Text Mining and Sentiment Classification for Logistics Enterprises Evaluation Based on BERT

Lihang Cheng[1], Siyuan Guo[2], Yifan Liu[3] and Yi Zhuang[4]
[1]*School of Advanced Manufacturing, Fuzhou University, Jinjiang, China*
[2]*Department of Electrical Engineering, College of Mechanical and Electrical Engineering, Fujian Agriculture and Forestry University, Fuzhou, China*
[3]*Department of Mechanical Engineering, School of Mechanical and Electrical Engineering, Anhui University of Science & Technology, Huainan, China*
[4]*Department of Information Management and Information Systems, School of Management Science and Engineering, Shandong University of Finance and Economics, Jinan, China*

Keywords: Text Mining, Sentiment Analysis, BERT, Evaluation of Logistics Enterprises.

Abstract: In the era of community internet and intelligent industry, evaluation text data, as a novel alternative data resource, is widely utilized by the industrial and commercial sector. In this paper, we innovatively stand in the perspective of logistics industrial informatics, consider evaluation text and sentiment features as the key information reflecting the satisfaction of logistics enterprises, and construct experiments using pre-trained models, and consider them as one of the normalized data for sentiment classification. In other words, deep learning techniques were utilized to analyze the user evaluations of each logistics enterprise on the microblogging platform, which were fed into the Bert model to discriminate the sentiment polarity, and were able to classify the predictions with a high degree of accuracy. It provides a path to further extract the distribution of emotional tendency and the evaluation theme words of logistics enterprises from the text data, which expands the perspective and dimension of users' choice of logistics enterprises, and also helps the logistics enterprises to improve their services based on the evaluation, and helps the development of the logistics industry and the evaluation research system of logistics management from the side of alternative data mining and analysis.

1 INTRODUCTION

With the development of the Internet community, the demand for big data in the Industrial Internet of Things is increasing, and the degree of information integration with other fields is enhanced, and the exploration and utilization of alternative data in the industry is also increasing. In the current era of e-commerce, prediction and evaluation methods for alternative data can fully serve the process of product and service provision and selection.

For the logistics industry, text as a kind of alternative data, rich inventory and availability, but also an important information resource, can be a more comprehensive reflection of the logistics satisfaction of an enterprise and the level of service, so the means of text mining for the enterprise or the end customer to provide a new perspective for the evaluation of the strength of the enterprise logistics. It can help customers choose high-quality enterprises, logistics enterprises themselves can also locate customer pain points, understand the demand, and then targeted to enhance certain aspects of the ability to improve service quality. At the macro-logistics level, it can also provide guiding suggestions for the entire logistics industry, and utilize two-way adjustment interactions based on sentiment analysis between logistics enterprises and customers to promote high-quality development of the industry in the information age.

This study aims at the latest Chinese Internet evaluation data, in the Sina microblogging platform using the crawler program to obtain the recent thousands of evaluation text about each logistics enterprise, call the BERT pre-training model to fully capture the complex relationship and emotional information in the text, use the data for the logistics

Cheng, L., Guo, S., Liu, Y. and Zhuang, Y.
Text Mining and Sentiment Classification for Logistics Enterprises Evaluation Based on BERT.
DOI: 10.5220/0012928600004536

In *Proceedings of the 1st International Conference on Data Mining, E-Learning, and Information Systems (DMEIS 2024)*, pages 113-118
ISBN: 978-989-758-715-3

context of the BERT model fine-tuning, so that it can realize the emotional polarity of the evaluation text data of the logistics enterprise. After the model validation test, the relevant indexes show that the BERT model effectively completes the sentiment classification prediction of logistics enterprise evaluation, and then the Fine-tuning test with different data imbalance is conducted around the model tuning, and the results show that the model can achieve the best performance and complete the task of predicting the sentiment of the text of the logistics enterprise when the data imbalance is 30-40%.

