Sentiment analysis based on food e-commerce reviews

Suoxiang Zhang*
College of Information and Electrical Engineering, China Agricultural University, 100083, Beijing, China
Author’s e-mail address:S20183081344@cau.com.cn

Abstract. Customer feedback is of great significance for food manufacturers such as wine to improve product quality. With the increase of consumer users, user feedback collection and classification methods in the traditional food industry have become very time-consuming. Aiming at this problem, this paper proposes a method of embedding TD-IDF weighted Word2vec word vector into a bidirectional long and short-term memory network based on the attention mechanism.

1. INTRODUCTION
The wine production industry is a highly competitive market, and it is developing very rapidly. Wine producers and some wine sellers use traditional user feedback forms, which often require a lot of effort to classify the review text uploaded by users. The feedback from users on Online wine shop on wine evaluations can be a good source for our sentiment analysis. Users who have purchased wine labels multiple times are more sensitive to the taste of wine, and the evaluation data they provide are more valuable to manufacturers. We crawled the comment text of the above-mentioned users. For the sentiment classification task of feedback evaluation text, although some effective models have been proposed, there are still some problems, such as: how to make Word2vec better reflect the key features of the text, how to effectively use context information and semantic features. In order to solve the above problems, and at the same time, a TF-IDF[1] weighted Word2vec[2] word vector embedding is proposed for the user evaluation text containing a large number of characteristics describing the taste, aroma, color and vintage of the wine. At the same time, the attention mechanism[3] is used to help BILSTM[4] to carry out the text before and after the text. According to the characteristics of the text, the model we proposed captures the important words with distinguishing degree in the text, so as to perform more effective emotion classification work.

2. MODEL
2.1. TF-IDF
Some text. The main idea of TF-IDF: If a word or phrase appears frequently in an article and rarely appears in other articles, it is considered that the word or phrase has good classification ability and is suitable for classification. In fact, TF-IDF is TF*IDF, TF term frequency (Term Frequency), IDF inverse document frequency (Inverse Document Frequency). Term Frequency (Term Frequency, TF) represents the frequency at which the keyword $w$ appears in the document $D_i$:

$$TF_{w,D_i} = \frac{\text{count}(w)}{|D_i|}$$  \hspace{1cm} (1)
IDF Inverse Document Frequency (Inverse Document Frequency), where \( N \) is the total number of all documents, \( I(w, D_i) \) indicates whether the document \( D_i \) contains keywords, if it contains it, it is 1, and if it does not contain it, it is 0. If the word \( w \) does not appear in all documents, the denominator in the IDF formula is 0; therefore, IDF needs to be smoothed:

\[
IDF_w = \log \frac{N}{\sum_{i=1}^{m} I(w, D_i)}
\]

(2)

2.2. WORD2VEC

Place the figure as close as possible after the point where it is first referenced in the text. If there is a large number of figures and tables it might be necessary to place some before their text citation. If a figure or table is too large to fit into one column, it can be centred across both columns at the top or the bottom of the page. The architecture of the Word2Vec model uses a neural network to take the text body as the input, and the text vector space as its output. The generated word vector is a low-dimensional space vector that captures the meaning of the word semantics. There are two types of Word2Vec architecture models, namely the Skip-Gram model and the continuous bag of words (CBOW) model.

The Skip-gram model is introduced to predict the context word based on the current word, while the CBOW model introduced is to predict the current word based on the context word. The Word2Vec model architecture, CBOW model and Skip-Gram model can be seen in Figure 1.

![Fig. 1. Structure of Word2vec model.](image)

2.3. ATTENTION MECHANISM

When using the bidirectional long and short-term memory network model to predict different tags, not all text words in the context make the same contribution. The attention mechanism captures the important parts of the sentence to enhance accuracy.

The emotional polarity of a sentence is not only related to contextual information, but also highly related to viewpoint terms and aspect terms. But given a sentence, not all context words have the same contribution to the semantics of the sentence. To solve this problem, the attention mechanism is used to increase their importance by giving them more weight, thereby extracting these more important words. The attention mechanism can highlight the impact of input on output and optimize the traditional model by calculating the attention probability distribution. As shown in Figure 2
2.4. **BILSTM**

LSTM is a special recurrent network model, which overcomes the gradient explosion problem of the RNN model in the training process. The bidirectional long short-term memory network (BiLSTM) consists of the following: Two independent LSTMs can merge information from two directions.

