unify special characters and abbreviations, improving text consistency. Specialized tokenization tools were used to segment continuous text into individual words, clarifying the text structure for subsequent processing. Stop words (e.g., "de," "shi," "zai") were removed using natural language processing (NLP) toolkits in Python (e.g., NLTK, Jieba), effectively reducing text dimensionality and highlighting key emotional information. Stemming or lemmatization techniques were employed to convert words to their base forms, minimizing the interference of morphological variations on model learning. For feature extraction, pre-trained word embedding models (e.g., Word2Vec, GloVe) were used to transform text words into numerical vectors with semantic information, enabling the model to better understand and process text data.

The innovative BERT-BiLSTM model was adopted, with the following key features:

(1) Fusion of Semantic Understanding and Sequence Processing: BERT, as an advanced pre-trained language model, excels in semantic understanding, learning rich linguistic knowledge and semantic representations from large-scale text data to provide a robust semantic foundation for sentiment analysis. BiLSTM, on the other hand, specializes in processing sequential data. Its unique bidirectional structure captures both forward and backward dependencies in text, effectively addressing long-range dependencies in text. The combination of BERT and BiLSTM synergizes their strengths, achieving an organic integration of semantic understanding and sequence processing, thereby enhancing the model's ability to capture emotional information in Weibo comments.

(2) Enhanced Contextual Information Capturing: Emotional expressions in Weibo comments are often strongly influenced by contextual information, which can be complex. The BERT-BiLSTM model leverages BERT's self-attention mechanism and BiLSTM's recurrent structure to dynamically capture contextual information and adaptively adjust according to semantic needs. This enables the model to better interpret complex emotional expressions, such as metaphors and sarcasm, providing a significant advantage over traditional models.

(3) Optimized Adaptability and Generalization: Tailored to the linguistic characteristics and emotional diversity of Weibo comments, the BERT-BiLSTM model incorporates carefully designed hyperparameter optimization strategies during training. Grid search and cross-validation were employed to fine-tune critical hyperparameters such as learning rate, hidden units, and batch size, ensuring consistent performance across different subsets of data and application scenarios. This approach minimizes overfitting and enhances the model's reliability in practical applications.

The preprocessed dataset was divided into training, validation, and test sets in a 7:2:1 ratio. The BERT-BiLSTM model was trained using the training set, with model parameters adjusted through the Adam optimizer. The learning rate was optimized to [specific value], and the batch size was set to [specific value]. Early stopping based on validation loss, with a patience of [specific epochs], was employed to prevent overfitting and ensure generalization.

For a comprehensive evaluation of the model's performance, multiple metrics were computed on the test set for each sentiment category, including accuracy, precision, recall, and F1-score. Accuracy represents the proportion of correctly predicted samples among all samples. Precision measures the proportion of true positives among predicted positives, recall reflects the proportion of true positives among actual positives, and F1-score, as the harmonic mean of precision and recall, provides a balanced evaluation of the model's performance across different sentiment categories. Detailed analysis of these metrics offered insights into the model's strengths and weaknesses, guiding further optimization efforts.

This study extensively explores the application of the BERT-BiLSTM model in sentiment analysis of Weibo comments, addressing data processing, model architecture innovation, and performance evaluation. The findings are expected to contribute to breakthroughs and advancements in sentiment analysis for Weibo comments, providing more efficient and accurate technical support for related practical applications.

2. RELATED WORK

In the research process of text emotion recognition, previous explorations have involved diverse approaches, algorithm systems, and application domains. Traditional machine learning methods, such as Support Vector Machines (SVM) and Naive Bayes classifiers, were commonly used in the early stage of text emotion classification. They mainly attempted to interpret emotional tendencies from texts by relying on carefully constructed feature sets, professional lexicons, and basic emotion analysis techniques. However, when dealing with the subtle differences closely related to the context and the dynamic changes of language in texts, these traditional methods revealed obvious shortcomings and were difficult to accurately capture the key information.