2 RELEVANT LITERATURE

Text data, as a typical unstructured data, has also gradually started to be studied in the era of big data.Feldman R et al. first proposed the concept of text mining in their work (Feldman, Sanger, 2007) and introduced several methods to perform mining analysis. Since Chinese text does not come with its own disambiguation characters, many scholars have tried various methods in solving features, recognition and representation,. For example, scholars such as Xu G made a study on several different algorithms for Chinese participles and proposed an automatic Chinese participle system based on the forward maximal matching method (Xu, Hu, Wang, 2007) while Hu Yan et al. proposed a method for text feature extraction from the point of view of Chinese lexical properties (Hu, Wu, Zhong, 2007).

The analysis of text sentiment is also an important task and branching direction in text mining. The main methods of text sentiment analysis as Hu R mentioned in their review study (Hu, Rui, Zeng, 2018) from the earliest sentiment lexicon matching as Li J et al. used (Li, Xu, Xiong, 2010), to the traditional machine learning methods such as Bayes, SVM and so on as Hasan A and other scholars used (Hasan, Moin, Karim, 2018) and then evolved to the current deep learning methods based on various types of neural networks as Dang N C mentioned in his review study mentioned new methods such as CNN, RNN, LSTM models currently used for sentiment analysis (Dang, Moreno-García, De la Prieta, 2018) .

Nowadays, the mining of sentiment in text has also become more delicate and in-depth, and the term "fine-grained sentiment analysis" has been proposed, for example, Lai Y and other scholars have used CNN to realize fine-grained sentiment classification of microblog text (Lai, Zhang, Han, 2020) .

Treiblmaier and other scholars systematically reviewed the use of word cloud graphs, topic models and other methods for text data extraction and sentiment analysis in logistics and supply chain management (Treiblmaier, Mair, 2021), constructing a method system for logistics text mining analysis. While Singh et al. incorporated sentiment analysis indicators into the overall performance evaluation system of 3PL logistics enterprises (Singh et al., 2022). Hong and Lim, both groups of researchers based on the perspective of user satisfaction, respectively, utilized CNN (Hong et al., 2019) and Bi-LSTM models for text sentiment analysis to parse the elements of users logistics satisfaction in the specific scenarios of fresh e-commerce logistics and cold chain logistics, and put forward suggestions for improvement (Lim, Li, Song, 2021).

3 BERT-BASED SENTIMENT ANALYSIS MODEL

3.1 Structure of BERT

The BERT model was first proposed by Devlin J et al. of Google in 2018 (Devlin, Chang, Lee, et al., 2018), BERT is Bidirectional Encoder Representations from Transformers, and its basic structure mainly consists of the Encoder part of the Transformer model (Vaswani, Shazeer, Parmar, et al., 2017). It consists of a fully-connected stack of multiple Encoder units and uses two pre-training tasks, MLM and NSP, which have excellent performance in processing short and medium texts and consider the contextual bi-directional interaction. The basic structure of the model is shown in Fig.1.

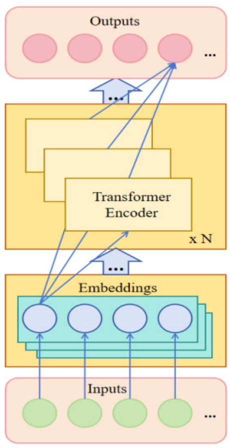

Figure 1: The structure of BERT.

3.2 Transformer Encoder

The Transformer Encoder is a part of the Transformer model for feature extraction. It consists of multiple Encoder Layers, and each Encoder Layer contains two sub-layers: a multi-head self-attention mechanism layer and a fully connected feedforward layer. Its specific structure is shown in Fig. 2.

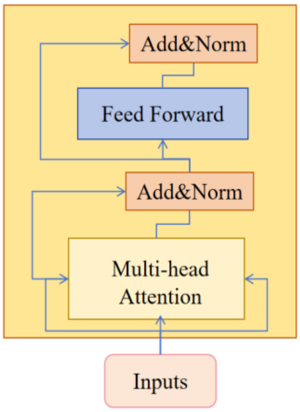

Figure 2: Processing of Transformer Encoder.

