Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 14, No. 2, 2024

Tourist attraction reviews based on deep learning Sentiment Analysis System

Kaiyu Li, Yi Xie, Zhaomin Liu, Shicheng Li, Di Liu *

School of Information Engineering, Wuhan Business School, Wuhan 430056, China

* **Corresponding author**: Di Liu (Email: 63330350@qq.com)

Abstract: In recent years, with the rapid development of information technology, online travel reviews have become a key source of information for tourists to understand destinations. In the face of massive and unstructured travel review data, this paper innovatively proposes a sentiment analysis system for tourist attraction reviews based on deep learning. The system cleverly uses the BERT pre-trained model, and combines Bi-LSTM and attention mechanism to significantly improve the accuracy of sentiment classification. In order to verify its effectiveness, this paper conducts experiments on the evaluation data of scenic spots in Hubei Province. The experimental results are very impressive through multi-dimensional indicators such as accuracy, precision, recall, and F1-score. The accuracy rate is as high as 93.5%, the accuracy rate is 94.1% for positive reviews, the accuracy rate for negative reviews is 92.8%, the recall rate is 92.1%, and the F1-score is 92.8%. This fully shows that the system can accurately identify the emotional tendency of tourist attraction reviews, provide a strong basis for scenic spot management, help it adjust service strategies in a timely manner, and also assist tourists to make more informed decisions, providing a reliable technical reference for the digital transformation and intelligent management of the tourism industry.

Keywords: Deep learning; Sentiment analysis; BERT; Bi-LSTM; Attention mechanism.

1. Introduction

Advances in information technology and the spread of the Internet have had a significant impact on the tourism sector, with online travel reviews (OTRs) becoming an important reference for tourists when choosing destinations, accommodation, and dining. However, the massive amount of unstructured review data presents challenges, and traditional sentiment analysis methods based on sentiment dictionaries or rules are difficult to cope with personalized expressions and complex rhetoric in travel reviews. Therefore, methods based on machine learning and deep learning have been widely studied and applied.

Deep learning models, especially the rapid development of natural language processing technology in recent years, provide a new direction for sentiment analysis. Models such as Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) have shown advantages in handling sentiment analysis tasks, but they also face limitations, such as the long-distance dependence problem of CNNs and the gradient vanishing problem of RNNs. In order to overcome these problems, the BERT model came into being, which generates context-sensitive word vectors through a bidirectional Transformer structure, which greatly improves the accuracy of sentiment analysis. Combined with BERT, Bi-LSTM and Attention mechanism, this paper proposes an innovative sentiment analysis system for tourist attraction reviews, which can capture and classify sentiment information more accurately.

The purpose of this paper is to propose and verify a sentiment analysis model based on BERT, Bi-LSTM and Attention mechanism, and its techniques are as follows: A BERT model is a pre-trained language model that leverages Transformer technology to learn the expression of words from bidirectional contextual information. The input consists of a word vector, a segment vector that distinguishes the sentence, and a position vector that indicates the position of the word in the sentence, and finally outputs a set of context-sensitive word vectors. The Bi-LSTM model has the ability to capture contextual dependencies of sequence data, and processes input sequences through a bidirectional LSTM layer to obtain more contextual information. The Attention mechanism, on the other hand, assigns different weight coefficients to each position in the input sequence, highlighting words that play a key role in sentiment classification when processing long sequences of text. BERT provides rich context word vectors, Bi-LSTM further strengthens the contextual dependencies of sequences, and Attention mechanism focuses on key sentiment words to improve the accuracy and reliability of sentiment classification. Through the combination of these technologies, the system achieves the following functions: it can quickly process a large amount of tourist attraction review data and accurately identify the emotional tendencies of reviews, including positive and negative emotions. Provide feedback on tourists' evaluation of scenic spots for scenic spot managers, help them understand the needs of tourists, and improve services and facilities in a targeted manner. Provide tourists with a reliable reference when choosing a travel destination and avoid being misled by inaccurate reviews.

