applied sciences

MDPI

Article

A Method of Sentiment Analysis and Visualized Interaction Based on Ernie-Tiny and BiGRU

Yiqi Xiu [1,†], Xinlei Liu [1,†], Jingjing Qiu [1], Tangjun Qu [1], Juan Liu [1,2] and Yulong Bian [1,2,*]

[1] School of Mechanical, Electrical and Information Engineering, Shandong University, Jinan 264209, China; 202000800555@mail.sdu.edu.cn (Y.X.); qjj314095399@163.com (J.Q.); 202237571@mail.sdu.edu.cn (T.Q.); zzzliujuan@sdu.edu.cn (J.L.)
[2] Engineering Research Center of Digital and Media Technology, Ministry of Education, Jinan 250101, China
* Correspondence: bianyulong@sdu.edu.cn
† These authors contributed equally to this work.

Abstract: Sentiment analysis of subjective texts in social media is beneficial to help people adjust and intervene in a negative mental state in time, which is significant to mental health care. At present, limited by the accuracy of word segmentation, sentiment analysis of subjective text has difficulties in dealing with context, sentence patterns, and word co-occurrence. This paper aims to propose an efficient method of semantic feature representation and sentiment analysis, thereby providing a basis for sentiment visualization and interactive applications. Based on Ernie-Tiny and BiGRU, this paper proposes a sentiment analysis model ET_s_BG+p to solve problems in analyzing Chinese subjective texts' complex semantics, diverse sentence patterns, and shortness. The model inputs the semantic features obtained via Ernie-Tiny into BiGRU and then splices the output with the sentence vectors of Ernie-Tiny to form final text features and perform sentiment classification. Experiments are performed on a dataset integrating text comments from Weibo, takeaway, and e-commerce platforms. The results show that the model proposed in this paper performs best in most of the evaluation indicators compared with baseline models such as CNN, BiLSTM, and GRU. The experiments show that the accuracy of the model on the dataset built in this research is 84.30%, the precision is 83.95%, the recall rate is 88.35%, and the F1 value is 85.98%. At the same time, based on ET_s_BG+p, this paper develops a prototype visual display platform that integrates functions such as text input, sentiment analysis, and agent interaction, which can provide support for daily emotion monitoring and adjustment.

Keywords: sentiment analysis; deep learning; Ernie-Tiny; BiGRU; visualized interaction

check for updates

Citation: Xiu, Y.; Liu, X.; Qiu, J.; Qu, T.; Liu, J.; Bian, Y. A Method of Sentiment Analysis and Visualized Interaction Based on Ernie-Tiny and BiGRU. *Appl. Sci.* **2023**, *13*, 5961. https://doi.org/10.3390/app13105961

Academic Editor: Antonio Fernández-Caballero

Received: 19 April 2023
Revised: 9 May 2023
Accepted: 9 May 2023
Published: 12 May 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With the information boom dominated by user-generated Internet content, a large number of user comments have been generated on the Internet. These user comments are subjective texts with personal emotions and experiences. Much valuable information can be obtained through analyzing these comments. For example, merchants can master user satisfaction with the help of relevant product comments, and potential users can evaluate products through viewing product reviews. However, with the rapid growth of the number of comments, manual analysis is no longer advisable. Therefore, information technology is utilized to mine the sentiment tendency contained in the texts, and text-based sentiment analysis technology has come into being.

The implementation method of text-based sentiment analysis has gone through four stages: sentiment dictionary, machine learning, deep learning, and pre-trained model fine-tuning. International research on text sentiment analysis is more focused on English texts, while there are differences in sentence structure and expression between Chinese and English. Chinese semantics are more complex, influenced by many factors such as context

and syntax, so it is more difficult to extract semantic features and analyze sentiment for Chinese [1].

Among the commonly used text sentiment analysis methods, the sentiment dictionary method [2] is simple and fast. The typically used Chinese sentiment dictionaries include HowNet and the Chinese sentiment vocabulary ontology library of Dalian University of Technology. However, the process of constructing and updating dictionaries requires a lot of manpower and material resources. In addition, the method is separated from the context, unable to understand the different meanings of the same word, and the transferability is weak.

In terms of text sentiment analysis methods based on machine learning, the commonly used methods include support vector machines [3], Naive Bayes, and the maximum entropy model [4,5]. Zhao et al. [6] introduced multiple redundant markers and used conditional random fields to analyze text sentiment. The experimental results are better than the standard model; Chen et al. [7] proposed the XGBoost machine learning algorithm, using a large-scale parallel boosted tree, which is several times faster than common methods, and began to be used on text data [8]; HAN et al. [9] proposed the XGB-CNN model based on a convolutional neural network and XGBoost, and used XGBoost instead of Softmax for regression analysis, which greatly improved the classification accuracy. In general, machine learning is better due to mining text feature modeling, but the model suffers from the problem of sparse semantic features which seriously affects the model's effectiveness.

Today, text-based sentiment analysis methods are no longer limited to sentiment dictionaries and machine learning methods. Deep learning and fine-tuning methods based on pre-trained models bring new opportunities to sentiment analysis tasks [10]. Recent studies have shown that they have better generalization ability and stronger migration ability, and constantly refresh the optimal performance of Natural Language Processing (NLP) tasks [11], which can provide reliable technical support for sentiment analysis of subjective texts [12].

In recent years, deep learning methods have been more commonly used in Natural Language Processing. In sentiment analysis tasks, the mainstream methods of deep learning include onvolutional Neural Networks (CNNs) [13], Recurrent Neural Networks (RNNs) [14], the attention mechanism model [15,16], etc. The Long Short-Term Memory (LSTM) [17,18] network, proposed on the basis of a recurrent neural network, solves the problem of vanishing gradient and exploding gradient to some extent [19]. The Gate Recurrent Unit (GRU) [20] further simplifies LSTM and improves computational efficiency. Later, researchers proposed the Bi-directional Long Short-Term Memory (BiLSTM) [21] and Bi-directional Gated Recurrent Unit (BiGRU), which can relate contextual content and model sequence information in both directions. At the same time, hybrid neural networks complement each other's strengths and bring better results than using a basic model alone [22]. For example, Li et al. [23] used a model combining CNN and BiLSTM in the sentiment analysis of movie recommendations. CNN was used to extract local features, and BiLSTM was used to obtain the context; Basiri et al. [24] proposed a bi-directional CNN–RNN model based on the attention mechanism, which used an independent BiLSTM layer and a BiGRU layer to extract context, and obtained better evaluation results.