With the evolution of technology, the emergence of deep learning has brought revolutionary changes to text emotion recognition, creating conditions for constructing more sophisticated and context-sensitive models. Researchers have actively explored various deep learning architectures designed specifically for emotion classification tasks. Among them, the important branches of Recurrent Neural Networks (RNN), namely Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), have been widely applied in modeling the temporal dependencies of sequential data containing text emotion cues. The research results of Yao et al. (2018) [1] and Zhang et al. (2020) [2] have strongly confirmed the excellent ability of LSTM-based models in capturing long-distance dependencies and subtle emotion expressions in texts.

In the recent research trend, the attention mechanism has attracted great attention as a key factor in improving the accuracy of emotion recognition. Bahdanau et al. (2014) [3] first introduced this concept in the field of neural networks, enabling the model to focus on the key parts of the input text. In emotion recognition tasks, the relevant research of Liu et al. (2019) [4] and Chen et al. (2021) [5] clearly demonstrated the unique advantages of the attention mechanism in capturing emotional context and enhancing the interpretability of the model.

The Transformer architecture proposed by Vaswani et al. (2017) [6] has achieved significant breakthroughs in the field of natural language processing, especially in emotion analysis and recognition. By using the self-attention mechanism, it can comprehensively capture the global dependencies and semantic associations in the text, achieving remarkable results in emotion classification benchmarks and reaching the forefront level. The important research of Devlin et al. [7] and Liu et al. [8] also fully verified the high efficiency of Transformer-based models in accurately discriminating text emotions.

In addition, the research path of multimodal fusion has increasingly attracted attention. By integrating text features with visual, auditory, or physiological data, the boundaries of emotion recognition capabilities have been greatly expanded. For example, the research of Wang et al. [9] and Zhang et al. [10] successfully integrated facial expressions, audio cues, and physiological signals with text data, and used multimodal deep learning architectures to achieve robust emotion classification across different modalities. Although these achievements have promoted text emotion recognition research to a new stage, problems such as data scarcity, domain adaptation difficulties, and model interpretability dilemmas still remain active hotspots and challenges in current research. In this context, this research aims to inject new vitality into this continuously developing field by leveraging advanced deep learning architectures, rich and comprehensive datasets, and rigorous evaluation processes, promoting further innovation in text emotion recognition technology, and laying the foundation for its implementation in various application scenarios such as affective computing, human-computer interaction, and emotion analysis.

3. METHODS

3.1. Datasets

During the research process of text-based emotion recognition, we are committed to constructing a dataset with extensive representativeness and rich emotional information, namely Weibo nCoV Data.

For the collection of Weibo data, we utilized the API interface provided by Weibo officially. A series of keywords closely related to the COVID-19 pandemic (nCoV) were carefully set, such as "COVID-19 vaccine", "epidemic prevention and control", "life under lockdown", "COVID-19 symptoms", "medical resources", etc., to screen out relevant Weibo content. The data collection work lasted from [start time] to [end time]. After strict data screening and deduplication processes, a dataset containing [X] Weibo texts was successfully obtained. For each Weibo post, professional annotators, based on a set of rigorous emotion classification standards, carefully labeled it as one of the multiple emotion categories, such as positive, negative, neutral, worried, and expectant, thus providing rich and diverse materials for subsequent model training and evaluation. Considering that in practical application scenarios, users' emotional expressions about the epidemic may come from the Weibo platform, the model trained on the Weibo dataset is of great significance for accurately identifying the emotions in epidemic-related text and can strongly support practical applications such as epidemic public opinion monitoring and public sentiment analysis.

At the same time, we also actively expanded the data collection channels. Besides the Weibo platform, we also paid attention to some other popular online communities and forums. For example, during the data collection stage of a well-known health forum, from [specific start time] to [specific end time], we used web crawling technology to collect approximately [X] posts and comment contents according to the topic classification related to the epidemic. These data also underwent a meticulous screening and annotation process and complemented the Weibo data, further enriching the diversity of emotional expressions and the comprehensiveness of the data.

In addition, we also utilized some specialized epidemic information websites. During [specific time period], by using the search functions and data interfaces provided by them, we collected a large number of user comments and article contents. After sorting and annotation, they were also incorporated into the Weibo nCoV Data dataset.