This layer inputs a sequence of inputs as Query, Key and Value into an attention mechanism, which then multiplies and sums the attention weights α with the Value to produce an output representation O. And it can be implemented as a multi-head mechanism by splitting the inputs into multiple sub-vectors. The formula is as follows:

$$\begin{aligned}&\mathrm{MultiHead}(Q,K,V)\\&=\mathrm{Concat}\,(\,head_1,\dots,\ head_h)W^O\end{aligned} \tag{1}$$

$$head_i=\mathrm{Attention}\left(QW_i^Q,KW_i^K,VW_i^V\right) \tag{2}$$

$$\mathrm{Attention}\,(Q,K,V)=\mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{3}$$

where, head denotes the number of heads, W^Q, W^K and W^V is the weight matrixs for linear transformation. W^O is the weight matrix that splices the multiple results and obtains the final output through a linear transformation. d_k is the dimension of the key.

3.2.2 Add and Normalize

The output of the Multihead Self-Attention Mechanism layer requires residual linking and layer normalization. Residual linking is to add the inputs to the outputs to reduce the loss of information, while layer normalization is to normalize all the feature dimensions of each sample to improve the stability of the model and the speed of convergence. The formulas is as follow:

$$\mathrm{LayerNorm}\,(x+\mathrm{MultiHead}\,(Q,K,V)) \tag{4}$$

where, x is the input vector.

3.2.3 Feed Forward Networks

In the fully-connected feedforward layer, the output representation is transformed by a combination of two linear transformations and an activation function (usually is Relu). so as to increase the nonlinear and representational capabilities of the model and better capture the semantic information in the sequence. The formula is as follows:

$$\mathrm{FFN}\,(x)=\max(0,xW_1+b_1)W_2+b_2 \tag{5}$$

where W_1,b_1 and b_2 are the weight matrices and bias vectors of linear transformations.

Both residual connectivity and layer normalization are used in the fully connected feedforward layer. The formula is as follows

$$\mathrm{LayerNorm}\,(x+\mathrm{FFN}(x)) \tag{6}$$

3.3 Tokenization Based on BERT

In the BERT model, the input data is processed by means of word embedding, which requires three steps of processing to obtain a valid text vector representation. The first is the Token Embedding phase, where each word or token is converted into a fixed dimensional embedding vector to capture the semantic relationships between words. Next is Segment Embedding for distinguishing semantics and associations between different sentences. Finally Position Embedding, which considers the order of words in the text. The input layer of the BERT model is shown in Fig.3.

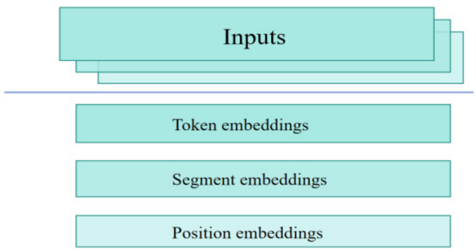

Figure 3: Tokenization Based on BERT.

3.4 Pre-Training Strategies for BERT

As a pre-trained model, BERT's "bi-directional" comprehension and powerful natural language processing performance mainly come from its pre-training phase, which uses a large amount of unlabeled textual data for two major training tasks to learn contextual semantics and sentence associations through the Masked Language Model (MLM) and the Next Sentence Prediction (NSP) tasks to learn contextual semantics and sentence associations. In the MLM task, the model learns to understand the missing words in the text by predicting the masked words in a "Mask" fashion, using contextual words and probabilities. In the NSP task, the model predicts coherence and semantic associations between two sentences, improving the model's ability to understand sentence-level semantics.

3.5 Fine-Tuning and Model Calling

When we use the BERT model to deal with the downstream tasks in the experimental design, we are actually calling the BERT model, which has been unsupervised trained on a large amount of corpus, to carry out Fine-tuning, i.e., to introduce the dataset of evaluation of logistics enterprises to be re-trained on the basis of the pre-trained model to adjust the corresponding hyper-parameters and optimize the performance in the classification of the evaluation text sentiment, so as to make the model better able to optimize the performance of evaluation text sentiment classification, so that the model can better adapt to and complete the task of evaluation text sentiment analysis and mining in the field of logistics.