However, complete deep learning is time-consuming, and text representation is still dominated by word vectors, leading to difficulty in capturing implicit semantic features. The BERT pre-trained model proposed by Google [25], with a multi-layer bi-directional Transformer [26] as the main body, extracts contextual semantic features and highlights sentence focus using two pre-trained methods, the Masked Language Model (MLM) and Next Sentence Prediction (NSP). Downstream fine-tuning based on the BERT pre-trained model can be applied to new tasks and save a lot of retraining time. However, BERT's strategy of masking, with Chinese character granularity, obviously destroys Chinese entities and does not conform to the characteristics of Chinese text. In contrast, the Ernie model proposed by Baidu [27] learns syntactic and semantic information from these units implicitly through masking units such as phrases and entities. The Ernie model outperforms models such as BERT on various Chinese Natural Language Processing tasks. The lightweight Ernie

model Ernie-Tiny, proposed immediately after, is compressed from 12 layers to 3 layers, which significantly reduces model parameters, improves training speed, and provides good text analysis, making it more suitable for practical applications and development. In recent years, some researchers have been devoted to using Ernie's pre-trained model together with Bi-directional Gated Recurrent Unit (BiGRU) to perform Chinese text analysis. LEI et al. [28] used Ernie and BiGRU for Chinese text classification tasks. Cheng et al. [29] proposed an improved Chinese short text classification method based on the Ernie_BiGRU model. In terms of various evaluation metrics for the text classification task, the joint use of Ernie and BiGRU performed well.

However, traditional Chinese text analysis methods are limited by the accuracy of word separation and have difficulties in dealing with factors such as context, syntax, and word co-occurrence relationships, which makes Chinese text sentiment analysis less effective. Moreover, the pre-trained model has huge parameters and consumes many computational resources. Therefore, this study intends to propose a new subjective Chinese text sentiment analysis model to address the problems of insufficient semantic feature extraction and low model efficiency. Since text sentiment analysis technology is less applied in the field of mental health care, this study developed an interactive application prototype that visualizes text sentiment analysis results based on this model. The application prototype can help people understand their emotional state in daily life in a timely and intuitive manner, which is important for people's daily mental health. The main contributions of this paper include:

1. A subjective text sentiment analysis model ET_s_BG+p based on Ernie-Tiny and BiGRU is proposed, which can achieve efficient semantic feature representation and sentiment analysis for Chinese subjective texts with complex semantics, diverse sentence forms, and short texts.
2. Based on the ET_s_BG+p model, an interactive application prototype integrating text input, sentiment analysis, intelligent interaction, and other functions is developed to provide a reference for sentiment visualization and interactive application design.

2. Materials and Methods

2.1. ET_s_BG+p Model Development

2.1.1. Overall Design Idea of the Model

This paper proposes the ET_s_BG+p sentiment analysis model with the Ernie-Tiny pre-trained model. BERT's word-based masking lacks the ability of global information modeling in Chinese Natural Language Processing tasks. To solve this problem, this model improves the masking strategy. Training on massive data based on the improved masking strategy enables the model to learn the complex semantics hidden in the text and distinguish polysemy. As a variant of LSTM, GRU is simpler than the standard LSTM model, and runs faster with fewer tensor operations. So, the model in this paper further enriches emotional features via using BiGRU to link the context. Therefore, this paper expects that this combination can achieve better results.

2.1.2. Ernie-Tiny

As a miniaturized Ernie, Ernie-Tiny is reduced from a 12-layer to a 3-layer Transformer coding structure, with a speed-up of nearly 4.3 times, which is more in line with practical applications. The specific structure is shown in Figure 1.

Ernie-Tiny uses subword granularity to replace word granularity as input granularity. The specific input consists of a three-part overlay of Token Embedding, Segment Embedding, and Positional Embedding. In Figure 1, Chinese for "People's Republic of China" is input into Ernie-Tiny. Position Embedding encodes the position information of the input word into a feature vector. Segment Embedding divides the input words into two categories: A and B. Token Embedding Token Embedding tags input words and converts these tags into vectors. In terms of pre-trained method, Ernie's MLM is different from BERT. The model uses three masking strategies: basic-level masking randomly masks

basic Chinese characters, phrase-level masking randomly masks Chinese phrases, and entity-level masking randomly masks entities in sentences. Through predicting the mask positions, the model can learn external entity knowledge and semantic relations, which can lead to better understanding of Chinese text.

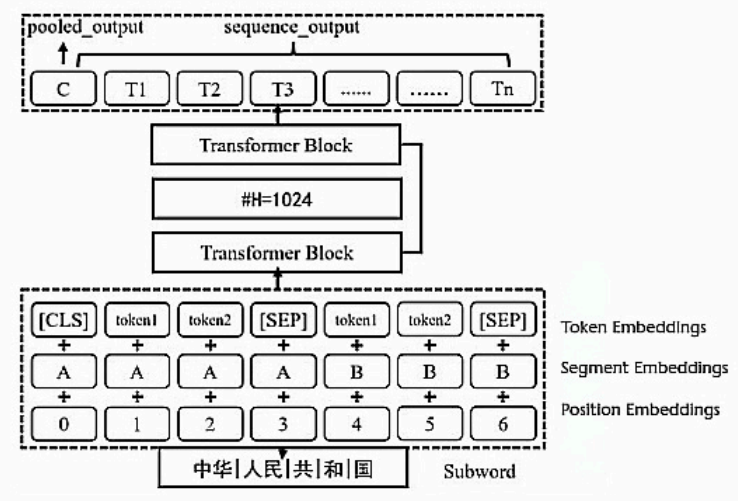


Figure 1. The structure of Ernie-Tiny.

