

KC-ISA: An Implicit Sentiment Analysis Model Combining Knowledge Enhancement and Context Features

Minghao Xu, Daling Wang*, Shi Feng, Zhenfei Yang, Yifei Zhang

Northeastern University, Shenyang, China

2071825@stu.neu.edu.cn, {wangdaling, fengshi}@cse.neu.edu.cn

2001851@stu.neu.edu.cn, zhangyifei@cse.neu.edu.cn

Abstract

Sentiment analysis has always been an important research direction in natural language processing. The research can be divided into explicit sentiment analysis and implicit sentiment analysis according to whether there are sentiment words in language expression. There have been many research results in explicit sentiment analysis. However, implicit sentiment analysis is rarely studied. Compared with explicit sentiment expression, implicit sentiment expression usually omits a lot of knowledge and common sense, and context also has an important impact on implicit sentiment expression. In this paper, we use a knowledge graph to supplement implicit sentiment expression and propose a novel **Implicit Sentiment Analysis** model combining **Knowledge enhancement and Context features** (dubbed KC-ISA). The KC-ISA model can effectively integrate external knowledge and contextual features by the coattention mechanism. Finally, we conduct experiments on the SMP2019 implicit sentiment analysis dataset. Moreover, to verify the generality of the model, we also conduct experiments on two common sentiment analysis datasets. The results on three datasets show that our proposed KC-ISA model can achieve better results on text sentiment analysis.

1 Introduction

With the rapid development of the Internet, the Internet has become the media form with the widest audience and the fastest reflection of social dynamics in today's society. Online social media has become an excellent platform in which all users can easily participate, share and communicate. Therefore, a large number of texts that reflect the user's subjective consciousness have begun to appear, and the texts with the user's sentiment tendencies are also increasing.

Text sentiment analysis is the process of analyzing, summarizing, and reasoning subjective senti-

ment texts. From the expression level of text, it can be divided into explicit sentiment text and implicit sentiment text according to whether the text contains explicit sentimental words. Liu (2012) firstly divided sentiment into subjective opinion and factual implication opinion. The former is a subjective statement that gives the polarity of sentiment, while the latter is an implicit expression of sentiment through an objective statement. Liao et al. (2019) defined implicit sentiment sentences as "language fragments that express subjective sentiment but do not contain explicit sentiment words". As the basic research of sentiment analysis, the analysis of explicit sentiment text has abundant relevant research results. However, the emotions reflected by people's behavior and feelings towards objective things are often rich and abstract and are usually expressed in subtle ways. For example, using objective statements to express one's sentiment, or using rhetorical expressions to describe them vividly, is called implicit sentiment expression. That means the implicit sentiment expression depends on something other than the sentiment words. So implicit sentiment analysis first needs to discover these dependencies. According to Li et al. (2021b), about 30% of sentences contain implicit sentiment.

By reviewing the relevant paper (Liao et al., 2019) and analyzing the data, we summarized that implicit sentiment sentences have the following three distinct dependencies:

(1) Context Dependency

Below are two paragraphs of implicit sentiment expression text E1 and E2, each of which contains two sentences S1 and S2.

E1-S1: This orange is only sold for 1 yuan a pound. E1-S2: It was so cheap!

E2-S1: This orange is only sold for 1 yuan a pound. E2-S2: It must be very bad!

E1-S1 and E2-S1 are the target sentences for sentiment analysis. In E1-S1, because the price

* Corresponding author

of oranges is low, the speaker has the intention to buy and expresses a positive sentiment. In E2-S1, because the price is too low, the speaker doubts the taste of orange and expresses a negative sentiment. The same sentence expresses different sentiments due to the difference between E1-S2 and E2-S2. Therefore, contextual information is an important factor affecting implicit sentiment expression. If we want to accurately analyze implicit sentiment expression, it is crucial to integrate contextual features.

(2) Sentiment Target Dependency

Let's look at the next two implicit sentiment texts.

E3: In the race, he runs like a cheetah.

E4: In the race, he runs like a turtle.

E3 and E4 are exactly the same except for the sentiment target, but the first sentence is positive and the second sentence is negative. The two sentences express completely opposite implicit sentiment only by the change of the sentiment target, indicating that the implicit sentiment is also closely related to the sentiment target.

(3) Knowledge and Common Sense Dependency

Short text reviews usually omit a lot of knowledge and common sense, and it is difficult to judge the true sentiment polarity from the text itself. For example, in E3 and E4, to accurately judge the sentiment of these two sentences, we also need to have certain knowledge and common sense, and we must know that cheetahs run fast, but turtles are slow. But the target sentence we want to analyze does not contain such background common sense. The external knowledge just makes up for these shortcomings. It contains the relationship between sentiment words and non-sentiment words and evaluation objects in different fields.

