

## Article

# Improving Sentiment Classification of Restaurant Reviews with Attention-Based Bi-GRU Neural Network

Liangqiang Li, Liang Yang \* and Yuyang Zeng

Business and Tourism School, Sichuan Agricultural University, Chengdu 611830, China; lilq@sicau.edu.cn (L.L.); 2019228010@stu.sicau.edu.cn (Y.Z.)

\* Correspondence: yl0119@outlook.com or 2019228018@stu.sicau.edu.cn

**Abstract:** In the era of Web 2.0, there is a huge amount of user-generated content, but the huge amount of unstructured data makes it difficult for merchants to provide personalized services and for users to extract information efficiently, so it is necessary to perform sentiment analysis for restaurant reviews. The significant advantage of Bi-GRU is the guaranteed symmetry of the hidden layer weight update, to take into account the context in online restaurant reviews and to obtain better results with fewer parameters, so we combined Word2vec, Bi-GRU, and Attention method to build a sentiment analysis model for online restaurant reviews. Restaurant reviews from Dianping.com were used to train and validate the model. With F1-score greater than 89%, we can conclude that the comprehensive performance of the Word2vec+Bi-GRU+Attention sentiment analysis model is better than the commonly used sentiment analysis models. We applied deep learning methods to review sentiment analysis in online food ordering platforms to improve the performance of sentiment analysis in the restaurant review domain.



**Citation:** Li, L.; Yang, L.; Zeng, Y. Improving Sentiment Classification of Restaurant Reviews with Attention-Based Bi-GRU Neural Network. *Symmetry* **2021**, *13*, 1517. <https://doi.org/10.3390/sym13081517>

Academic Editors: Juan Luis García Guirao, Waldemar Karwowski, Boback Parsaei and Hamid R. Parsaei

Received: 24 July 2021

Accepted: 11 August 2021

Published: 18 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 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/>).

**Keywords:** online restaurant reviews; Bi-GRU; sentiment analysis; attention

## 1. Introduction

The widespread adoption of Web 2.0 has provided an environment for consumers to engage in expression, creativity, communication, and sharing. Consumers are able to post reviews on online ordering platforms (e.g., Yelp, TripAdvisor, Dianping.) in order to express their opinions about restaurants, vent their emotions, and engage in social activities. Merchants often encourage consumers to actively participate in reviews, and massive user-generated restaurant reviews give consumers the opportunity to fully express their needs while helping merchants provide real-time and personalized service [1,2]. According to a 2019 BrightLocal survey, approximately two-thirds of consumers have posted reviews of local establishments, with an average of nine reviews per person per year [3]. Due to the intangible and complex nature of goods and services in the restaurant industry, consumers rely heavily on reviews from other customers to evaluate service quality before spending money [4]. Restaurant reviews express the composition of consumers' emotional needs and are an important source of information that consumers can refer to [5]. In the pre-consumer information search phase, consumers tend to search for a large number of restaurant reviews from other users to reduce the perceived uncertainty and perceived risk caused by information asymmetry [6].

Due to a large amount of unstructured information available on the Web, collecting as well as aggregating product review information is a challenging task, which requires the use of automated methods to help researchers collect as well as analyze data, and many previous studies have used sentiment analysis to mine consumer attitudes [7]. The object of sentiment analysis can be in the form of speech, text, images, etc. Restaurant reviews are usually presented as text, so the sentiment analysis in most of papers focuses on text-based sentiment analysis [8]. Consumers usually form a general perception of a restaurant by reading existing restaurant reviews in the pre-purchase information-seeking

stage, and the huge amount of restaurant review information obviously exceeds consumers' information processing ability, and reading fewer reviews has a higher probability of generating misperceptions [9]. This requires the platform to provide an efficient way of processing to quickly identify the emotional information contained in restaurant reviews.

There are two main categories of current classification methods. The first is the method based on sentiment lexicon, which mainly judges the sentiment tendency of a text based on the number of sentiment words appearing in the text; the other method is based on machine learning, including Support Vector Machine, Naïve Bayes, K nearest neighbor algorithm, etc. [10–13]. The limitations of previous studies are found through comparison: (1) lexicon-based, machine learning methods rely on accurate sentiment dictionaries and data preprocessing, and traditional word characterization methods do not take into account contextual information, making sentiment analysis less effective [14]; (2) online ordering platform reviews have strong domain characteristics, such as “Service”, “Comfortable”, “Enjoyable” and other words, and online ordering platform reviews contain many expressions and meaningless words. Research using sentiment dictionaries or semantic knowledge bases relies on language-specific external resources, and this approach has poor applicability in the face of different domains. It is difficult to consider the full range of specialized vocabulary using traditional sentiment analysis methods.

To efficiently and accurately identify the sentiment in restaurant reviews, we fully consider the advantages of Word2vec and Bi-directional Gated Recurrent Unit (Bi-GRU), and add attention mechanism in neural network. First, we preprocessed online restaurant reviews. Secondly, the distributed word vector representation method, Word2vec, is used to train word vectors. Finally, a restaurant review sentiment classifier was constructed using Bi-GRU. This paper contributes from the following two aspects.

- We used Word2vec for word vector representation and attention mechanism in Bi-GRU for sentiment analysis, which improves the efficiency of sentiment analysis;
- We took full advantage of Bi-GRU's symmetric update to apply it to online restaurant review sentiment analysis, considering the contextual dependencies in online restaurant reviews.

The rest of the paper is organized as follows. Section 2 lists related work on restaurant review sentiment and sentiment analysis methods. Section 3 includes the research framework of this paper and the algorithms. Section 4 provides the detailed steps to construct a sentiment classifier and shows the results of experiments. Section 5 elaborates the conclusion. Section 6 discusses the limitations of this paper and future works.

## 2. Literature Review

In this paper, we combined attention mechanism and Bi-GRU for sentiment analysis of reviews on online ordering platforms. In this section, we introduce online restaurant reviews and the related works about sentiment analysis methods.

### 2.1. Online Restaurant Reviews

Consumers usually consider restaurant reviews when making restaurant selection decisions because they complement other information provided by merchants, such as restaurant descriptions, expert opinions, and personalized needs generated by automated recommendation systems [15]. Consumers who read restaurant reviews will rely on their previous experiences to perceive the attitudes expressed in the reviews, and by continuously reading restaurant reviews consumers will form an overall perception of the store and eventually influence their purchase behavior. Restaurant review sentiment reflects the general perceptions and attitudes of other consumers about the restaurant, and consumers often decide to go to a reputable restaurant after searching for online restaurant reviews [4].