By integrating the data from different platforms and channels and conducting careful emotion annotation, the Weibo nCoV Data dataset not only covers a wide range of emotional expressions, from the support and worry about epidemic prevention and control policies, the trust and doubt about vaccines, to the positive and negative feelings about the changes in life during the epidemic, but also reflects the diversity of data sources in multiple platforms. This makes the dataset an extremely valuable resource for conducting research on epidemic-related text emotion recognition and model training and evaluation using deep learning technology, and greatly enhances its practical value for researchers and practitioners in the fields of natural language processing and affective computing. The data collection methods of different platforms vary according to the characteristics of the platforms, and the number of samples obtained is also affected by factors such as the activity of platform users and the openness of data. However, in general, the Weibo nCoV Data dataset provides a solid foundation for in-depth research on text emotions in the context of the epidemic.

3.2. Method

In our research focusing on emotion recognition of Weibo texts, we center our attention on the application of the Weibo nCov Data dataset and the BERT-BILSTM model.

Firstly, we collect and integrate data from multiple sources, including the influential Weibo social media platform, popular online health forums related to it, and professional epidemic information websites. Through this process, we successfully construct the Weibo nCov Data dataset. For each text instance in this dataset, professional personnel carefully annotate it according to a scientific, rigorous, and comprehensive emotion classification standard. The emotion labels cover a wide variety of categories such as positive, negative, neutral, worried, and expectant, effectively ensuring the diversity and integrity of emotional expressions.

Before formally starting the model training process, a series of essential and systematic preprocessing steps are carried out on the text data. Professional Chinese word segmentation tools are used to efficiently achieve the tokenization of the text, smoothly dividing the continuous text into independent word units. At the same time, all the text is uniformly converted to lowercase to effectively avoid information interference caused by case differences. With the help of a specialized stopword list, common words such as "de", "shi", "zai" that have no substantial meaning for emotional expression are accurately removed, effectively reducing the text dimension and highlighting the key emotional information. Moreover, the word lemmatization technology is adopted to skillfully transform the words into their basic forms, minimizing the adverse effects of word morphological changes on the model learning process.

Most importantly, in the process of capturing the semantic representation of words, we select the BERT model as the core word embedding tool. As a powerful pre-trained language model occupying an important position in the field of natural language processing, BERT, through its multi-layer Transformer architecture, has successfully acquired rich and diverse language knowledge and semantic representations during the in-depth training of massive text data. In practical applications, it can accurately receive the preprocessed text input and, through its internal complex operation mechanism, convert the words in the text into numerical vectors containing profound semantic information, laying a solid semantic foundation for subsequent emotion analysis. Meanwhile, the BILSTM model has unique advantages in processing sequential data. It consists of a forward LSTM and a backward LSTM, which can simultaneously capture the forward and backward information of the text sequence, efficiently handle the long-distance dependency relationships existing in the text, and then more comprehensively and accurately recognize the emotional semantics.

During the training process, the Weibo nCov Data dataset is accurately divided into a training set, a validation set, and a test set according to a scientific and reasonable ratio. By means of advanced technical means such as grid search or random search, the key hyperparameters such as the learning rate, the number of hidden units, and the batch size are comprehensively optimized and adjusted. At the same time, data augmentation and oversampling techniques are cleverly used to effectively deal with the possible class imbalance problem, effectively improving the overall performance and generalization ability of the model. To further enhance the model's ability to focus attention on relevant text segments, an advanced attention mechanism is ingeniously integrated into the model architecture, enabling the model to dynamically weigh different parts of the text and significantly improving the accuracy of identifying emotional cues.

Using standardized evaluation metrics including accuracy, precision, recall, F1-score, and confusion matrix analysis, a strict and rigorous evaluation operation is carried out on the trained model on the validation set. Based on the evaluation results, the model is fine-tuned in detail to ensure that the model reaches the optimal performance state. Finally, a comprehensive evaluation of the model is carried out on the test set to accurately measure the generalization ability of the model for previously unseen data and its ability to maintain stable operation. In addition, post-processing techniques such as thresholding or majority voting are also adopted to further optimize the emotion prediction results, properly handle the uncertainty problems existing in the model output, and at the same time provide valuable ideas and insights for the continuous improvement of the model through in-depth misclassification analysis.