Ernie-Tiny's Transformer encoder is shown in Figure 2, including its Self-Attention Mechanism, Add&Normalize Layer, and Feed Forward Neural Network. *X*1 and *X*2 in Figure 2 represent the word vectors of the two words that make up the sentence. The input word vector adds positional information to the Self-Attention Mechanism, calculates the dependency relationship between the current word and each word in the sentence, and then updates the weight of the word to capture the entire sentence information and highlight the sentence focus. As a sentence start token, [CLS] has no practical significance. After self-attention, it can represent the whole sentence feature most fairly, which is often used in classification tasks. Through three layers of Transformer encoding, Ernie-Tiny can finally obtain the sentence vector and sequence vector of the text. The sentence vector is the last layer [CLS] with dimensions of *batch_size* $\times$ 1024; the sequence vector is the vector representation of each token in the last sequence layer with dimensions of *batch_size* $\times$ *sequence_length* $\times$ 1024.

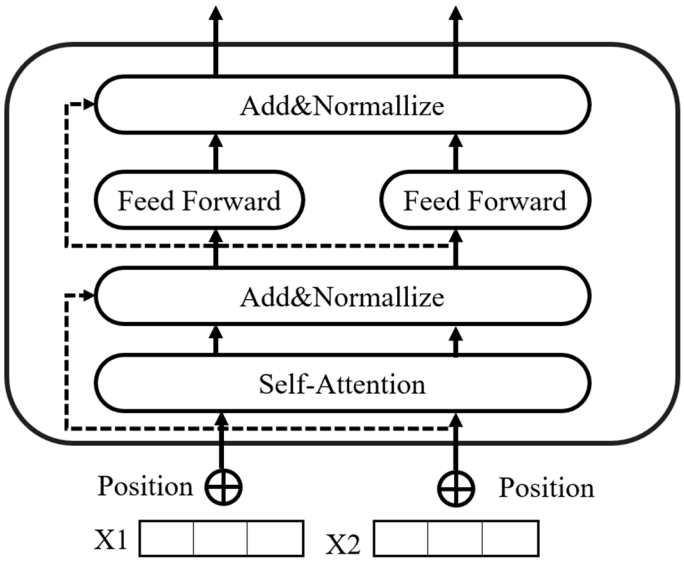


Figure 2. Transformer encoder structure.

2.1.3. The BiGRU

GRU is a variant of LSTM, which synthesizes the forget gate and the input gate of LSTM into the update gate. There is only an update gate and a reset gate in the GRU model, which is simpler and faster than the standard LSTM model.

The structure of the BiGRU model is shown in Figure 3. The BiGRU model consists of forward GRU and reverse GRU together. T is the input sequence vector. $\overrightarrow{h}_i$ is the vector of the output of the sequence i after the forward GRU calculation, which is jointly influenced by the input vector T_i and the previous moment $\overrightarrow{h}_{i-1}$. $\overleftarrow{h}_i$ is the vector output after the reverse GRU calculation, and the influence factor is the same as the forward GRU, which is calculated as follows:

$$\overrightarrow{h}_i = f\left(T_i, \overrightarrow{h}_{i-1}\right) \tag{1}$$

$$\overleftarrow{h}_i = f\left(T_i, \overleftarrow{h}_{i-1}\right) \tag{2}$$

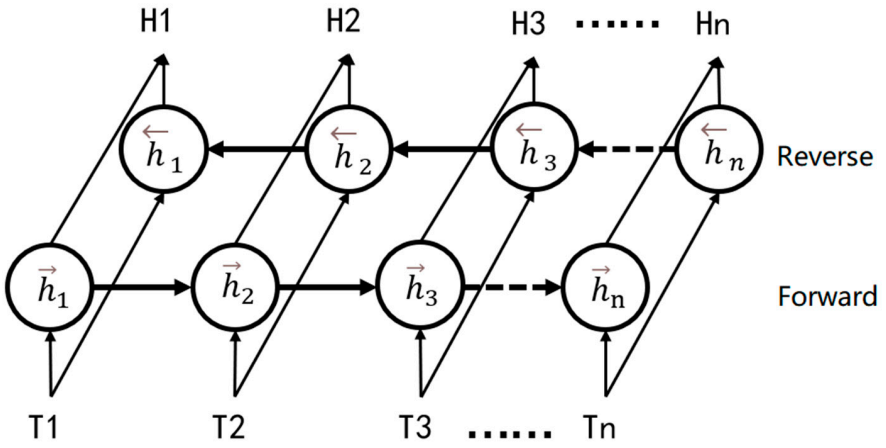


Figure 3. The structure of BiGRU.

The forward and reverse outputs are stitched together to obtain the output vector H of the BiGRU model.

2.1.4. ET_s_BG+p

The ET_s_BG+p model proposed in this paper is shown in Figure 4. E denotes the embedding vector generated from the input. It is mainly processed in two layers: The first layer is processed using Ernie-Tiny, whose pre-training process involves learning a large high-quality Chinese corpus, which is more adapted to Chinese entities, syntax, and semantics, and is conducive to Chinese subjective text sentiment analysis tasks after downstream fine-tuning; the sentence vector representing the semantics of the whole sentence and the sequence vector representing each token in the last layer are obtained via Ernie-Tiny. The second layer is processed with BiGRU, which can further mine the semantic relationships between above and below. The sentence vector output from Ernie-Tiny represents whole sentence semantics, and the sequence vector is processed using BiGRU to further extract text features; the information features of both can be spliced and fused to obtain better text representation. Therefore, the output of BiGRU is spliced with the sentence vector of Ernie-Tiny and finally passed into the output layer through the fully connected layer, and the final sentiment classification is performed using the Softmax function. The Softmax formula is as follows:

$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{c=1}^{C} e^{z_i}} \tag{3}$$