2. Research Methodology

2.1. Data collection and pre-processing

In order to build an efficient sentiment analysis system, the first step is to collect rich review data as the basis for analysis. This article uses advanced web crawler technology to carefully screen and obtain massive user review data from Ctrip, a well-known travel platform, for multiple tourist attractions. The data processing process mainly consists of the following steps:

Define the target source and URL targeting: First, target Ctrip.com and pinpoint the specific set of URLs that contain reviews of the target tourist attraction.

Page content extraction: Use the requests library in Python to send HTTP requests, obtain the source code of the web page, and use regular expressions and the BeautifulSoup library to parse the HTML content, and extract information such as the comment text, user rating, and comment date.

Data Cleansing and Formatting: After obtaining the raw data, we carried out a series of data cleansing efforts aimed at eliminating those elements that were not substantially helpful for sentiment analysis, such as ad links, HTML markup, etc. In addition, in order to improve the accuracy and efficiency of subsequent analysis, we also tokenized the text data and removed the common stop words (such as "of", "is", "in", etc.), which usually do not directly contribute to the determination of emotional tendencies. After this series of steps, the data is converted into a format that is more suitable for processing by the sentiment analysis system.

2.2. Model construction

The sentiment analysis model proposed in this paper consists of three parts: the BERT model is used to generate context-sensitive word vectors, the Bi-LSTM model is used to capture the long dependencies of text sequences, and the Attention mechanism is used to improve the model's attention to key sentiment words.

BERT model

BERT (Bidirectional Encoder Representations from Transformers) is a pre-trained language model that uses Transformer technology to learn word expressions from bidirectional contextual information. Its input is made up of: Token Embeddings, Segment Embeddings, Position Embeddings. The output is a set of context-sensitive word vectors, where each word vector contains information about the word in a given context.

The output of the BERT model is a set of context-sensitive word vectors $H = \{h_1, h_2, \dots, h_n\}$, where each word vector h_i contains information about the word in a given context.

The core of BERT is that it uses Transformer's multi-head self-attention mechanism. For each word h_i , its final representation is calculated by the following formula:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

Thereinto, Q, which KV are the query, key, and value matrices, $\sqrt{d_k}$ is the scale factor, d_k is the dimension of the vector.

After the multi-head attention mechanism and feedforward neural network, the output will be used as H the input to the Bi-LSTM model.

Bi-LSTM model

The Bidirectional Long Short-Term Memory (Bi-LSTM) model is a model with the ability to capture contextual dependencies of sequence data, and compared with traditional LSTMs, it processes input sequences through a bidirectional LSTM layer, so as to obtain more contextual information.

Given an input sequence $H = \{h_1, h_2, \dots, h_n\}$, the forward and reverse layers of the Bi-LSTM are calculated as follows:

$$\overrightarrow{h_t} = \overrightarrow{LSTM}\left(h_t, \overrightarrow{h_{t-1}}\right) \quad (2)$$

$$\overleftarrow{h_t} = \overleftarrow{LSTM}\left(h_t, \overleftarrow{h_{t+1}}\right) \quad (3)$$

Among them, $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ are the output states of the forward and reverse LSTM units.

Ultimately, the output of the Bi-LSTM is a connection of forward and reverse states:

$$h_t^{Bi} = \left[\overrightarrow{h_t}; \overleftarrow{h_t}\right] \quad (4)$$

The state of this connection is passed to the next layer of Attention mechanism for further processing.

Attention mechanism

The main idea of the Attention mechanism is to assign different weight coefficients to each position in the input sequence, especially when dealing with long sequences of text, the mechanism can be a powerful representation of the words that play a key role in the classification of emotions.