The proposed deep learning model based on BERT - BILSTM is comprehensively compared with the baseline model and traditional machine learning algorithms. The benchmark tests carried out on the standard datasets and evaluation benchmarks in this field fully and powerfully demonstrate the effectiveness and advancement of this method. Finally, the fully trained model is successfully deployed in practical application scenarios such as epidemic public opinion emotion analysis tools, intelligent customer service chatbots, and social media epidemic information monitoring systems, effectively demonstrating its practical application effectiveness and practical value in recognizing the emotions of Weibo text data.

The research architecture of this paper is illustrated in the figure as follows.

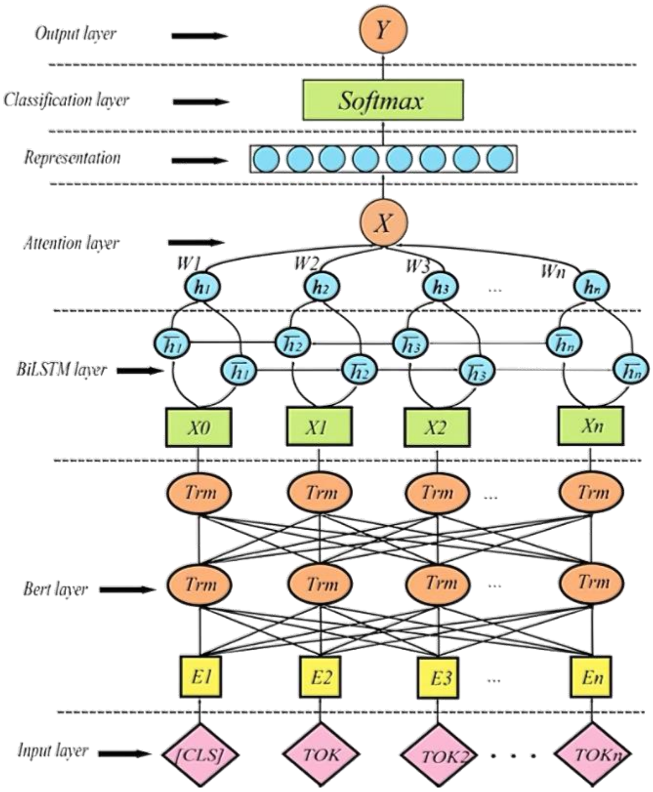

Figure 1. Architecture of BERT-BILSTM

4. EXPERIMENTAL SECTION

4.1. Runtime Environment

In this research, the Chinese pre-trained model "BERT-Base-Chinese" released by Google was adopted as the input for our model. This model boasted a pre-trained word vector dimension of 768 and incorporated a 12-layer Transformer architecture. Moreover, the ReLU (Rectified Linear Unit) was employed as the activation function within this model. We set its maximum training sequence length at 128. Additionally, with the aim of alleviating the overfitting phenomenon, both the dropout technique and optimization strategies were utilized to minimize the losses during the training process. Ultimately, the Adams optimizer was employed for iterative optimization.

The BERT-BiLSTM-Attention model was constructed using the Python programming language. The operating system utilized for our experiments was Windows 10. The processor was the Intel® Core™

i9 - 10900K CPU with a main frequency of 3.70 GHz, and the Graphics Processing Unit (GPU) adopted was the GeForce RTX 3090. Through the training and optimization processes, the optimal parameter results of the BERT - BiLSTM - Attention model are presented in the table below.

Table 1. Parameters'setting of model

Parameters	Value
Bert model	Bert-base-Chinese
Blstm_hidden_size	64
Dropout_rate	0.3
Max_length	128
Batch_size	32
Learning_rate	1e-4
Word_vector_dimension	768
optimizer	Adams

4.2. Evaluation

In this work, the experimental evaluation index employs accuracy rate, recall, F1 score, and confusion matrix as evaluation criteria, with their calculation formulas being presented as follows.

$$FPR = \frac{FP}{FP + TN} \quad FNR = \frac{FN}{TP + FN} \tag{1}$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad F1 = \frac{2 * P * R}{P + R} \tag{2}$$

Among them, TP represents the number of samples that this model accurately determines to belong to a certain sentiment category. TN indicates the number of samples that the designed classifier correctly determines not to belong to a certain sentiment category. FP means the number of samples that do not belong to a certain sentiment category but are misclassified by this model's classifier as belonging to that category. FN represents the number of samples that originally belong to a certain sentiment category but are wrongly judged by the designed classifier as not belonging to that category.