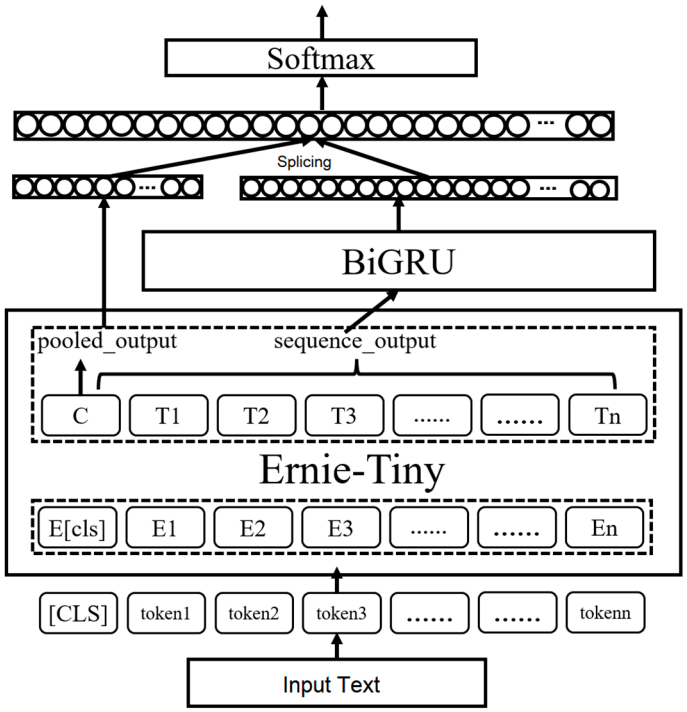


Figure 4. ET_s_BG+p structure.

z_i is the output value of the i node and C is the number of output nodes, i.e., the number of categories for classification. The Softmax function enables converting the output values of multiple categories into the likelihood of belonging to each category.

The loss function used in this model is the cross-entropy loss function, with the following equation:

$$\text{loss}(x, class) = -\log\left(\frac{\exp(x[calss])}{\sum_i \exp(x[i])}\right) = -x[class] + \log\left(\sum \exp(x[i])\right) \quad (4)$$

x is the output of the last layer of the network, and the shape is [*batch_size, class*]. *class* is the label passed in, i.e., the index value of the category.

2.2. Database

Considering the characteristics of subjective texts, this paper integrates four datasets—waimai_10k [30], online_shopping_10_cats [31], simplifyweibo_4_moods [32], and weibo _senti_100k [33]—from takeaway platforms, e-commerce platforms, and microblogs as the original data sources. The dataset waimai_10k is a corpus of Chinese take-out reviews. The dataset online_shopping_10_cats contains data from various e-commerce platforms, i.e., more than 60,000 comments covering 10 categories (books, tablets, mobile phones, fruits, shampoo, water heaters, Mengniu, clothes, computers, hotels). the datasets simplifyweibo_4_moods and weibo_senti_100k contain data from Weibo comments. In this paper, data pre-processing includes standardizing data format, de-duplication, filtering, length limiting to refine data, and traditional simplification steps. The final experimental dataset is an integration of these four datasets. Additionally, the original data labels are normalized into two categories through de-duplication and merging: 0 indicates negative emotions such as anger, disgust, fear, and sadness; 1 indicates positive emotions such as happiness, fondness, appreciation, and praise. The final dataset contains 160,296 positive data and 133,620 negative data, which are distributed as shown in Figure 5. The whole dataset was divided into a training set, validation set, and test set according to the ratio 8:1:1, as shown in Table 1.



Figure 5. Bipolar data distribution.

Table 1. Dataset division.

Dataset	Number
training set	235,132
validation set	29,392
test set	29,392

2.3. Experimental Environment and Parameter Settings

Considering factors such as the native platform and speed of the model, model training tests were mainly carried out on the Baidu intelligent cloud (AI Studio), using its Intel Xeon Gold 6148 CPU and Tesla v100 GPU. At the same time, in order to match the platform, the deep learning framework used Paddle 2.2.2 to build the model [34].

The Ernie-Tiny pre-trained model sets the learning rate using LinearDecayWithWarmup periodic decay function. The initial learning rate was set to 5. The parameter batch_size affects the speed of training and the number of iterations. The parameter batch_size was set at 100. Due to the fast convergence of the pre-trained model, this research finds that the optimum is reached at the 3rd to 4th round via comparing the results from epoch = 1 to epoch = 7. Since too many training rounds can easily lead to overfitting problems, the epoch was finally set at 3. The specific parameter settings are shown in Table 2.

Table 2. Parameter settings.

Parameter	Value
learning rate	5×10^5
batch_size	100
epoch	3
max_length	200
Optimizer	AdamW

2.4. Evaluation Criteria and Baseline Model

To compare the performance of each model, four evaluation metrics commonly used in text classification tasks are used in this paper, namely *Accuracy*, *Precision*, *Recall* and *F*1 (F-Score). Since the positive and negative portions of the dataset used in this paper are not balanced, there will be gaps in *Precision* and *Recall*. Therefore, *F*1 is used as the most significant metric to assess the predictive ability of the model. *F*1 is the harmonic mean of

Precision and *Recall*, and it can be used to reconcile *Precision* and *Recall* to make *Precision* and *Recall* compatible. The calculation formula of *F*1 is as follows:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (5)$$

To test the performance of the models proposed in this paper, some relevant baseline models are introduced for comparison experiments based on the evaluation criteria. The following are the models involved in this comparative experiment:

1. CNN: The model uses the pre-trained model of PKUseg [35], an open-source word separation tool of Peking University on the Weibo corpus, for text word separation after data pre-processing. Then, Tencent word vectors are used to construct the word embedding matrix, and each word is converted into a 200-dimensional word vector, which is input into the two-layer convolutional neural network. After the convolution layer and the pooling layer, the output layer is finally passed through the fully connected layer, and the Softmax function is used to complete the sentiment polarity two-classification task.
2. BiLSTM: After converting text data into word vectors according to the above method, the word vectors are input into the BiLSTM model. Finally, the output layer is passed through the fully connected layer, and the Softmax function is used for text sentiment classification.
3. GRU: Similar to the method used above, the model uses GRU for text sentiment classification.
4. ET_P: Sentence vectors representing the semantic features of whole sentences are obtained using Ernie-Tiny processing and passed into the output layer through the fully connected layer. The Softmax function is used for text sentiment classification.
5. ET_s: The model calculates the mean value of the output sequence vector of the last layer of Ernie-Tiny, and then transmits it to the output layer through the fully connected layer. The Softmax function is used for text sentiment classification.
6. ET_s_BiGRU: The sequence vector obtained through Ernie-Tiny processing is input into BiGRU, and then the output of BiGRU is passed into the output layer through the fully connected layer. Finally, text sentiment classification is performed using the Softmax function.
7. ET_s_BG+p: The model proposed in this paper.