Most existing studies examine consumer psychology and behavior in terms of online restaurant reviews, and some of the most relevant research publications in the field are listed in Table 1.

**Table 1.** Literature about online restaurant reviews.

Study	Author	Theme	Method	Conclusion
[16]	Nakayama et al., (2019)	Heterogeneity of customer dining experience.	Quantitative statistics on the distribution of user comment sentiment	There may be significant differences in user behavior between countries on cross-cultural social commerce platforms.
[17]	Jurafsky et al., (2014)	Explore the narratives consumers use to frame positive and negative emotions online.	Text mining	Negative reviews are more likely to reflect characteristics associated with trauma narratives, while positive reviews are more likely to use long pieces of narrative to emphasize the linguistic capital of the reviewer.
[18]	Jia. (2020)	Discover and compare the motivation and satisfaction of restaurant visitor customers from different cultural backgrounds.	Probabilistic theme model	Chinese tourists are less inclined to downgrade restaurants and more intensely fascinated by the food on offer, while American tourists are more inclined to seek out fun and less uncomfortable with crowds.
[19]	Meek et al., (2021)	What normative and informative features of Online restaurant reviews affect the perceived usefulness of the ratings indicated by “likes”.	Content analysis	Heuristics can provide important filters for leads.
[20]	Tian et al., (2021)	Measuring consumer sentiment related to food and emotional responses from online review data.	Dictionary-based sentiment analysis, empirical analysis	Consumers used more positive words than negative sentiment in their reviews, more emotional words were used when discussing restaurant service than food.
[21]	Li et al., (2021)	How trial time distance affects review evaluation consistency.	Empirical analysis.	Significant moderating effects of empirical value and type of review equipment on the relationship between temporal distance and review consistency.

Consumer emotional expression is prevalent in online reviews and other forms of computer-mediated communication [22]. Some scholars mined emotional information from online restaurant reviews to provide practice guidance. Luo and Xu applied a deep learning approach to analyze aspect restaurant sentiment during the COVID-19 pandemic period and found that the deep learning model achieved better results overall compared to machine learning algorithms [23]. Micu et al., used Naïve Bayes to classify the sentiment of restaurant reviews, which helps marketers to grasp the characteristics and interests of consumers [24]. Some scholars have studied the methodological perspective of sentiment analysis of restaurant reviews. Kim et al. used word co-occurrence method to calculate the co-occurrence frequency of words in sentences and assigned the highest scoring implicit features to the sentences, while the author introduced a threshold parameter to filter potential features with low scores, and the results showed that this threshold-based approach has good performance for sentiment analysis [25]. Li et al., calculated the sentiment intensity

of online reviews using a text mining method and through empirical analysis they found that positive emotions had a negative impact on reviews, while negative emotions had a positive impact, in addition expressing angry emotions was more useful than expressing positive emotions [26]. Krishna et al., used machine learning methods to perform sentiment analysis on online restaurant reviews, and SVM achieves optimal results based on a specific data set [27].

Although many studies have paid attention to analyzing online restaurant reviews sentiment to help merchants on the platform to improve their services, there are still some questions: (1) Can the accuracy of a restaurant review sentiment classifier be further improved? (2) Does the method and efficiency of sentiment analysis of online restaurant reviews in Chinese differ from other languages due to the more ambiguous expressions in Chinese?

## 2.2. Sentiment Analysis Method

Sentiment analysis, also known as opinion mining, is a computational study of people's needs, attitudes, and emotions toward an entity [28]. Sentiment analysis is able to obtain the positive or negative sentiments of evaluation subjects and their intensity, and the results of sentiment analysis can be useful in many fields, such as online sentiment opinion analysis, topic monitoring, word-of-mouth evaluation of massive products, and so on.

Feature selection is a fundamental task in the field of sentiment analysis, and effective feature selection from subjective texts can significantly improve the efficiency of sentiment analysis [29,30]. Many scholars have conducted research from the feature perspective to find an effective feature selection method. Zhang et al., selected N-char-grams and N-POS-grams as potential sentiment features and used Boolean weighting method to calculate feature weights, and the results showed that the feature characterization method they chose was able to obtain better accuracy [30]. Hogenboom et al., used a vectorized representation based on text structure for multi-domain English text sentiment analysis, and the conclusion showed that this method works better than word-based feature representation [31]. Sentiment analysis is more domain-sensitive and product feature selection can be seen as the identification of domain-specific named entities, which leads to the fact that most sentiment analysis methods require domain-specific knowledge to improve the performance of the system. Most of the existing studies on feature selection have limitations, and the efficiency of sentiment analysis decreases significantly once it is removed from a specific domain.

Many studies used sentiment dictionaries as well as machine learning methods to analyze restaurant reviews [12,32], and although relatively good results have been achieved, the data processing effort is relatively high and the domain is less transferable. Meanwhile, deep learning-based sentiment analysis methods are gaining popularity as deep learning provides automatic feature extraction as well as richer representation performance and better performance [33]. Abdi et al., proposed a deep learning-based approach to classify user opinions expressed in reviews (called RNSA), which overcomes the disadvantages of traditional methods that lose temporal as well as positional information and achieves good results in sentence-level sentiment classification [34]. Al-Smadi used Long Short Term Memory (LSTM) to achieve sentiment analysis of reviews of Arabian hotels in two ways, first, by combining Bi-directional Long Short Term Memory (Bi-LSTM) and conditional random fields for the formulation of opinion requirements classification, and second, sentiment analysis using LSTM, which showed that both outperformed the previous baseline study [35].

In the field of sentiment analysis, many scholars have used methods based on sentiment dictionaries or traditional machine learning. The results of these methods are not satisfactory, as the performance of the model is heavily dependent on the feature selection strategy and the tuning of the parameters. Deep learning includes Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), and other network structures [29]. Deep learning-based sentiment analysis models use neural network to learn to extract complex features from data with minimal external contributions, and it has achieved good performance in natural language processing [36]. Compared



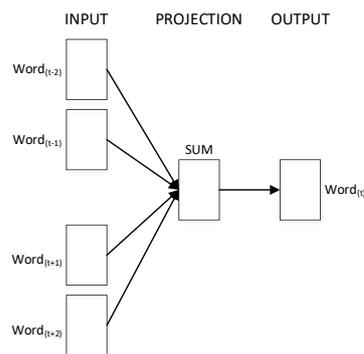


Figure 2. CBOW Framework.