\[
X = \left[ x_{t-1} \right]_{t}
\]

\[
f_t = \sigma(W_f \cdot X + b_f)
\]  
\[
i_t = \sigma(W_i \cdot X + b_i)
\]  
\[
o_t = \sigma(W_o \cdot X + b_o)
\]  
\[
c_t = f_t \cdot c_{t-1} + i_t \cdot \tanh(W_c \cdot X + b_c)
\]  
\[
h_t = o_t \cdot \tanh(c_t)
\]

\(i_t, f_t, o_t, c_t\) represent the input gate, forget gate, output gate, memory cell at time \(t\), \(W_i, W_f, W_o\) and \(W_c\) represent the weight matrix corresponding to different control gates, \(b_i, b_f, b_o, b_c\) represent offset vector, and \(c_t\) represents the intermediate state of input, \(x_t\) represents the input vector at time \(t\), \(h_t\) represents the output result at time \(t\), \(\odot\) represents the dot multiplication operator.

The structure of BILSTM is shown in Figure 3 and the basic framework of the TF-IDF-weighted Word2vec word vector embedding based on the self-attention bidirectional long and short-term memory network model is shown in Figure 4.
3. EXPERIMENTS AND RESULTS

According to the customer information of multiple Online wine shop, we collected 25,514 reviews. At the same time, based on user ratings, we divided them into three categories, namely: good reviews, bad reviews and medium reviews. Equations should be centred and should be numbered with the number on the right-hand side.

Table.1 data set

| Reviews                                                                 | Ratings          |
|------------------------------------------------------------------------|------------------|
| 这是一款果味浓郁的葡萄酒，有着刚刚好的酸度，清新的水果味道，淡淡的柠檬味和淡淡的质感，准备好喝了。(This is a fruity wine with just the right acidity, fresh fruit flavor, light lemon flavor and light texture. Ready to drink.) | Good review      |
| 这款酒口感柔和，口感饱满，但不稳定。(This wine has a soft mouthfeel, full mouthfeel, but unstable.) | Medium review    |
| 口感相当清淡，不喜欢(The taste is quite light. I don’t like it)        | Bad review       |

3.1. Experiment setting

The parameter settings in the experiment are shown in Table 2.

Table.2 parameter settings

| Parameter        | Value |
|------------------|-------|
| max len          | 128   |
| embedding dim    | 128   |
| batch size       | 64    |
| epoch            | 10    |
| optimizer        | softmax |
| dropout          | 0.2   |

3.2. Evaluation Index

This article uses the accuracy (P), recall (R) and F-score of event detection to evaluate the performance of the model:

\[
Precision = \frac{T_p}{T_p + F_p}
\]  
(9)
\[
Recall = \frac{T_p}{T_p + F_N}
\]  
\[
F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}
\]

Tp indicates True Positives, Fp indicates False Positives and FN indicates False Negatives.

### 3.3. Experiment Results

In the same experimental environment, in order to verify the performance of the Our model, we use the BILSTM[5] and CNN[6] as a comparison experiment.

| Model      | P    | R    | F1    |
|------------|------|------|-------|
| Our Model  | 0.8359 | 0.8535 | 0.8446 |
| CNN        | 0.7707 | 0.7509 | 0.7607 |
| BILSTM     | 0.795  | 0.796  | 0.795  |

As shown in Table 3: Our Model has achieved good results and has great advantages compared with BILSTM and CNN.

### 4. CONCLUSION

Compared with the BILSTM and CNN algorithms, this algorithm has greater advantages and can complete the task of wine sentiment classification. In the future, we will further optimize the model based on the characteristics of Chinese wine texts.

### References

[1] You E S, Choi G H , Kim S H . Study on Extraction of Keywords Using TF-IDF and Text Structure of Novels[J]. Journal of the Korea Society of Computer & Information, 2015, 20(2): 439-441.

[2] Mikolov T , Chen K , Corrado G , et al. Efficient Estimation of Word Representations in Vector Space[J]. Computer Science, 2013.

[3] Devlin J , Chang M W , Lee K , et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding[J]. 2018.

[4] Hochreiter S, Schmidhuber J. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8):
1735-1780.

[5] Kiperwasser E, Goldberg Y. Simple and Accurate Dependency Parsing Using Bidirectional LSTM Feature Representations[J]. Transactions of the Association for Computational Linguistics, 2016, 4:313-327.

[6] Krizhevsky A, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks[C]// NIPS. Curran Associates Inc. 2012.