3. Results and Discussion

3.1. Model Performance Evaluation

The specific results of the experiment are shown in Table 3. Due to the unbalanced dataset, positive texts are more easily recognized compared to negative texts, and the Recall is generally higher than the Precision. In the model without Ernie-Tiny, both BiLSTM and GRU outperformed the CNN model in all metrics, with the *F*1 of BiLSTM being 4% higher than that of CNN, which proves that considering contextual information can improve the model's textual emotion recognition ability. The overall effect of BiLSTM and GRU is not much different. Because the GRU network's complexity is lower than that of BiLSTM and the *F*1 of GRU is only 0.27% lower than that of BiLSTM, the GRU cost performance is relatively higher, so the model proposed in this study uses BiGRU. Compared with the Ernie-Tiny series models and other basic models, the Ernie-Tiny series models score relatively higher in each evaluation metric. The specific performance is that the Accuracy, *Precision*, *Recall* and *F*1 are increased by at least 4.02%, 3.28%, 0.65%, and 3.42%, respectively. This shows that the Ernie-Tiny pre-trained model can extract the semantic features of Chinese text more effectively than the collocation word vector. A comparison of each model's performance is shown in Figure 6.

Table 3. Comparison of experimental results.

Model Name	Accuracy (%)	Precision (%)	Recall (%)	F1(%)
CNN	75.26	76.11	79.62	77.83
BiLSTM	79.23	78.24	85.77	81.83
GRU	79.14	78.77	84.54	81.56
ET_p	83.70	82.05	**89.74**	85.72
ET_s	83.25	82.05	88.70	85.25
ET_s_BiGRU	83.94	**84.48**	86.42	85.44
ET_s_BG+p	**84.30**	83.95	88.35	**85.98**

Note: Bold indicates the optimal value.

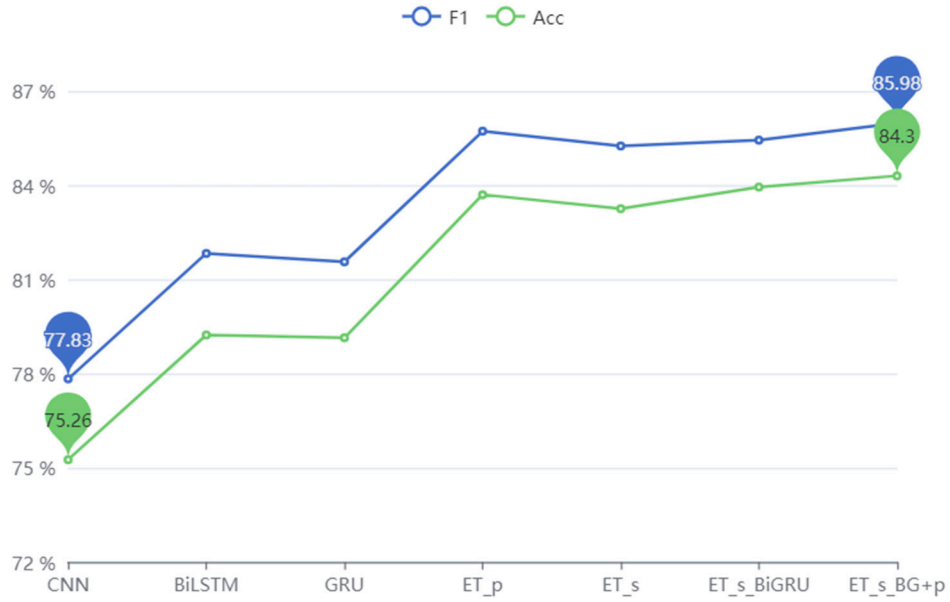


Figure 6. Model performance comparison line chart.

Comparison among the models using Ernie-Tiny: ET_p is slightly better compared to ET_s, with an *F*1 value improvement of 0.47%, indicating that for the dataset used in this paper, the sentence vector representing the semantic features of whole sentences in Ernie-Tiny is more representative of the overall features of the text than the averaged sequence vector and is more suitable for classification. The ET_s_BiGRU model has the highest accuracy, the smallest gap between accuracy and recall, and the best balance, with 0.69% improvement in accuracy and 0.19% improvement in *F*1 value compared to ET_s, which proves the benefit of using BiGRU. ET_s_BG+p achieves the highest optimal accuracy and *F*1 value in the whole experiment, integrating the advantages of ET_p and ET_s_BiGRU, and proves the value and significance of the model proposed in this paper. The model effect difference comparison is shown in Figure 7.

In addition, comparing the running time of each model, the running time of the Ernie-Tiny series models is shorter than that of other basic models, which shows that the pre-trained model is efficient. The running time within the Ernie-Tiny series of models is positively correlated with the complexity of the model, but there is no significant difference.

3.2. Root Cause Analysis

In Natural Language Processing (NLP) tasks, Convolutional Neural Networks (CNNs) can effectively extract local features of text vectors. However, due to the affinity dependence of natural language on structure, text classification based on CNN will ignore the contextual meaning of words, which limits the accuracy of text classification [36]. A Recurrent Neural Network is an ideal choice for processing dynamic input sequences that are common in NLP. For serialized input, Recurrent Neural Networks (RNNs) can effectively integrate

adjacent position information [20,37] and process various tasks in NLP. Both Bi-directional Long Short-Term Memory (BiLSTM) and the Gate Recurrent Unit (GRU) are variants of RNN, so BiLSTM and GRU perform better than the CNN model in the experiment.