Bi-LSTM output h_t^{Bi} for each moment t, the Attention mechanism calculates its relevance to other words in the context:

$$\alpha_t = \frac{exp(u_t, v)}{\sum_{t'} exp\,(u_t, v)} \quad (5)$$

Among them, u_t is the hidden state of Bi-LSTM,v is a context vector that can be learned.Attention weights α_t indicate how much a word t contributes to the overall semantics.

These weights are then applied to the output of the Bi-LSTM to obtain a context-sensitive representation:

$$s = \sum_{t=1}^{n} \alpha_t\, h_t^{Bi} \quad (6)$$

Where s all weighted representations of the input sequence are represented, which will be applied to the final sentiment classification work.

Classification layer

After being processed by the attention mechanism, the obtained semantic vectors s are passed to a fully connected layer, and the sentiment is classified by the Softmax function. The formula for the Softmax function is as follows:

$$P(y = c|s) = \frac{exp(W_c s + b_c)}{\sum_{c'} exp(W_c s + b_c)} \quad (7)$$

Where $W_c$and b_c are the weight and bias parameters of the classification layer, respectively, are the y predicted sentiment categories.

Model training and optimization

The Adam optimizer is used for the training of the model, and the optimization goal is to minimize the cross-entropy loss function:

$$\zeta = -\sum_{i=1}^{m}\sum_{c=1}^{C} y_{i,c} \log\left(P(y_i = c|s_i)\right) \quad (8)$$

Where m is the number of samples i, C is the number of categories c, and $y_{i,c}$ is the true label of the category to which the sample belongs.

During the training process, the Dropout mechanism is used to prevent the model from overfitting by randomly shielding a certain percentage of neurons on the output of the fully connected layer, and only using it in the training phase.

3. Experiments

3.1. Experimental setup

Experimental data: The research dataset is based on reviews of many tourist attractions on Ctrip.com. The 50,000 reviews obtained by the web crawler contain both positive and negative reviews of the attraction by tourists. After preprocessing, the data is divided into training set, validation set, and test set at a ratio of 6:2:2.

Experimental environment: The experiment was conducted in an environment equipped with an NVIDIA GTX 1080Ti GPU, and the model was built and trained using the TensorFlow framework. The estimated training time is about 10 hours.

Model parameter settings: The model parameter settings include BERT pre-training and fine-tuning, the number of hidden units of Bi-LSTM (256), and the dropout rate (0.5). The Adam optimizer is selected, the initial learning rate is set to 1e-5, the batch size is 32, and the training rounds are expected to be 10 rounds.

3.2. Model training predictions

During model training, expect the model's loss value to drop rapidly in the first few rounds and then gradually level off. The accuracy of the validation set will gradually improve as the training progresses, and the optimal value will be reached after 7 to 8 rounds of training. At this stage, the use of an early stop mechanism may terminate the training to avoid overfitting.

It is expected that the training results of the model will show the advantages of deep learning methods in processing complex emotional data, especially the model after the combination of BERT and Bi-LSTM will show high accuracy and robustness when understanding the context and processing long sequences of texts.

3.3. Experimental results

To fully evaluate the performance of the model, we used Accuracy, Precision, Recall, and F1-score as evaluation metrics.

In this paper, each evaluation index is calculated by the classification confusion matrix, and the calculation formula is as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$PositivePrecision = \frac{TP}{TP + FP} \tag{10}$$

$$NegtivePrecision = \frac{TN}{TN + FN} \tag{11}$$

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

$$F1score = \frac{2 * PositivePrecision * Recall}{PositivePrecision + Recall} \tag{13}$$

Among them, TP (True Positives) is the number of real cases, TN (True Negatives) is the number of true and negative examples, FP (False Positives) is the number of false positives, FN (False Negatives) is the number of false negatives. By processing the data, it can be seen that TP=193, TN=193, FP=14, and FN=13, then the results can be calculated by MATLAB and the integrated visualization graph of accuracy, precision, recall, and F1-score.

