

## 摘要

The review text of the e-commerce platform of agricultural products contains the rich and complex emotions of the product consumers, and reflects the user's consumer experience and views on the product. Compared with the features and advantages of English text segmentation structure, the analysis of Chinese sentiment polarity in complex context has always been a difficult problem in the field of deep learning. This paper explores the issue of sentiment polarity analysis of short texts of agricultural Chinese reviews by constructing three deep learning models of Text-CNN, Bi-LSTM and BERT fine-tuning. Experimental results show that, compared with the Bi-LSTM and BERT fine-tuning models, the Text-CNN model's accuracy, precision, recall and F1 value are all about 3 to 8 percentage points higher, confirming that Text-CNN is processing agricultural products Short text comment data has a high accuracy of sentiment classification, which can be used as a model for sentiment classification of agricultural product e-commerce sales review data, providing an important reference for agricultural product production and operation enterprises to accurately distinguish user feedback and enhance market competitiveness.

## 提出的方法

### 1. 算法

BERT (Bidirectional Encoder Representations from Transformers) is a pre-trained language representation model proposed by Devlin. BERT is a Transformer encoder that implements multi-layer bidirectional based on Transformer. Unlike other recent language representation models, BERT aims to pre-train deep bidirectional representations by jointly adjusting the context in all layers. The pre-trained BERT means that it can be fine-tuned through an additional output layer, and is suitable for the construction of the most advanced model for a wide range of tasks, without the need for major architectural modifications for specific tasks.

Bi-LSTM (Bi-directional Long-Short Term Memory) is a kind of LSTM, which is composed of two-way LSTM, which is a combination of two-way recurrent neural network and long- and short-term memory. At each time t, the input will be provided to these two neural networks in completely opposite directions at the same time, and the output will be jointly determined by the two unidirectional recurrent neural networks. In the field of sentiment analysis, due to the advantages that LSTM can choose to learn or forget, the LSTM neural network model can better capture the long-distance dependencies. And LSTM modeling cannot encode back-to-front problems, but Bi-LSTM can better capture the two-way semantic dependence.

Text-CNN was first proposed by Kim Y. The model consists of an input layer, a convolutional layer (feature extraction), a pooling layer (down-sampling), a fully connected layer (classification) and an output layer, and the word vector is constructed to solve related applications of NLP to sentiment analysis tasks. The core idea of Text-CNN is to use multiple different convolution kernels to extract the local key information in the text and accurately capture the local features of the text. The so-called local feature in the text is a sliding window composed of several words. The advantage of convolutional neural network is that it can automatically filter and combine the sliding window to obtain semantic information of different abstract levels.



Fig. 1. BERT fine-tuning classification model based on pretraining

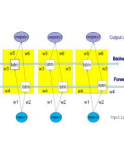


Fig. 2. Bi-LSTM neural network based on deep learning

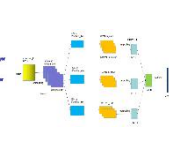
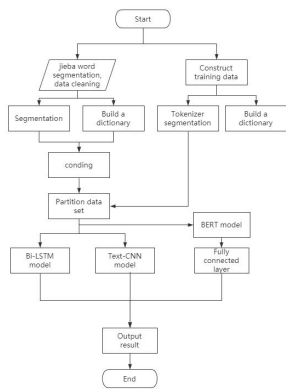


Fig. 3. Text-CNN convolutional neural network

### 2. 算法流程



## 实验

### 1. 数据介绍

This article uses the R language to build a crawler framework to crawl the agricultural product consumption review data of the e-commerce platform, including bananas, peaches, dragon fruit, potatoes, tomatoes, cucumbers, green vegetables and other agricultural products to build a total of 23418 agricultural product review data Data set for testing and verification of emotional binary classification

Table I  
The composition of agricultural product reviews data set

Dataset	Pos	Neg
Train	11910(80%)	6824(80%)
Test	2978(20%)	1706(20%)
Total	14888(63.58%)	8530(36.42%)

### 2. 实验结果

Table II  
Final result

Metric Model	Type	Accuracy /%	Precision /%	Recall /%	F1-score %
BERT	pos	95.39	94.02	96.77	95.38
	neg		96.79	94.05	95.40
Text-CNN	pos	99.92	99.90	99.95	99.92
	neg		99.95	99.90	99.93
Bi-LSTM	pos	93.60	91.70	95.62	93.62
	neg		95.59	91.65	93.58

The accuracy of the Text-CNN convolutional neural network reached 99.92%, which is significantly higher than the accuracy of 95.39% of the BERT fine-tuned classification model and the accuracy of 93.60% of the Bi-LSTM bidirectional long-short-term memory neural network. In the results of the training data set, the accuracy of the Bi-LSTM two-way long and short-term memory neural network is 93.60%, but the accuracy is a bit poor compared to the other two models. On the agricultural product review dataset with short text sentences, the classification model is inferior to the other two models in terms of accuracy, precision, recall, and F1 value. In addition, it can be seen that the accuracy, precision, recall, and F1 value of the Text-CNN classification model are about 3 to 8 percentage points higher than the other two models. The classification effect between the three models is obviously gradient. That is, the Text-CNN model is significantly better than the BERT fine-tuned classification model, and the BERT fine-tuned classification model is significantly better than the Bi-LSTM model. In summary, the sentiment classification effect of agricultural product e-commerce review data based on the Text-CNN convolutional neural network model is the best.

## 结论

Compared with traditional sentiment analysis methods, this article uses deep learning multiple models to explore the review data of agricultural product e-commerce platforms for sentiment classification comparison. In the end, the Text-CNN convolutional deep neural network model we tested got unexpected results: the accuracy of emotion classification reached 99.2%, and compared with the Bi-LSTM bidirectional long short-term memory neural network, it was about 6 percentage points higher. The comparison of the BERT classification model is about 4% higher, which proves the excellent sentiment classification performance of Text-CNN in short texts such as agricultural Chinese comment data. The research results of this article provide a valuable and practical reference in the field of sentiment classification of agricultural product review data.

## 主要参考文献

- [1] Zhou.H, Wu.G, "Research on Sentiment Analysis of Chinese E\_x0002\_Commerce Comments Based on Deep Learning", Journal of Physics Conference Series, vol. 1237, pp. 022002, 2019.
- [2] Xiong, Shufeng and Wang, Kuiyi and Ji, Donghong and Wang, Bingkun, "A Short Text Sentiment-Topic Model for Product Reviews", Neurocomputing, vol. 297, Mar. 2018.
- [3] Huang, Jiafeng and Xue, Yun and Hu, Xiaohui and Jin, Huixia and Lu, Xin and Liu, Zhihuang, "Sentiment analysis of Chinese online reviews using ensemble learning framework", 2019.
- [4] Liu, Li Zhen and Liu, Hao and Wang, Han Shi and Song, Wei and Zhao, Xin Lei, "Generating domain-specific affective ontology from Chinese reviews for sentiment analysis", Journal of Shanghai Jiaotong University (Science), vol. 1237, no. 1, pp. 32-37, 2015.

## 联系方式

联系人: 贺子康  
 手机: 18677387275  
 邮箱: 516471706@qq.com