4 EXPERIMENT

4.1 Data Access

In this study, more than 500 pieces of text data were crawled by python crawler program using keywords of different logistics enterprises' names respectively, which covered a total of 9 mainstream enterprises in China and contained fields as shown in Table 1.

Table 1: Sample fields.

Time	Weibo_ID	Gender	Text

4.2 Data Preprocessing

As the crawler program directly crawls the original text of the comments, the text carries a large number of invalid characters (e.g., emoticons, topic tags, special symbols, etc.), which are preprocessed by a data cleaning program to keep the most valuable text information in order to facilitate the efficacy of the BERT model for learning and classification.

Table 2: Labeling standards.

Sentiment	Label
Positive	2
Neutral	1
Negative	0

For fine-grained model classification, the BERT classification model constructed in the previous section has to be pre-trained using data with annotations, and for the data cleaned in the previous step, we classified 3 categories of sentiment according to the criteria in Table 2.

4.3 Model Training

After all the data have been pre-processed, this study divides the data into a training set, a test set and a validation set in the ratio of "8:1:1", and captures the sentiment information in the evaluation text of logistics companies through training, which is essentially Fine-tuning for the BERT pre-trained model for it to learn.

5 RESULTS

5.1 Model Performance

Metrics such as accuracy, precision, recall, and F1 score of the model on the test set are calculated to determine the performance of the model. The common metrics for evaluating the training effect of the model are Accuracy and Log-Loss, the higher the accuracy, the better the model's classification effect.

Since the BERT model still belongs to the composite attention network structure, the adjustment of the parameters is realized through the back-propagation of the loss function values, which measures the gap between the predicted and actual values of the model. Therefore, the smaller the log-loss value is, the better the model's classification effect is. By setting the number of training rounds, the model can be trained multiple times, thus allowing the model to be optimized gradually. The following figure shows the curve of the change of the accuracy and loss values during the training process.

From the above figure, it can be seen that with the increase of the number of model training, the accuracy of the BERT model gradually increases, and finally reaches 0.8795, and the log loss value is reduced to close to 0, indicating that the model has been effectively trained. Meanwhile, from the validation set, when the epoch reaches 3, the difference between the loss value and the training set is only 0.3, which indicates that the model does not have overfitting phenomenon and has a certain degree of generality.

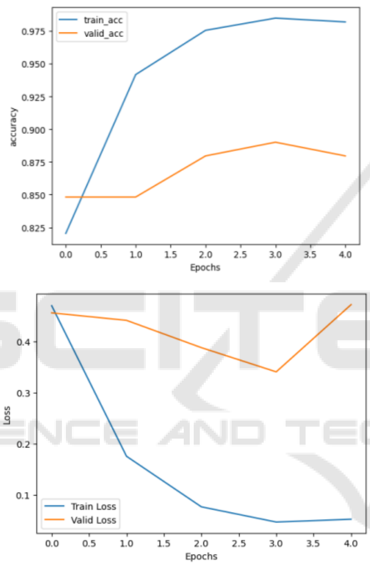

Figure 4: Changes in classification accuracy and loss values of the BERT model.

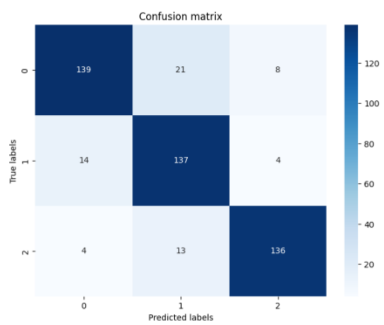

Figure 5：BERT model confusion matrix.

The following are the confusion matrices for positive (2), neutral (1) and negative (1) classification, and the size of the classification performance indicators (Pre, Recall, F1) for each of the above models can be calculated based on the values of each block of the confusion matrix values. At the same time, it can be seen that the color of the diagonal block from the matrix is significantly darker than the other blocks, which intuitively shows that the BERT model for sentiment classification of logistics enterprises trained in this study has good classification performance.