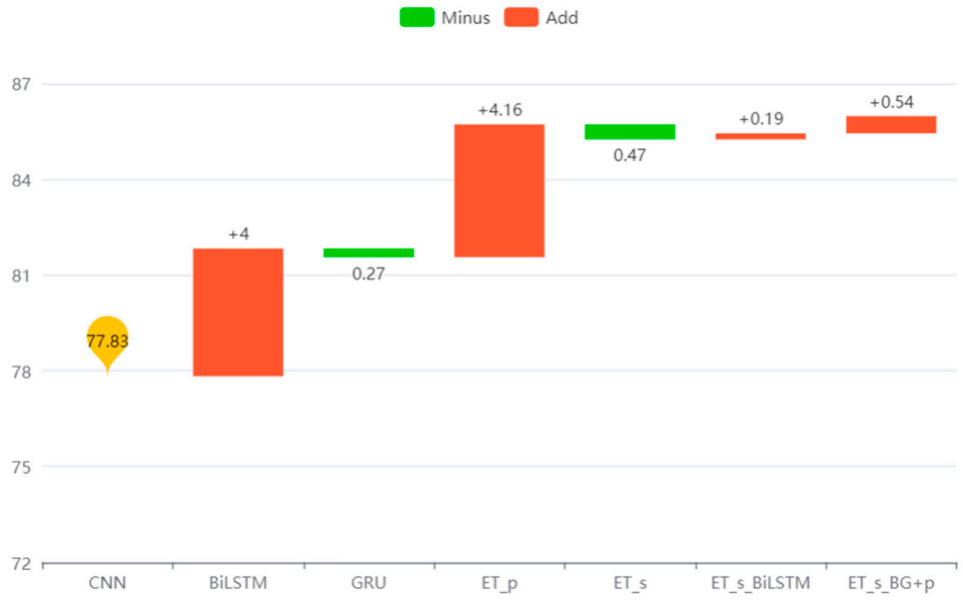


Figure 7. Model *F*1 difference comparison.

At present, pre-trained models have achieved the most advanced results in various NLP tasks. In general, pre-trained language models are trained on a large amount of text data in a self-supervised manner, and then fine-turned during downstream tasks or directly deployed through zero/few-shot learning without task-specific fine-tuning. Such pre-trained language models have become the new paradigm for Natural Language Processing tasks. The pre-trained language model Ernie [27] has proved to be effective for improving the performance of sentiment classification tasks [38]. Therefore, the performance of each evaluation metric of the models using Ernie-Tiny is better than other basic models.

3.3. Interactive Application Prototyping

3.3.1. Platform Framework Design

The visualization platform framework designed for this study is divided into four parts as follows:

1. The algorithm model part is based on the ET_s_BG+p sentiment analysis model, which calculates the sentiment value of the front-end input text and stores it in the database.
2. The data storage part is implemented using Python's own lightweight database SQLite.
3. The interaction interface part uses the Flask framework to connect the front and back ends to achieve the interaction between data and the user interface.
4. The front-end application part uses HTML, CSS, JavaScript, and other technologies to build direct interaction functions with users.

After the user enters text on the platform, ET_s_BG+p can perform sentiment analysis on the text and store the resulting sentiment index in the database. Then, the visualization tool can generate a visual chart through calling the sentiment index. The specific architecture of this visualization platform system is shown in Figure 8.

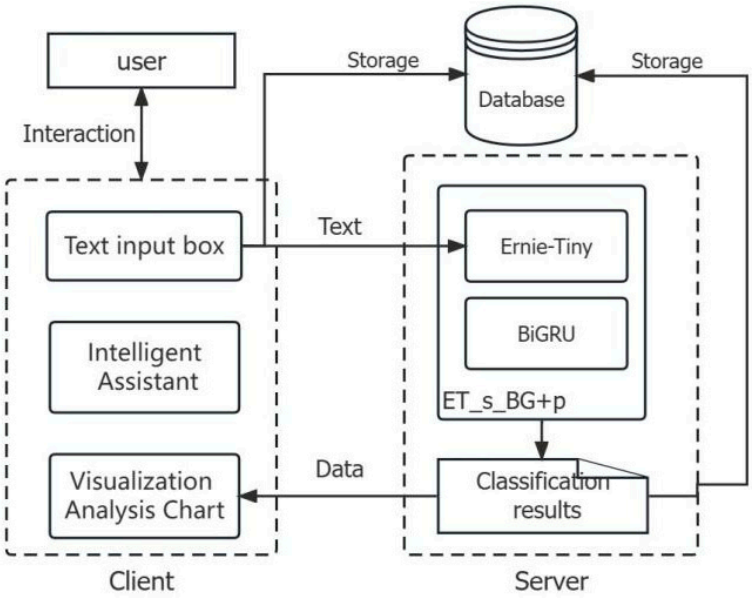


Figure 8. System architecture.

3.3.2. Interaction Function Design

To improve the usefulness of the visualization platform and optimize user experience, three main interaction functions are designed in this study, namely, text manipulation, visual sentiment analysis graphs, and an emotional feedback adjustment method based on an intelligent assistant.

1. Text operation

The application is designed with a simple text operation interface, as shown in Figure 9. After the user writes text in the input box and clicks the "Finish" key, the application will generate sentiment data through sentiment analysis and store the sentiment data in the database. At the same time, the application provides the functions of viewing history, deleting and modifying, etc. The home page only shows the text records of the last seven entries. If the user needs to view historical data, they can query by month or day. Only the first 100 words of each record are displayed, hiding the rest of the text. After the user clicks the "View Details" key, the application will pop up a modal box to display the text details, and the user can delete or modify the content in this modal box.

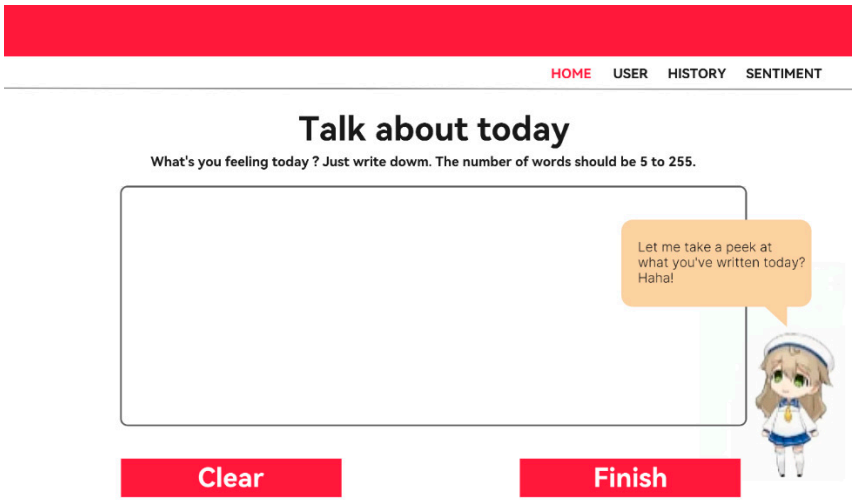


Figure 9. The input interface.