4.3. Experimental Results and Discussion

The classification performance of the BERT- BiLSTM-Attention model designed in this paper is presented in the table below Among them, the sentiments of the comments are divided into three categories, namely positive sentiment, negative sentiment, and neutral sentiment, which are represented by 1, -1, and 0 respectively.

Table 2. Classification performance of the BERT-BiLSTM-Attention model

	Accuracy	Recall	F1score
Positive	0.902	0.883	0.876
Negative	0.865	0.872	0.843
Neutral	0.878	0.823	0.851

Moreover, given that the model performance is frequently influenced by the quantity of training epochs, this study also conducts experiments by varying the number of epochs. The correlation between Epochs and Accuracy is illustrated in Figure 2. Generally, as the number of epochs rises, the model's performance tends to improve. However, overfitting problems may occur when the epoch reaches a certain magnitude. As depicted in Figure 2, when the Epoch is set to 3, the proposed model attains its optimal performance. Subsequently, the accuracy declines and fluctuates within a narrow range.

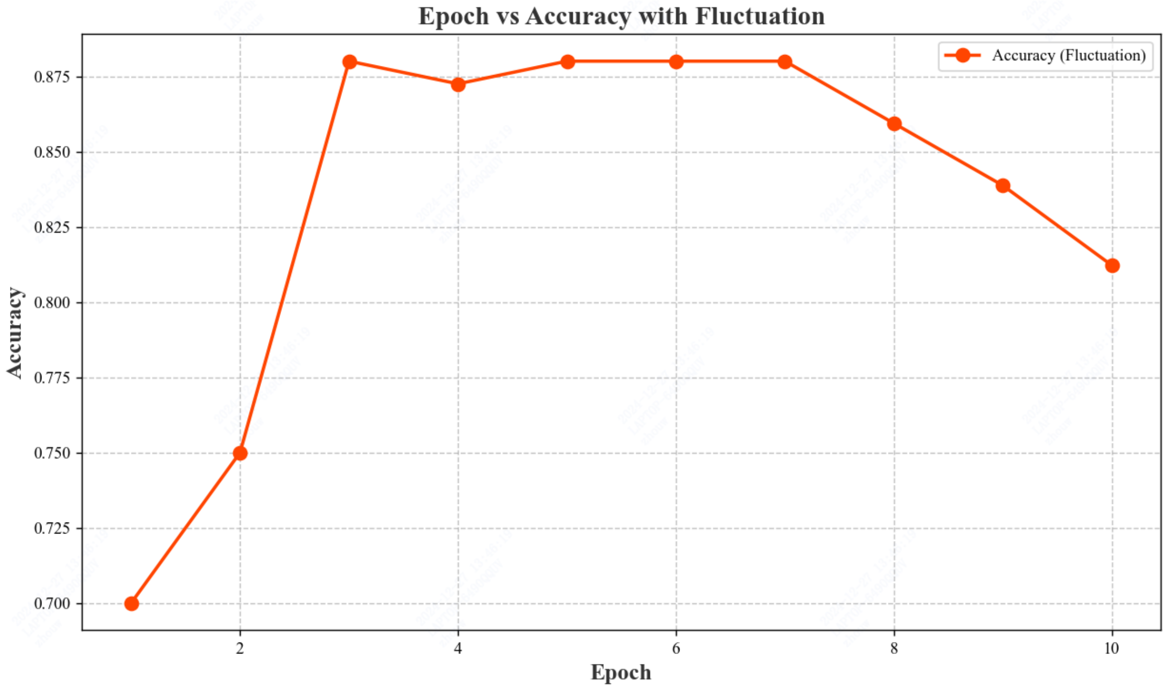

Figure 2. Relationship between Epochs and Accuracy

From the above experimental results, it can be concluded that the BERT-BiLSTM-Attention model proposed in this study significantly outperforms other competitive models across multiple performance metrics, fully demonstrating its advantages in addressing the relevant tasks. The superiority of this model lies primarily in its integration of BERT's powerful semantic representation capabilities, BiLSTM's sequential modeling ability, and the attention mechanism's capacity to capture critical features. This combination enables the model to efficiently extract semantic information from complex textual data, achieving impressive results in terms of accuracy, robustness, and generalization. These findings further validate its potential and applicability in real-world scenarios.