### 3.2. Bi-GRU

In this article, the restaurant review sentiment classifier was constructed using the Bi-dimensional Gated Recurrent Unit (Bi-GRU) approach. Next, Recurrent Neural Network (RNN), Gated Recurrent Unit (GRU), and Bi-GRU are briefly introduced.

RNN uses a feedback loop where the output of each step is fed back into the Recurrent Neural Network therefore influencing the next output, a process that is repeated in each subsequent step. Such a feedback mechanism allows Recurrent Neural Network to dynamically learn sequence features and thus improve the efficiency of sentiment analysis. The computational equation is as follows:

$$s_t = f(Ux_t + Ws_{t-1}) \tag{1}$$

$$o_t = g(Vs_t) \tag{2}$$

where  $s_t$  denotes the value of the hidden layer,  $f, g$  denotes the activation function,  $U$  denotes the weights of  $x_t$ ,  $W$  denotes the weights matrix of  $s_{t-1}$ ,  $V$  denotes the weights matrix of the hidden layer.

Chung et al., proposed a GRU model with similar experimental results to LSTM, but with a simpler structure and more efficient computational process [38]. Like the input-output structure of RNN, GRU is influenced by the current input  $x^t$  and the hidden state  $x^t$  passed from the previous node. The Rate Recurrent Unit solves the gradient explosion problem in a simpler structure by introducing reset gate  $r$  and update gate  $z$ , as shown in Figure 3.

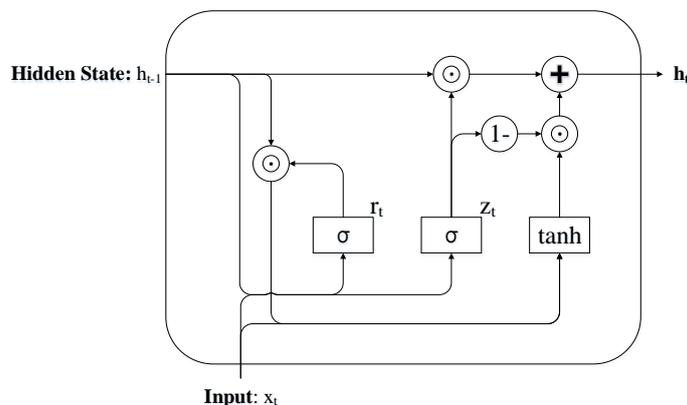


Figure 3. Construction of GRU.

First, the input from the current node and the state transmitted down from the previous node are used to obtain the reset as well as update the gating state, which is calculated as follows:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \tag{3}$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (4)$$

Secondly, after obtaining the gate signal use reset gate to record the current moment state, the specific calculation formula is as follows:

$$\tilde{h}_t = \tanh(W_{\tilde{h}} \cdot [r_t * h_{t-1}, x_t]) \quad (5)$$

The last step is to update the memory. The specific calculation formula is as follows:

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (6)$$

where  $t$  denotes a certain moment,  $\sigma$  denotes the activation function,  $W$  denotes the weight,  $r_t$  denotes the reset gate  $t$  at the moment,  $z_t$  denotes the update gate  $t$  at the moment, and  $h_t$  denotes the activation state at the moment  $t$ .

While Bi-GRU allows the hidden layer to capture historical and future contextual information, Bi-GRU takes into account both preceding and following sentence dependencies on top of GRU, which is usually applied in text classification tasks. Bi-GRU's operation mechanism is shown in Figure 4. At each step, the same weight matrix is multiplied with the input or the hidden layer at the previous time point and the processing has symmetry. This symmetry ensures that the neural network can fully take into account the context and ultimately improves the classification of the model.

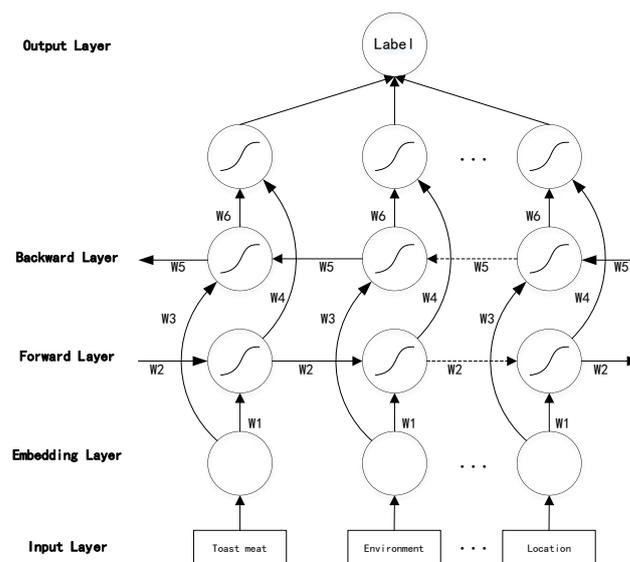


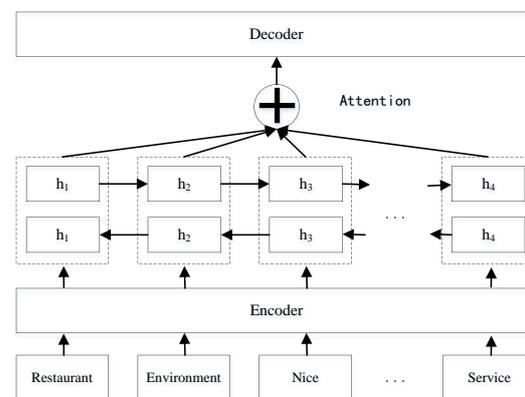
Figure 4. Operation Mechanism of Bi-GRU.

### 3.3. Attention Mechanism

The attention mechanism was originally derived from the human visual attention mechanism and was later applied to the field of artificial intelligence [39]. The attention mechanism is a simple method to encode sequential data based on the importance score assigned to each unit. As an information resource allocation scheme, it is widely used in various information streamlining tasks [40]. Deep learning models based on the attention mechanism can capture global and local connections flexibly, making the model less complex and with fewer parameters, improving the efficiency of model training.