Therefore, to address the three distinct dependencies of implicit sentiment expression, we need model context, sentiment target, and external knowledge. Especially, for knowledge and common sense dependency, we introduce a knowledge graph as external common sense to supplement implicit sentiment sentences to enrich the expression of sentiments. Inspired by the Coattention (Xiong et al., 2017) mechanism, in this paper, we propose KC-ISA, a novel Implicit Sentiment Analysis model combining Knowledge enhancement and Context features, which can effectively integrate external knowledge and contextual features to com-

plement sentiment expressions. Moreover, by the pre-trained model, KC-ISA can utilize external knowledge and contextual features more effectively. Extensive experiments show that our proposed KC-ISA model can effectively improve the accuracy of implicit sentiment analysis.

In summary, the main contributions of this paper are as follows:

First, we propose KC-ISA model for three distinct dependencies of implicit sentiment analysis. Moreover, we combined KC-ISA model with the pre-trained model BERT as KC-ISA-BERT. KC-ISA model can effectively integrate contextual features and incorporate external knowledge into implicit sentiment analysis.

Second, we conduct experiments on implicit sentiment dataset, and the results show that our proposed model can effectively improve the accuracy of implicit sentiment analysis.

Third, considering the real review data is a mixture of explicit and implicit data, we further conducted experiments on two common sentiment analysis datasets. The experimental results further validate the effectiveness and general applicability of our KC-ISA model.

The datasets and code have been uploaded to <https://github.com/AnonymousColing2022/KC-ISA>

2 Related Work

Related work mainly includes two parts, one is implicit sentiment analysis, and the other is integrating external knowledge in natural language processing tasks.

2.1 Implicit Sentiment Analysis

Liu (2012) first classified sentiment analysis into explicit and implicit sentiment analysis. Liao et al. (2019) focused on the recognition of fact-implicit sentiment at the sentence level and proposed a multi-level semantic fusion method based on representation learning to learn recognition features. Greene and Resnik (2009) used grammatical structure to establish language-driven features and implicit sentimental associations for sentences without obvious sentimental indicators but still able to express sentiments or opinions, and used similarity calculation to improve the effect of text sentiment analysis. Wei et al. (2020) proposed a BILSTM model with multi-polar orthogonal attention for implicit sentiment analysis. Compared with

traditional single-attention models, multi-polar attention can identify the differences between words and sentimental orientations. Wang et al. (2021) proposed an implicit sentiment sentence discrimination method fused with context information and established a multi-level orthogonal attention C-MPOA model fused with context information. Zuo et al. (2020) obtained the features of implicit sentiment sentences and context through GCN, and proposed a context-specific heterogeneous graph convolutional neural network(CsHGCN) to solve the problem of missing sentimental words. Zhou et al. (2021) represented an event as the combination of its event type and the event triplet. Based on such event representation, they proposed a model with a hierarchical tensor-based composition mechanism to detect sentiment in text. The previous works mainly focused on the structure of the sentence itself and its context. To our knowledge, it is the first time that we introduced the knowledge graph as external knowledge into the implicit sentiment analysis task and then integrated the contextual features to enable the model to better understand the implicit sentiment contained in the sentence, thereby improving the accuracy of the implicit sentiment analysis.

2.2 Integrating External Knowledge

Not only in sentiment analysis, but in many natural language processing tasks, there are a lot of works on how to introduce external knowledge for knowledge enhancement. Wang et al. (2018) introduced the information of the knowledge graph in the sorting task of the recommendation system and established a Ripple Set through the user’s historical information, that is, the nodes corresponding to the knowledge graph, and each step will be propagated to the adjacent nodes, thereby improving the recommendation accuracy and variety. Liu et al. (2018) added the embedding of the input entities in the knowledge graph in the text matching task as additional supervision information to the model, which can well help the model achieve higher-level semantic understanding. Zhang et al. (2019) modified the original encoder structure of BERT and added entity embedding input. In this way, external knowledge was introduced in pre-training process to improve the effect of downstream tasks. Wang et al. (2020) introduced external knowledge in natural language inference tasks. They searched for the entities corresponding to the subject, predicate,

and object in the two sentences in the knowledge graph, found the path between the two corresponding entities, then obtained the representation of the corresponding knowledge through BILSTM, and finally combined it with the original text to infer. Instead of directly adding external knowledge, Sun et al. (2019) modified BERT’s mask strategy. Through entity-level mask and phrase-level mask, the model implicitly learns relevant knowledge and longer semantically dependent information, such as the relationship between entities, the properties of the entity, and the type of the events, which can make the model have better generalization and adaptability. There are also many related works on integrating external knowledge in natural language processing tasks, but less research on implicit sentiment analysis. To our knowledge, it is the first time that we have introduced both knowledge graph and contextual features into implicit sentiment analysis task, which perfectly matches the dependencies of implicit sentiment expression (context dependency, sentiment target dependency and knowledge and common sense dependency), which can significantly improve the accuracy of implicit sentiment analysis.