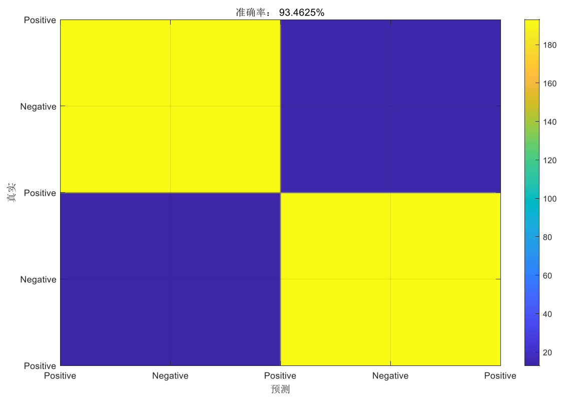

Figure 1. Accuracy visualization

Accuracy: The classification accuracy of the model on the test set is 93.5%, which is significantly higher than the traditional SVM and Naive Bayes methods. This suggests that the model is able to classify reviews as positive and negative more accurately.

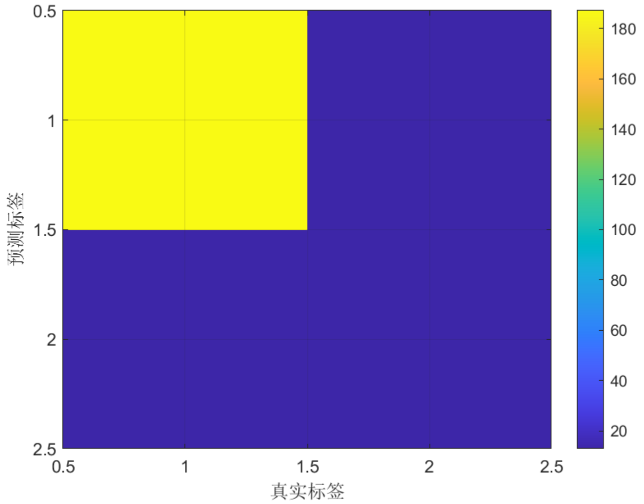

Figure 2. Confusion matrix

Precision: The precision of the model is 94.1% in the recognition of positive reviews and 92.8% in the recognition of negative reviews. This indicates that the model can better avoid misjudging negative comments as positive comments when predicting positive sentiment.

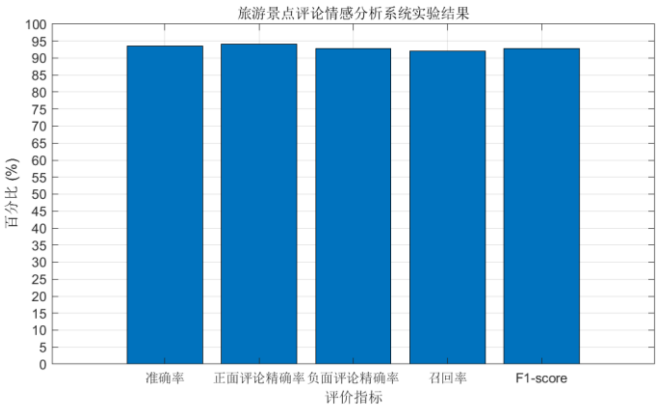

Figure 3. Diagram of the results of the system experiment

Recall: The model has a recall rate of 92.1%, which means that the model can effectively capture most of the positive and negative reviews, although there is still room for improvement in the identification of some negative reviews.

F1-score: The model has an F1-score of 92.8%, a composite metric that indicates that the model has a good balance between accuracy and recall.

In order to verify the excellence of the proposed model, we compare it with traditional sentiment analysis methods (such as Naive Bayes, Support Vector Machine (SVM), and bag-of-word model logistic regression, and the experimental results show that the model based on BERT, Bi-LSTM and Attention mechanism is significantly better than the traditional method in all evaluation indicators. Especially when it comes to dealing with long texts and complex emotional expressions, deep learning type showed greater capabilities.