Specifically, the attention mechanism assigns different weights to the input in the model, which can quickly extract the key information from the data to improve the robustness of the results. For example, if the input words of the sentiment classification model are "Restaurant", "Environment", and "Nice", the attention mechanism will take the word probability distribution of 0.2, 0.3, and 0.5 into account in the output of the model. The attention mechanism takes the word probability distributions 0.2, 0.3, and 0.5 into account

in the output of the model, which ultimately improves the quality of the sentiment analysis. The model after the introduction of the attention mechanism is shown in Figure 5.



**Figure 5.** Construction of Attention Mechanism.

The underlying form of the attention mechanism is shown below:

$$\begin{aligned} e_i &= a(u, v_i) \\ \alpha_i &= \frac{e_i}{\sum_i e_i} \\ c &= \sum_i \alpha_i v_i \end{aligned} \quad (7)$$

where  $u$  is the matching feature vector based on the current task for interaction with the context.  $v_i$  is the feature vector for a time stamp in the time series,  $e_i$  is initial attention score without normalization,  $\alpha_i$  is the attention score after normalization operation, and  $c$  is the contextual feature for the current time stamp, and it can be calculated by the summation of the attention score multiplied by the feature vector  $v$ .

In this paper, we used Bi-GRU to analyze the sentiment of reviews in online food ordering platforms, taking into account the pretext features as well as post text features to improve the accuracy of the results. During the training process, we used dropout to randomly remove neurons in the hidden layer to prevent overfitting and make the model more generalizable, and used softmax in the output layer to map the results to the range of 0~1. Finally, binary cross-entropy was used as the loss function with the following equation:

$$L = \frac{1}{N} \sum_i L_i = \frac{1}{N} \sum_i -[y_i \cdot \log(p_i) + (1 - y_i) \cdot \log(1 - p_i)] \quad (8)$$

where  $y_i$  denotes the label of sample  $i$ , and  $p_i$  denotes the probability that the sample is predicted to be positive.

### 3.4. Model Evaluation Metrics

Confusion matrices are commonly used in the task of two-classification supervised learning to determine the gap between predicted and true values, in the form shown in Table 2 [41]. A single confusion matrix metric is difficult to measure the merit of the model. Therefore, Precision, Recall, and F1-Score were used as the evaluation metrics for model performance in the research setting of this paper.

**Table 2.** Confusion Matrix.

Actual Class	Predicted Class	
	Positive Class	Negative Class
Positive class	True Positive (TP)	False Negative (FN)
Negative class	False Positive (FP)	True Negative (TN)

The calculation of Precision is shown below:

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

The calculation of Recall is shown below:

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

The F1-Score is commonly used in statistics to measure the performance of a dichotomous model, which takes into account both accuracy and recall, and the F1-Score is calculated as follows:

$$F_1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (11)$$

## 4. Experimental Results

### 4.1. Data Description

The experimental data in this paper comes from Dianping.com (accessed on 1 May 2021), which is now the leading local lifestyle consumption platform in China. We randomly crawled a total of 35,248 reviews from 130 stores by crawlers, which contain information on fields such as username, taste rating, environment rating, service rating, review content, and review time. An example of reviews is shown in Figure 6. The field of online reviews usually considers the textual sentiment of online reviews to be consistent with the digital review ratings [42]. In this paper, the average of taste ratings, environment ratings, and service ratings was taken as the composite score and judged, with positive sentiment polarity if the rating is greater than 3 and negative sentiment polarity if the rating is less than or equal to 3. Finally, 26,703 positive sentiment reviews and 8545 negative sentiment reviews were obtained. The descriptive statistics of the review data are shown in Table 3. The distribution of the ratings ranged from 0.5 to 5, the sentiment polarity of the reviews ranged from 0 to 1, and the length of the review text ranged from 5 to 2093. The length distribution of positive and negative reviews is shown in Figure 7.

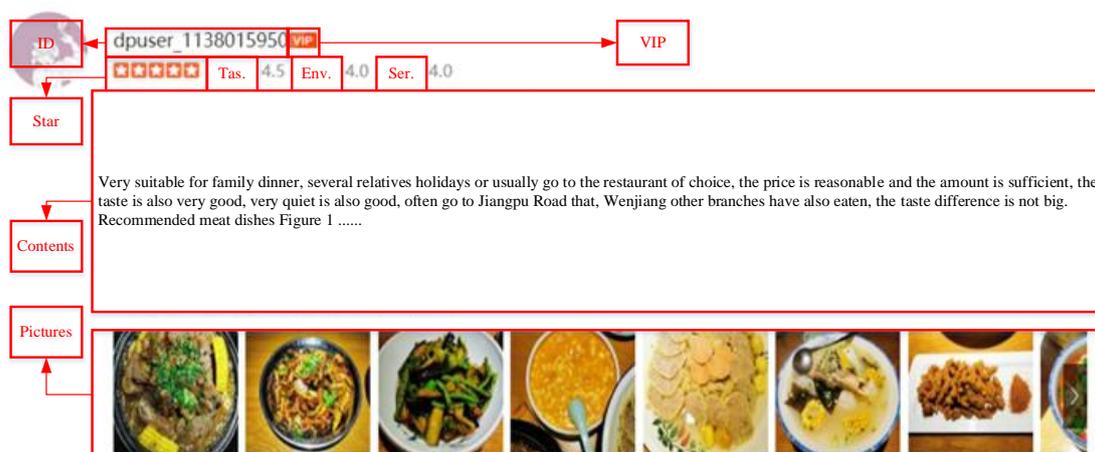
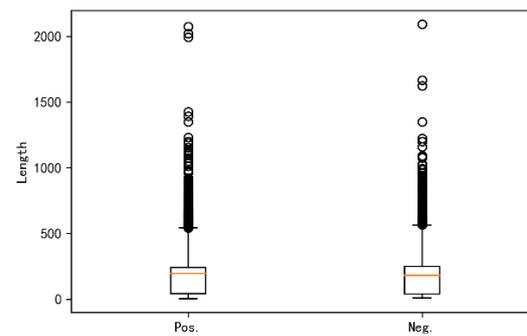


Figure 6. Review Score and Review Screenshot.