2. Visual sentiment analysis

The application uses ECharts to design sentiment analysis charts and provides three visual charts, as shown in Figure 10. Mood polarity graphs show the probability of emotional polarity for each text, and the range of the displayed curves can be adjusted through the timeline. Mood distribution pie charts show the distribution of the number of positive and negative diaries. The mood swing graphs show the ratio of positive diaries to the total number of diaries for a month, which better reflects the user's mood direction.

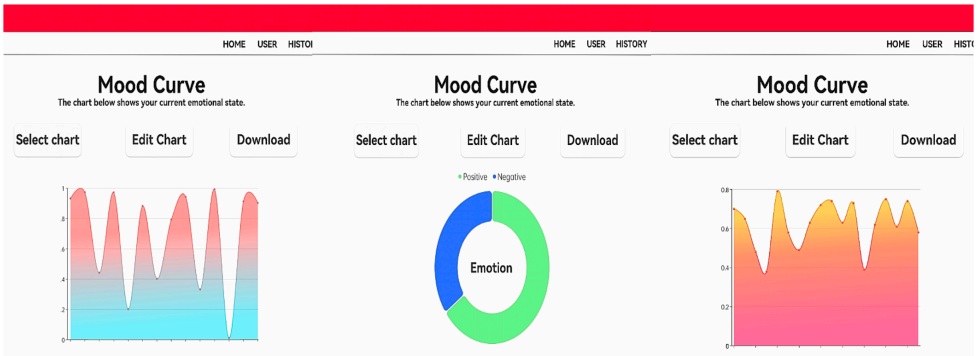


Figure 10. Visual Sentiment Charts.

3.4. Value Evaluation and Improvement Direction of the Application Prototype

The application prototype was developed based on the model proposed in this study to demonstrate the potential of the model in practical applications. The prototype is able to identify the user's emotional state based on subjective text input and visualize the emotions in various graphical forms, which can help the user understand their emotional state in real time and adjust their emotions in time. The application prototype shows the application potential of text sentiment analysis technology in the field of mental health care.

At present, the function of this application prototype is relatively simple, and the user's emotional state derived from text sentiment analysis is not detailed enough. In the future, the model will be improved so that the application will have more detailed emotion recognition capabilities. At the same time, the form of emotion visualization can be richer, for example, using animation to show the user's emotional state. In addition, the application will be augmented with a real-time comment function, which will help users to regulate emotions through giving users corresponding suggestions based on their emotional state.

4. Conclusions

This paper proposes the ET_s_BG+p model for Chinese text sentiment analysis, and compares the model with other baseline models based on the integrated dataset. The experimental results show that the model outperforms other baseline models in most performance metrics and scores 85.98 in the most important evaluation metric, F1. In addition, an interactive application prototype for text sentiment visualization is developed based on this model. Since the dataset used in this study has only two sentiment labels, further refinement of the dataset sentiment category labels will be performed in subsequent studies. In addition, this model needs to continue to be improved and enhanced in its ability to generalize unbalanced data.