3 Proposed Model

We summarized three dependencies of implicit sentiment expression earlier, i.e. context dependency, sentiment target dependency, and knowledge and common sense dependency. Since there are no explicit sentiment words in implicit sentiment sentences as a guide, traditional sentiment analysis methods are not very effective. Therefore, we proposed KC-ISA model for the three distinct dependencies. The model can effectively model the relationship between the target sentence and the context, and incorporate external knowledge, which can effectively enrich the implicit sentiment expression of the sentence and improve the accuracy of implicit sentiment analysis.

3.1 Problem Definition

For a given sentence $S = \langle C, T \rangle$, where C is context and T is target sentence, our goal is to judge the sentiment polarity of the target sentence T based on the context C and external knowledge fusion when the target sentence has no sentimental words. For example, for the previously mentioned implicit sentiment texts E1 and E2, E1-S1 and E2-S1 are target sentences T , and E1 and E2 are the

contexts C of E1-S1 and E2-S1, respectively. The other two implicit sentiment texts E3 and E4 contain only target sentences, and we choose themselves as their contexts. In sections 3.2 and 3.3 below, we elaborate on our proposed model KC-ISA to solve this problem.

3.2 Structure of KC-ISA Model

The overall structure of KC-ISA model is shown in Figure 1. From the figure, we can see that KC-ISA model is divided into three parts: the Contextual Features Fusion module, the Knowledge Fusion module and the BiAffine module.

Contextual Features Fusion module

(1) **Input layer:** Segment the context C and the target sentence T , and then convert each word in the sentence into a vector through the pre-trained Baidu Encyclopedia embedding (Li et al., 2018) to obtain $S_c = \{w_{c1}, w_{c2}, \dots, w_{cn_c}\} \in R^{n_c \times d_e}$ and $S_t = \{w_{t1}, w_{t2}, \dots, w_{tn_t}\} \in R^{n_t \times d_e}$, where n_c is the maximum length of the context, n_t is the maximum length of the target sentence, and d_e is the dimension of embedding.

(2) **BILSTM layer:** The BILSTM is commonly used to deal with various tasks in natural language processing, and the use of BILSTM can effectively capture the relationship of each word in a sentence. We encode the obtained S_c and S_t sentences through the BILSTM layer and get $H_c = \{h_{c1}, h_{c2}, \dots, h_{cn_c}\} \in R^{n_c \times 2d}$ and $H_t = \{h_{t1}, h_{t2}, \dots, h_{tn_t}\} \in R^{n_t \times 2d}$ via formula (1),

$$h_n = BILSTM(h_{n-1}, h_{n+1}, w_n) \quad (1)$$

where w_n is trainable parameter, d is the number of BILSTM hidden layer units.

(3) **Coattention layer:** Coattention is a complex attention mechanism proposed by Xiong et al. (2017), and used to solve the problems in question answering. Taking reading comprehension as an example, Coattention is like a skill that we use when doing reading comprehension: read with questions. Look at the questions first, and then go to the targeted reading of the text to find the answers. In this module, we use the Coattention mechanism to extract effective information from the context to enrich the target sentence. The calculation process of Coattention is as follows:

First, convert the previously obtained H_c and H_t through a layer of nonlinear network to obtain the encoder vector via formula (2) and (3),

$$H_c = \tanh(W_c H_c + b_c) \quad (2)$$

$$H_t = \tanh(W_t H_t + b_t) \quad (3)$$

where W_c , W_t , b_c and b_t are trainable parameters.

Second, use H_c and H_t to calculate the correlation matrix L according to formula (4).

$$L = H_c(H_t)^T \in R^{n_c \times n_t} \quad (4)$$

Third, the correlation matrix L can be used to calculate the attention score A_c of each word in the context C to the target sentence T , and the attention score A_t of each word in the target sentence T to each word in the context C via formula (5) and (6).

$$A_c = \text{softmax}(L) \in R^{n_c \times n_t} \quad (5)$$

$$A_t = \text{softmax}(L^T) \in R^{n_t \times n_c} \quad (6)$$

Then we get the corrected vector C_c of the context after the attention calculation via formula (7).

$$C_c = A_c H_t \in R^{n_c \times 2d} \quad (7)$$

Moreover, we use the vector of the context H_c to concatenate the corrected vector C_c , and then obtain the corrected target sentence vector C_t after the attention calculation via formula (8),

$$C_t = A_t [H_c; C_c] \in R^{n_t \times 4d} \quad (8)$$

where C_t is the co-dependent representation of the context C and the target sentence T , also known as the Coattention vector of the context and the target sentence.