**Table 3.** Descriptive Statistics.

	Count	Mean	Std	Min	50%	Max
Taste	35,248	4.03	1.12	0.5	4	5
Environment	35,248	3.96	1.13	0.5	4	5
Service	35,248	3.95	1.23	0.5	4	5
Polarity	35,248	0.76	0.43	0	1	1
Len	35,248	175.61	136.75	5	195	2093

**Figure 7.** Length Distribution of Positive and Negative Reviews.

The Chinese restaurant review data were processed by word separation using the jieba library in Python language, and the meaningless words were removed using the HIT stop words list [43]. The reviews after word separation are shown in Table 4.

**Table 4.** Results of Word Segmentation (Example).

Number	Raw Reviews	Reviews after Word Segmentation
1	Today at lunch time went to eat the king shrimp, the overall feeling is still good. Shrimp is a large large, generally larger than the shrimp outside a circle. The service is also relatively fast, ordered 5 min on the up. then the price is expensive, three people ate 400 yuan ~ and in addition to shrimp nothing else to eat. Would have liked to match some snacks ah, porridge ah, cold dishes ah, brine ah ..... through no. Only the barbecue, noon is not supplied. Overall still good.	Lunch/eat/king/shrimp/overall/feel/good/shrimp/large/large/outside/shrimp/a circle/service/order/minute/price/expensive/three/eat/400/yuan/shrimp/eat/would have/want/match/match/snacks/porridge/cold dish/brine/through/barbecue/noon/supply/good
2	The taste can only be said to be so-so, will not take a special detour to eat, soup from beginning to end have not been added, the waiter just started to see after never mind our table, eat to the middle to find even the dipping saucer are not! Three people set menu dishes are also too shabby, shrimp and meatballs such as meat are counted by the number, drinks cannot be changed outside the set menu of other drinks, I personally feel that every aspect of even the small red robe is not comparable, hope to improve.	taste/can only/say/special/detour/eat/soup/from begin to end/have not been added/waiter/never/mind/table/eat/find/dip/saucer/set/dish/too/shabby/shrimp/meatball/meat/number/count/drinks/change/set/outside/drinks/every aspect/even small/red robe/not comparable/improve

#### 4.2. Experimental Setup

In this paper, the word vector was trained using the Gensim library, a third-party library in Python, where the size of window was set to 5, the dimensionality of the word vector was set to 300 dimensions, the learning rate was set to 0.01, and the rest of the parameters used the default initial settings. The dimensionality of the lexicon was reduced

to two dimensions using principal component analysis and visualized as shown in Figure 8. The distribution of restaurant review length is shown in Figure 9. By the statistics of review word length, we found that more than 90% of the review length is below 90, so we constructed the sentence vector embedding matrix with length 90, and the values in the matrix are the corresponding word indexes.

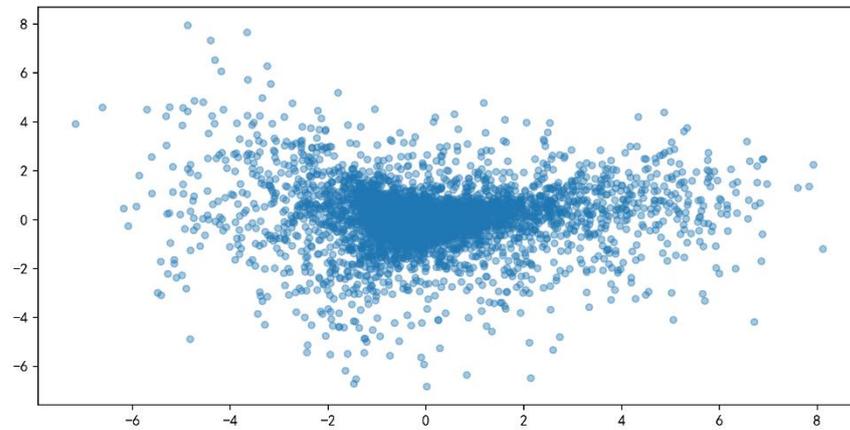


Figure 8. Word Vector Visualization.

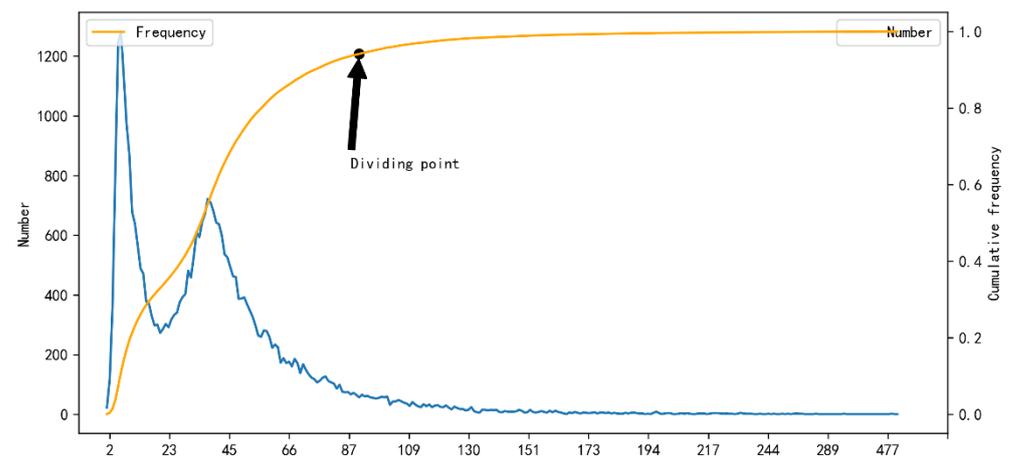


Figure 9. Review Length Distribution.

The Tensorflow deep learning framework was used to build the Word2vec+Bi-GRU+Attention deep learning model, and the training and test sets were divided into a 4:1 ratio. Dropout parameter allows the deep neural network model to ignore certain features during the training process to reduce the overfitting problem. To verify the impact of the dropout parameter on the model performance, we tested the accuracy of the model when dropout is 0.1~0.9, as shown in Figure 10, we can find that the accuracy of the model is highest when dropout is 0.2. Figure 11 gives a comparison of the model performance under three types of batch\_size settings, and the best when the batch\_size is 128 was considered comprehensively. To improve the adaptability of the model training process in different subsets, we set a certain proportion of the validation set. Figure 12 gives a comparison of the model performance under the three types of validation\_split parameter settings, the proportion of the validation set in the cross-validation is set to 0.4 and the model has a better performance. The detailed settings of each parameter in the neural network model are given in Table 5.