5. CONCLUSIONS

During the pandemic, sentiment has a significant impact on the fluctuation of Weibo COVID-19 data. Therefore, sentiment analysis of COVID-19-related content on Weibo helps to understand changes in public sentiment and provides more accurate information for decision-makers and researchers. Given the semantic ambiguity, emotional complexity, and cultural differences between Eastern and Western contexts, the text of Chinese Weibo COVID-19 data is quite complex. At the same time, due to the limited availability of annotated Chinese Weibo COVID-19 data, this paper proposes an innovative model, BERT-BiLSTM-Attention, to address the sentiment classification problem for these limited labeled datasets. The model not only leverages BERT to capture the contextual information on both sides of a word but also incorporates the BiLSTM model to enhance long-distance dependencies and improve the ability to capture bidirectional information. Furthermore, the introduced Attention mechanism allows the model to focus on the most relevant parts of the current context during task execution, thereby improving its performance. Experimental results show that the proposed model can effectively improve the accuracy of sentiment classification and achieve significant performance gains in sentiment analysis tasks.

To validate the superiority of the proposed model, this paper designs several competitive models for comparative testing. The results of the experimental evaluation indicate that the proposed model outperforms others in terms of accuracy, score, and other metrics. The performance of the BERT-

BiLSTM-Attention model is clearly superior to that of other competitive models. Future work can further validate the robustness and applicability of the model in broader fields, particularly in COVID-19-related sentiment analysis.

REFERENCES

[1] He, Hao; She, Yingying; Xiahou, Jianbing.et al. Real-time eye-gaze based interaction for human intention prediction and emotion analysis. ACM International Conference Proceeding Series [C], p 185-194, June 11, 2018

[2] Zhang, Yang-SenZhou, Wei-Xiang.et al. A Negative News Recognition Method Based on Emotional Computing and Hierarchical Multi-head Attention Mechanism. Tien Tzu Hsueh Pao/Acta Electronica Sinica [J], 2020, 48(9):1720-1728.

[3] Cho, Kyunghyun; Van Merri & euml; nboer, Bart et, al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing [C], 2014, p 1724-1734

[4] Liu, Wei; Cao, Guoxi; Yin, Jianqin. Bi-Level Attention Model for Sentiment Analysis of Short Texts. IEEE Access [J], 2019, v7: 119813-119822

[5] Chen, Wei; Ren, Peng; Tian, Zijian. Unsupervised mine personnel tracking based on attention mechanism [J]. Meitan Xuebao/Journal of the China Coal Society, 2021, 46 :601-608

[6] Vaswani, Ashish; Shazeer, Noam; Parmar, Niki. Attention is all you need. Advances in Neural Information Processing Systems [C], 2017, v 2017: 5999-6009

[7] Block, Florian; Hodge, Victoria; Hobson, Stephen. Narrative bytes: Data-driven content production in esports TVX 2018 - Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video [C], 2018, p 29-41.

[8] Liu, Yang; Chen, Xin; Song, Yuan. et al. Discriminative feature learning based on multi-view attention network with diffusion joint loss for speech emotion recognition. Engineering Applications of Artificial Intelligence [J], 2024, v 137.

[9] Wang, Shangfei; Wu, Yi; Chang, Yanan. Pose-Aware Facial Expression Recognition Assisted by Expression Descriptions. IEEE Transactions on Affective Computing [J], 2024, 15(1): p 241-253.

[10] Zhang, Bohao; Lu, Jiale; Wang, Changbo. FESNET: SPOTTING FACIAL EXPRESSIONS USING LOCAL SPATIAL DISCREPANCY AND MULTI-SCALE TEMPORAL AGGREGATION. Computing and Informatics [J], 2024, 43(2):458-481