Author Contributions: Conceptualization and methodology, all authors; validation, Y.X. and X.L.; analysis, Y.X. and X.L.; writing—original draft preparation, Y.X. and X.L.; supervision, X.L.; writing—review and editing, Y.X., X.L., J.Q., T.Q., J.L. and Y.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Ministry of Education of Humanities and Social Science project (22C10422009) and the Young Scholars Program of Shandong University, Weihai (20820211005).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Peng, H.; Cambria, E.; Hussain, A. A review of sentiment analysis research in Chinese language. *Cogn. Comput.* **2017**, *9*, 423–435. [CrossRef]
2. Ahmad, M.; Aftab, S.; Muhammad, S.; Waheed, U. Tools and techniques for lexicon driven sentiment analysis: A review. *Int. J. Multidiscip. Sci. Eng.* **2017**, *8*, 17–23.
3. Vapnik, V. SVM method of estimating density, conditional probability, and conditional density. In Proceedings of the IEEE International Symposium on Circuits and Systems, Geneva, Switzerland, 28–31 May 2000; pp. 749–752.
4. Berger, A.L.; Pietra, V.J.D.; Pietra, S.A.D. A maximum entropy approach to natural language processing. *Comput. Linguist.* **1996**, *22*, 39–71.
5. Pang, B.; Lee, L.; Vaithyanathan, S. Thumbs up: Sentiment classification using machine learning techniques. *arXiv* **2002**, arXiv:cs/0205070 [cs.CL]. [CrossRef]
6. Zhao, J.; Liu, K.; Wang, G. Adding redundant features for CRFs-based sentence sentiment classification. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, Honolulu, HI, USA, 25 October 2008; pp. 117–126.
7. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery Data Mining, San Francisco, CA, USA, 13–17 August 2016.
8. Seyfioglu, M.S.; Demirezen, M.U. A hierarchical approach for sentiment analysis and categorization of Turkish written customer relationship management data. In Proceedings of the 2017 Federated Conference on Computer Science and Information Systems, Prague, NY, USA, 3–6 September 2017; pp. 361–365.
9. Han, T.; Yin, W.; Fang, M. Sentiment Analysis Model Based on Convolutional Neural Network and XGBoost. *J. Fudan Univ.* **2019**, *58*, 560–564. (In Chinese)
10. Wang, N.-Y.; Ye, Y.-X.; Liu, L.; Feng, L.-Z.; Bao, T.; Peng, T. Language Models Based on Deep Learning: A Review. *J. Softw.* **2021**, *32*, 1082–1115. (In Chinese) [CrossRef]
11. Wang, Q.; Li, B.; Xiao, T.; Zhu, J.; Li, C.; Wong, D.F.; Chao, L.S. Learning deep transformer models for machine translation. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 5 June 2019; pp. 1810–1822.
12. Qiu, X.P.; Sun, T.X.; Xu, Y.G.; Shao, Y.; Dai, N.; Huang, X. Pre-trained models for natural language processing: A survey. *Sci. China Technol. Sci.* **2020**, *10*, 1872–1897. [CrossRef]
13. Kim, Y. Convolutional neural networks for sentence classification. *arXiv* **2015**, arXiv:1510.03820.
14. Irsoy, O.; Cardie, C. Opinion mining with deep recurrent neural networks. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 25–29 October 2014; Association for Computational Linguistics: Stroudsburg, PA, USA, 2014; pp. 720–728.
15. Mnih, V.; Heess, N.; Graves, A. Recurrent models of visual attention. In Proceedings of the 27th International Conference on Neural Information Processing Systems—Volume 2 (NIPS'14), Montreal, QC, Canada, 8–13 December 2014; MIT Press: Cambridge, MA, USA, 2014; pp. 2204–2212.
16. Lin, L.; Li, J.; Zhang, R.; Yu, W.; Sun, C. Opinion mining and sentiment analysis in social networks: A retweeting structure-aware approach. In Proceedings of the 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing, London, UK, 8–11 December 2014; pp. 890–895.
17. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]
18. Palangi, H.; Deng, L.; Shen, Y.; Gao, J.; He, X.; Chen, J.; Song, X.; Ward, R. Deep sentence embedding using long short-term memory networks: Analysis and application to information retrieval. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2016**, *24*, 694–707. [CrossRef]
19. Bengio, Y.; Simard, P.; Frasconi, P. Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. Neural Netw.* **1994**, *5*, 157–166. [CrossRef] [PubMed]
20. Cho, K.; Van Merrienboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078. [CrossRef]
21. Graves, A.; Schmidhuber, J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw.* **2005**, *18*, 602–610. [CrossRef] [PubMed]
22. Minaee, S.; Kalchbrenner, N.; Cambria, E.; Nikzad, N.; Chenaghlu, M.; Gao, J. Deep Learning Based Text Classification: A Comprehensive Review. *arXiv* **2004**, arXiv:2004.03705. [CrossRef]
23. Li, S.; Yan, Z.; Wu, X.; Li, A.; Zhou, B. A method of emotional analysis of movie based on convolution neural network and bi-directional LSTM RNN. In Proceedings of the IEEE Second International Conference on Data Science in Cyberspace, Shenzhen, China, 26–29 June 2017; pp. 156–161.

24. Ehsan Basiri, M.; Nemati, S.; Abdar, M.; Cambria, E.; Acharrya, U.R. ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis. *Future Gener. Comput. Syst.* **2020**, *115*, 279–294. [CrossRef]
25. Devlin, J.; Chang, M.; Lee, K.; Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Minneapolis, MI, USA, 2–7 June 2019; pp. 4171–4186.
26. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, K.; Polosukhin, I. Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17), Long Beach, CA, USA, 4–9 December 2017; Curran Associates Inc.: Red Hook, NY, USA, 2017; pp. 6000–6010.
27. Sun, Y.; Wang, S.H.; Li, Y.K.; Feng, S.; Chen, X.; Zhang, H.; Tian, X.; Wu, H. ERNIE: Enhanced representation through knowledge integration. *arXiv* **2019**, arXiv:1904.09223.
28. Lei, J.; Qian, Y. Chinese-test Classification Method Based on ERNIE-BiGRU. *J. Shanghai Univ. Electr. Power* **2020**, *36*, 329–335+350. (In Chinese)
29. Cheng, X.; Zhang, C.; Li, Q. Improved Chinese Short Text Classification Method Based on ERNIE_BiGRU Model. *J. Phys. Conf. Ser.* **2021**, *1993*, 012038. [CrossRef]
30. Dataset: Waimai_10k. Available online: https://github.com/SophonPlus/ChineseNlpCorpus/blob/master/datasets/waimai_10k/intro.ipynb (accessed on 3 September 2014).
31. Dataset: Online_Shopping_10_Cats. Available online: https://github.com/SophonPlus/ChineseNlpCorpus/blob/master/datasets/online_shopping_10_cats/intro.ipynb (accessed on 3 September 2014).
32. Dataset: Simplifyweibo_4_Moods. Available online: https://github.com/SophonPlus/ChineseNlpCorpus/blob/master/datasets/simplifyweibo_4_moods/intro.ipynb (accessed on 3 September 2014).
33. Dataset: Weibo_Senti_100k. Available online: https://github.com/SophonPlus/ChineseNlpCorpus/blob/master/datasets/weibo_senti_100k/intro.ipynb (accessed on 3 September 2014).
34. Yu, K. Large-scale deep learning at Baidu. In Proceedings of the 22nd ACM International Conference on Information & Knowledge Management, Burlingame, CA, USA, 27 October–1 November 2013; pp. 2211–2212.
35. Luo, R.; Xu, J.; Zhang, Y.; Ren, X.; Sun, X. PKUSEG: A Toolkit for Multi-Domain Chinese Word Segmentation. *arXiv* **2019**, arXiv:1906.11455.
36. Li, Y.; Dong, H. Text sentiment analysis based on feature fusion of convolution neural network and bidirectional long short-term memory network. *J. Comput. Appl.* **2018**, *38*, 30753080. (In Chinese)
37. Ebrahimi, J.; Dou, D. Chain Based RNN for Relation Classification [EB/OL]. Available online: http://www.aclweb.org/anthology/N/N15/N15-1133.pdf (accessed on 3 September 2014).
38. Socher, R.; Perelygin, A.; Wu, J.; Chuang, J.; Manning, C.D.; Ng, A.; Potts, C. Recursive deep models for semantic compositionality over a sentiment treebank. In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, Seattle, WA, USA, 28 March 2013; pp. 1631–1642.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.