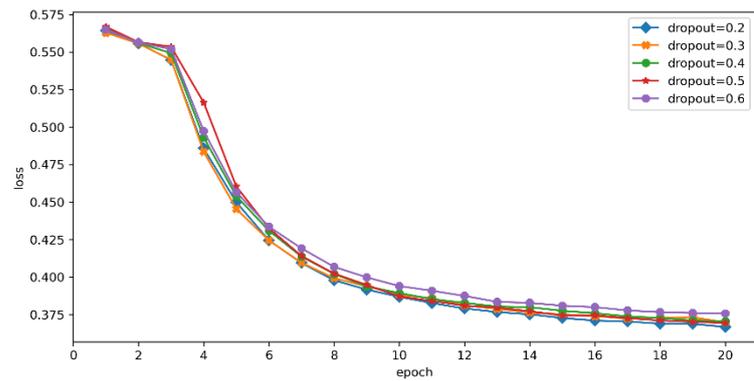


Figure 10. Impact of Dropout Parameters on Model Performance.

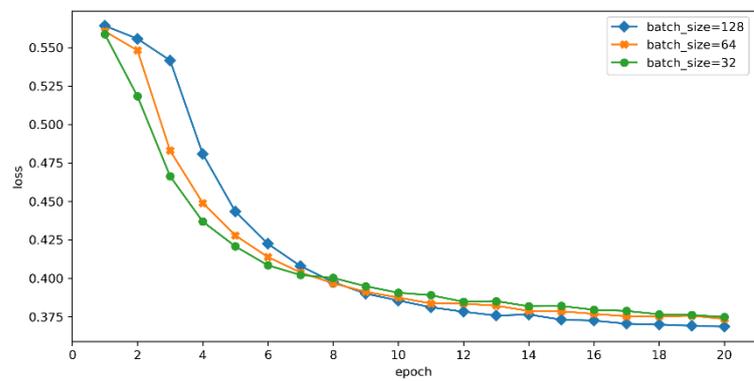


Figure 11. Impact of Batch\_size Parameters on Model Performance.

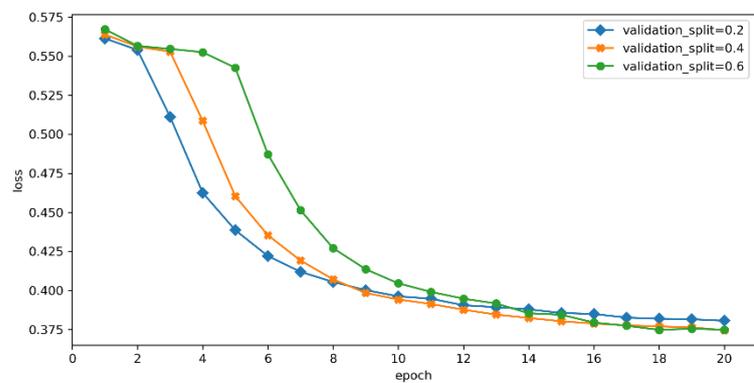


Figure 12. Impact of Validation\_split Parameters on Model Performance.

Table 5. Sentiment Analysis Neural Network Parameter Settings.

Parameter Setting	
Epoches	20
Batch_size	128
Hidden_size	50
Dropout	0.2
Loss	Crossentropy
Optimizer	Adam
Validation_split	0.4

#### 4.3. Baseline Model

To verify the validity of the sentiment classification model proposed in this paper, machine learning and deep learning methods were applied to the scenario of sentiment analysis of online ordering platform reviews, respectively. The baseline models we will apply are briefly introduced:

- K Nearest Neighbor (KNN). KNN is a classification algorithm whose basic principle is to compute the K values that are most similar to the centroid.
- Support Vector Machine (SVM). SVM can map the sample space to a functional space with high dimensionality by nonlinear mapping, converting an originally non-linearly separable problem into a linearly separable problem inside some feature space [44]. It has been proven to have good performance in sentiment analysis as well as efficiency.
- Convolutional Neural Network (CNN). Convolutional neural network can effectively consider information from different location sources, and they are widely used to solve problems such as image processing, natural language processing, including sentiment analysis, summary extraction, etc. [45].
- Bi-directional Long Short Term Memory (Bi-LSTM). Bi-LSTM fully considers context dependency and achieves good results in sentiment analysis [46].

#### 4.4. Experimental Results

Finally, their accuracy, recall, and F1 values were compared, and the results of the comparison are shown in Figure 13 and Table 6. It was found that the combined performance of the sentiment analysis model of W2v+Attention+Bi-GRU was better than the other models.

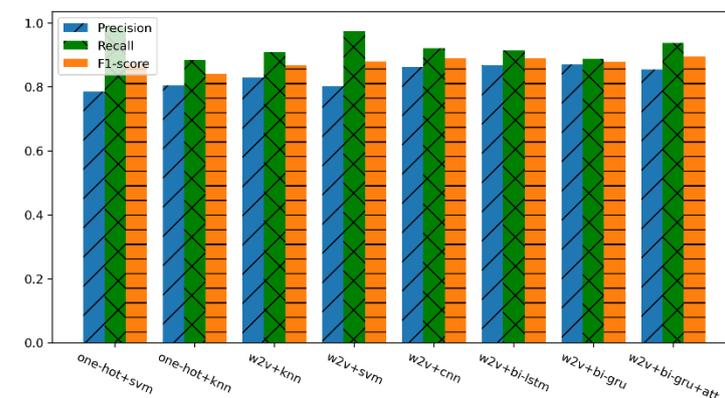


Figure 13. Model Comparison.

Table 6. Model Comparison.

Word Embedding	Model	Precision	Recall	F1-Score
One-hot	svm	78.46%	98.60%	87.39%
	knn	80.38%	88.32%	84.16%
W2v	knn	82.90%	90.90%	86.72%
	svm	80.18%	97.48%	87.99%
	cn	86.18%	92.11%	89.04%
	bi-lstm	86.67%	91.42%	88.98%
	bi-gru	87.02%	88.83%	87.92%
	bi-gru+att	85.51%	93.77%	89.45%

## 5. Discussion

As more and more unstructured restaurant reviews are exposed to consumers, how to perform rapid sentiment analysis and demand recognition on the text has become a

research hotspot. Based on the review data of Dianping.com (accessed on 1 May 2021) obtained by a web crawler, this paper used the Word2vec+Bi-GRU+Attention method to construct an online ordering platform review sentiment analysis model. It is found that the performance of the Word2vec+Bi-GRU+Attention method is higher than the commonly used sentiment analysis model.