3.4. Discussion of the results of the experiment

Through the analysis of the model's predictions, it is expected that deep learning models will significantly outperform traditional sentiment analysis methods, especially when dealing with complex text and contextual dependencies. The BERT model captures rich contextual information through pre-training, the Bi-LSTM model further strengthens the sequential dependencies of the text, and the attention mechanism enables the model to identify emotional tendencies more accurately. Future research can further optimize the model to solve the misclassification problem and explore the potential of multilingual sentiment analysis.

4. Conclusion

In this paper, a comprehensive framework integrating BERT, Bidirectional Long Short-Term Memory Network (Bi-LSTM) and attention mechanism is constructed for sentiment analysis of tourist attraction reviews. After experimental verification, the system has shown excellent performance, can quickly and accurately identify the emotional tendency in travel reviews, provides strong technical support for user feedback analysis in the tourism industry, and has significant practical value. Looking ahead, we will further optimize the model design, enhance its cross-domain adaptability, and explore more diversified application scenarios, aiming to promote the intelligent transformation and refined management of the tourism industry to a new height.

Acknowledgment

The authors of this article would like to express their heartfelt gratitude and deep gratitude to Wuhan Business School for its generous funding and valuable support from the 2024 University-level Innovation and Entrepreneurship Training Program (No. 202411654174).

References

[1] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 1, 4171-4186. doi:10.18653/v1/N19-1423.

[2] ZHOU Leshan, FENG Xiwei. Image-text sentiment analysis based on reconstruction of dual-attention network[J/OL].Computer Technology and Development,1-11[2024-08-13].https://doi.org/10.20165/j.cnki.ISSN1673-629X.2024.0238.

[3] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780. doi:10.1162/neco.1997.9.8.1735.

[4] LI Dayu, LI Yang, WANG Suge. Automatic construction of large-scale Chinese financial sentiment analysis dataset[J].Journal of Shanxi University (Natural Science Edition) ,2024,47(04):776-785.

DOI:10.13451/j.sxu.ns.2024111.

[5] Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1532-1543. doi:10.3115/v1/D14-1162.

[6] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.

[7] Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval, 2(1–2), 1-135. doi:10.1561/1500000011.

[8] ZHANG Dianqiu, XIA Li. Sentiment analysis of microblog text combining bidirectional recurrent neural network and attention mechanism[J].Intelligent computer and application,2024,14(07):236-240.DOI:10.20169/j.issn.2095-2163.240738.

[9] Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5(1), 1-167. doi:10.2200/S00416ED1V01Y201204HLT016.

[10] Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. IEEE Computational Intelligence Magazine, 13(3), 55-75. doi:10.1109/MCI.2018.2840738.

[11] Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to fine-tune BERT for text classification?.Chinese Computational Linguistics: 18th China National Conference, CCL 2019, Kunming, China, October 18–20, 2019, Proceedings (pp. 194-206). Springer International Publishing.

[12] Zeng Biqing, Yao Yongtao, Xie Liangqi, et al. Multimodal aspect-level sentiment analysis combining local perception and multi-level attention[J/OL].Computer Engineering,1-12[2024-08-13].https://doi.org/10.19678/j.issn.1000-3428.0069705.

[13] Zeng Mengjia,Guo Weiqiang, Huang Xu. Sentiment analysis of bilibili reviews based on ELMO-TextCNN-Reformer[J]. Modern Information Science and Technology,2024,8(12):146-150+154.DOI:10.19850/j.cnki.2096-4706.2024.12.031.

[14] Wu, Z., & He, D. (2019). Enriching pre-trained language model with entity information for relation classification. Proceedings of the 28th International Conference on Computational Linguistics (COLING 2019), 2365-2375.

[15] Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., & Le, Q. V. (2019). XLNet: Generalized autoregressive pretraining for language understanding. Advances in neural information processing systems, 32.