5.2 Data Distribution Tuning

Since the evaluation of logistics enterprises, as part of the service-oriented industry, is characterized by a "high satisfaction threshold" and data distribution imbalance, this study carries out the study of data distribution tuning by changing the inputs of the data volume of the pre-training model to Fine-tuning. The results show that the model achieves the maximum values of pre, acc and f1 when the data distribution imbalance is 30%-40%, which is where the best performance is obtained.

6 CONCLUSION

In this paper, we propose a pre-trained classification model using the BERT model to classify positive-neutral-negative sentiment scores of logistics enterprises, and all kinds of indexes can indicate its good classification performance. Meanwhile, we use the Fine-tuning means of the BERT model to study the tuning scheme when facing the problem of data imbalance, and the results obtained provide useful guidelines for the subsequent model and experimental optimization. The results obtained also provide useful guidelines for subsequent model and experimental optimization.

Based on the sentiment classification model proposed in this paper, all the text data crawled from the evaluation of logistics enterprises are analyzed for the prediction and classification of emotional tendency, and further theme extraction and cross-analysis, etc., so as to obtain the value information of the customer's satisfaction with the service and pain points of the logistics enterprises, which can be used as a reference for the selection of the logistics service providers, and at the same time provide a powerful support for the strategic decision-making and business improvement of the logistics enterprises.

REFERENCES

Feldman, R., Sanger, J., 2007. *The text mining handbook: advanced approaches in analyzing unstructured data.* Cambridge: Cambridge University Press.

Xu, G., Hu, X., Wang, Q., 2007. Research and implementation of Chinese word segmentation algorithm in text mining. *Computer Technology and Development*, (12), 122-124+172.

Hu, Y., Wu, H., Zhong, L., 2007. Research on lexical based feature extraction method in Chinese text classification. *Journal of Wuhan University of Technology*, (04), 132-135.

Hu, R., Rui, L., Zeng, P., et al. 2018. Text sentiment analysis: A review. In 2018 *IEEE 4th International Conference on Computer and Communications (ICCC)* (pp. 2283-2288). IEEE.

Li, J., Xu, Y., Xiong, H., et al., 2010. Chinese text emotion classification based on emotion dictionary. In 2010 *IEEE 2nd symposium on web society* (pp. 170-174). IEEE.

Hasan, A., Moin, S., Karim, A., et al., 2018. Machine learning-based sentiment analysis for twitter accounts. *Mathematical and computational applications, 23*(1), 11.

Dang, N. C., Moreno-García, M. N., De la Prieta, F., 2020. Sentiment analysis based on deep learning: A comparative study. *Electronics, 9*(3), 483.

Lai, Y., Zhang, L., Han, D., et al., 2020. Fine-grained emotion classification of Chinese microblogs based on graph convolution networks. *World Wide Web, 23*, 2771-2787.

Treiblmaier, H., Mair, P., 2021. Textual Data Science for Logistics and Supply Chain Management. *Logistics, 5*(3), 56.

Singh, S. P., Adhikari, A., Majumdar, A., et al., 2022. Does service quality influence operational and financial performance of third party logistics service providers? A mixed multi criteria decision making-text mining-based investigation. *Transportation Research Part E: Logistics and Transportation Review, 157*, 102558.

Hong, W., Zheng, C., Wu, L., et al., 2019. Analyzing the relationship between consumer satisfaction and fresh e-commerce logistics service using text mining techniques. *Sustainability, 11*(13), 3570.

Lim, M. K., Li, Y., Song, X., 2021. Exploring customer satisfaction in cold chain logistics using a text mining approach. *Industrial Management & Data Systems, 121*(12), 2426-2449.

Devlin, J., Chang, M. W., Lee, K., et al., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* preprint arXiv:1810.04805.

Vaswani, A., Shazeer, N., Parmar, N., et al., 2017. Attention is all you need. *In Advances in neural information processing systems* (Vol. 30).