The research in this article has certain theoretical and practical implications. First of all, in terms of theoretical implications, many scholars currently use professional sentiment dictionaries and machine learning methods to perform sentiment analysis on restaurant review texts. Traditional sentiment analysis methods rely on specific domain dictionaries, and it is influenced by the number of positive and negative words. The sentiment analysis model based on deep learning was proven to have better performance. This article uses the Word2vec+Bi-GRU+Attention method to perform sentiment analysis on Online restaurant reviews. After testing on the test set, it is found that in the environment of online ordering platforms, the comprehensive performance of Word2vec+Bi-GRU+Attention is better than the commonly used Machine learning methods and deep learning methods.

Secondly, in terms of practical implications, in the face of massive user reviews, sentiment analysis can provide consumers with decision support at a lower cost and faster speed. For example, when consumers choose a restaurant to dine at, they can select a higher quality restaurant by judging the ratio of positive reviews to negative reviews. They no longer need to read all the text, but simply combine keywords with emotional tendencies to quickly grasp the attitudes and opinions of reviewers. In addition, automated emotion recognition can enhance user satisfaction with the platform and ultimately increase consumer activity. Clustering reviews on different aspects, counting the ratio of positive reviews to negative reviews under each aspect, and consumers can choose a restaurant that suits their taste based on the distribution of different aspects of their emotions. By analyzing different aspects of a restaurant's sentiment, consumers no longer need to spend a lot of time reading and understanding each review to quickly grasp the restaurant's strengths and weaknesses. Meanwhile, for aspects where consumers have strong opinions, restaurants can make targeted improvements to improve consumer satisfaction with the restaurant.

## 6. Conclusions

The research in this article has some limitations. First, the number of positive reviews and the number of negative reviews is not balanced. The number of positive reviews is significantly higher than the number of negative reviews, which may cause deviations in the results. Second, this article uses ratings to determine the positive and negative sentiments of reviews, ignoring those high-scoring negative reviews and low-scoring positive reviews. This article judges the polarity of reviews based on restaurant review ratings. However, in reality, there is such a problem. Some consumers have given high scores but the polarity of the reviews is negative, while other consumers' behavior is just the opposite. The reader may use a satirical tone to comment.

In the future, we can consider using publicly available balanced datasets for training, or consider a combination of over-sampling and clustering techniques to make the samples more balanced. Furthermore, fine-grained machine learning methods can play an important role in identifying inconsistent reviews.

**Author Contributions:** Conceptualization, L.L.; Data curation, L.Y.; Formal analysis, L.Y.; Methodology, L.Y.; Project administration, L.L.; Resources, Y.Z.; Software, Y.Z.; Supervision, L.L.; Validation, L.Y.; Visualization, L.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Major Research Project of Innovative Groups in Guizhou Provincial Education Department, grant number Qian Jiao He KY Zi [2017]034.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

- Anaya-Sánchez, R.; Molinillo, S.; Aguilar-Illescas, R.; Liébana-Cabanillas, F. Improving travellers' trust in restaurant review sites. *Tour. Rev.* **2019**, *74*, 830–840. [CrossRef]
- Chen, Y.; Xie, J. Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Manag. Sci.* **2008**, *54*, 477–491. [CrossRef]
- Local Consumer Review Survey. Available online: <https://www.brightlocal.com/research/local-consumer-review-survey/> (accessed on 1 January 2019).
- Yang, S.-B.; Hlee, S.; Lee, J.; Koo, C. An empirical examination of online restaurant reviews on Yelp.com: A dual coding theory perspective. *Int. J. Contemp. Hosp. Manag.* **2017**, *29*, 817–839. [CrossRef]
- Marine-Roig, E.; Clave, S.A. A. A method for analysing large-scale UGC data for tourism: Application to the case of Catalonia. In *Information and Communication Technologies in Tourism 2015*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 3–17.
- Hong, H.; Xu, D.; Wang, G.A.; Fan, W. Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decis. Support Syst.* **2017**, *102*, 1–11. [CrossRef]
- Cambria, E.; Wang, H.; White, B. Guest Editorial: Big Social Data Analysis. *Knowledge-Based Syst.* **2014**, *69*, 1–2. [CrossRef]
- Mairesse, F.; Polifroni, J.; Di Fabbrizio, G. Can prosody inform sentiment analysis? Experiments on short spoken reviews. In Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, 25–30 March 2012; pp. 5093–5096.
- Zhang, Z.; Ye, Q.; Zhang, Z.; Li, Y. Sentiment classification of Internet restaurant reviews written in Cantonese. *Expert Syst. Appl.* **2011**, *38*, 7674–7682. [CrossRef]
- Fan, Z.-P.; Che, Y.-J.; Chen, Z.-Y. Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis. *J. Bus. Res.* **2017**, *74*, 90–100. [CrossRef]
- Nurifan, F.; Sarno, R.; Sungkono, K.R. Aspect Based Sentiment Analysis for Restaurant Reviews Using Hybrid ELMO-Wikipedia and Hybrid Expanded Opinion Lexicon-SentiCircle. *Int. J. Intell. Eng. Syst.* **2019**, *12*, 47–58.
- Xia, H.; Yang, Y.; Pan, X.; Zhang, Z.; An, W. Sentiment analysis for online reviews using conditional random fields and support vector machines. *Electron. Commer. Res.* **2020**, *20*, 343–360. [CrossRef]
- Su, Y.-J.; Hu, W.-C.; Jiang, J.-H.; Su, R.-Y. A novel LMAEB-CNN model for Chinese microblog sentiment analysis. *J. Supercomput.* **2020**, *76*, 9127–9141. [CrossRef]
- Kumar, K.N.; Uma, V. Intelligent sentiment-based lexicon for context-aware sentiment analysis: Optimized neural network for sentiment classification on social media. *J. Supercomput.* **2021**, 1–25.
- Mudambi, S.M.; Schuff, D. Research note: What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Q.* **2010**, *34*, 185–200. [CrossRef]
- Nakayama, M.; Wan, Y. The cultural impact on social commerce: A sentiment analysis on Yelp ethnic restaurant reviews. *Inf. Manag.* **2019**, *56*, 271–279. [CrossRef]
- Jurafsky, D.; Chahuneau, V.; Routledge, B.R.; Smith, N.A. Narrative framing of consumer sentiment in online restaurant reviews. *First Monday* **2014**. [CrossRef]
- Jia, S.S. Motivation and satisfaction of Chinese and US tourists in restaurants: A cross-cultural text mining of online reviews. *Tour. Manag.* **2020**, *78*, 104071. [CrossRef]
- Meek, S.; Wilk, V.; Lambert, C. A big data exploration of the informational and normative influences on the helpfulness of online restaurant reviews. *J. Bus. Res.* **2021**, *125*, 354–367. [CrossRef]
- Tian, G.; Lu, L.; McIntosh, C. What factors affect consumers' dining sentiments and their ratings: Evidence from restaurant online review data. *Food Qual. Prefer.* **2021**, *88*, 104060. [CrossRef]
- Li, H.; Qi, R.; Liu, H.; Meng, F.; Zhang, Z. Can time soften your opinion? The influence of consumer experience valence and review device type on restaurant evaluation. *Int. J. Hosp. Manag.* **2021**, *92*, 102729. [CrossRef]
- Lee, M.; Jeong, M.; Lee, J. Roles of negative emotions in customers' perceived helpfulness of hotel reviews on a user-generated review website: A text mining approach. *Int. J. Contemp. Hosp. Manag.* **2017**. [CrossRef]
- Luo, Y.; Xu, X. Comparative study of deep learning models for analyzing online restaurant reviews in the era of the COVID-19 pandemic. *Int. J. Hosp. Manag.* **2021**, *94*, 102849. [CrossRef]
- Micu, A.; Micu, A.E.; Geru, M.; Lixandriou, R.C. Analyzing user sentiment in social media: Implications for online marketing strategy. *Psychol. Mark.* **2017**, *34*, 1094–1100. [CrossRef]
- Schouten, K.; Frascar, F. Finding implicit features in consumer reviews for sentiment analysis. In *International Conference on Web Engineering*; Springer: Berlin/Heidelberg, Germany, 2014.
- Li, H.; Liu, H.; Zhang, Z. Online persuasion of review emotional intensity: A text mining analysis of restaurant reviews. *Int. J. Hosp. Manag.* **2020**, *89*, 102558. [CrossRef]

27. Krishna, A.; Akhilesh, V.; Aich, A.; Hegde, C. Sentiment analysis of restaurant reviews using machine learning techniques. In *Emerging Research in Electronics, Computer Science and Technology*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 687–696.
28. Medhat, W.; Hassan, A.; Korashy, H. Sentiment analysis algorithms and applications: A survey. *Ain Shams Eng. J.* **2014**, *5*, 1093–1113. [[CrossRef](#)]
29. Li, L.; Goh, T.-T.; Jin, D. How textual quality of online reviews affect classification performance: A case of deep learning sentiment analysis. *Neural Comput. Appl.* **2020**, *32*, 4387–4415. [[CrossRef](#)]
30. Zheng, L.; Wang, H.; Gao, S. Sentimental feature selection for sentiment analysis of Chinese online reviews. *Int. J. Mach. Learn. Cybern.* **2018**, *9*, 75–84. [[CrossRef](#)]
31. Hogenboom, A.; Frasinca, F.; de Jong, F.; Kaymak, U. Polarity classification using structure-based vector representations of text. *Decis. Support Syst.* **2015**, *74*, 46–56. [[CrossRef](#)]
32. Sun, Q.; Niu, J.; Yao, Z.; Yan, H. Exploring eWOM in online customer reviews: Sentiment analysis at a fine-grained level. *Eng. Appl. Artif. Intell.* **2019**, *81*, 68–78. [[CrossRef](#)]
33. Zhang, L.; Wang, S.; Liu, B. Deep learning for sentiment analysis: A survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, e1253. [[CrossRef](#)]
34. Abdi, A.; Shamsuddin, S.M.; Hasan, S.; Piran, J. Deep learning-based sentiment classification of evaluative text based on Multi-feature fusion. *Inf. Process. Manag.* **2019**, *56*, 1245–1259. [[CrossRef](#)]
35. Al-Smadi, M.; Qawasmeh, O.; Al-Ayyoub, M.; Jararweh, Y.; Gupta, B. Deep Recurrent neural network vs. support vector machine for aspect-based sentiment analysis of Arabic hotels' reviews. *J. Comput. Sci.* **2018**, *27*, 386–393. [[CrossRef](#)]
36. Araque, O.; Corcuera-Platas, I.; Sánchez-Rada, J.F.; Iglesias, C.A. Enhancing deep learning sentiment analysis with ensemble techniques in social applications. *Expert Syst. Appl.* **2017**, *77*, 236–246. [[CrossRef](#)]
37. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient estimation of word representations in vector space. *arXiv* **2013**, arXiv:1301.3781.
38. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv* **2014**, arXiv:1412.3555.
39. Niu, Z.; Zhong, G.; Yu, H. A review on the attention mechanism of deep learning. *Neurocomputing* **2021**, *452*, 48–62. [[CrossRef](#)]
40. Hu, D. An introductory survey on attention mechanisms in NLP problems. In *Proceedings of SAI Intelligent Systems Conference*; Springer: Berlin/Heidelberg, Germany, 2019.
41. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [[CrossRef](#)]
42. Hu, N.; Koh, N.S.; Reddy, S.K. Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales. *Decis. Support Syst.* **2014**, *57*, 42–53. [[CrossRef](#)]
43. Chinese Common Stop Words List. Available online: <https://github.com/goto456/stopwords> (accessed on 1 May 2021).
44. Zhang, W.; Kong, S.-X.; Zhu, Y.-C. Sentiment classification and computing for online reviews by a hybrid SVM and LSA based approach. *Clust. Comput.* **2019**, *22*, 12619–12632. [[CrossRef](#)]
45. Rhanoui, M.; Mikram, M.; Yousfi, S.; Barzali, S. A CNN-BiLSTM model for document-level sentiment analysis. *Mach. Learn. Knowl. Extr.* **2019**, *1*, 48. [[CrossRef](#)]
46. Fu, Y.; Liao, J.; Li, Y.; Wang, S.; Li, D.; Li, X. Multiple perspective attention based on double BiLSTM for aspect and sentiment pair extract. *Neurocomputing* **2021**, *438*, 302–311. [[CrossRef](#)]