Finally, a layer of BILSTM is used to fuse the original target sentence H_t with the Coattention vector C_t for getting the final vector U_t of fusion context and target sentence information via formula (9) and (10).

$$u_{tn} = BILSTM(u_{n-1}, u_{n+1}, [c_{tn}; h_{tn}]) \quad (9)$$

$$U_t = \{u_{t1}, u_{t2}, \dots, u_{tn_t}\} \in R^{n_t \times 2d} \quad (10)$$

Knowledge Fusion module

We choose XLORE (Wang et al., 2013) bilingual encyclopedia knowledge graph as an external knowledge base. XLORE is the first large-scale knowledge graph that balances Chinese and English knowledge. It currently contains more than 2.3 million concepts, more than 500,000 relationships, and more than 25 million instances, covering most of the common sense concepts. In addition to

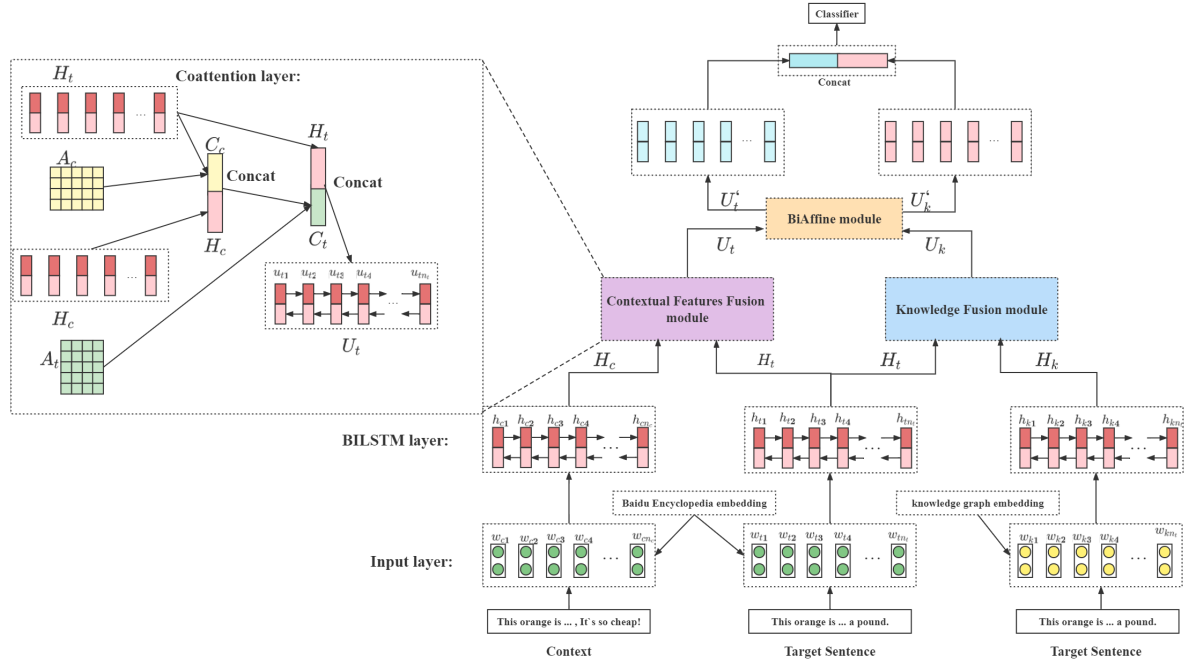


Figure 1: The overall structure of KC-ISA model

selecting the knowledge graph, it is also necessary to obtain the representation vector of knowledge. We use the openKE (Han et al., 2018) platform and select the TransE method to train the entities in XLORE and obtain the vector representation of knowledge entities.

(1) **Input layer:** The target sentence T is segmented, and then each word is searched for the corresponding vector in the trained knowledge graph representation vector. If there is no corresponding entity, this word vector is set to 0, and $S_k = \{w_{k1}, w_{k2}, \dots, w_{kn_t}\} \in R^{n_t \times d_k}$ is obtained, where n_t is the maximum length of the target sentence, d_k is the dimension of the trained knowledge graph vector.

(2) **BILSTM layer:** The function and calculation method of the BILSTM layer here are consistent with the Contextual Features Fusion module. After calculation, we can get $H_k = \{h_{k1}, h_{k2}, \dots, h_{kn_t}\} \in R^{n_t \times 2d}$, where d is the number of BILSTM hidden layer units.

(3) **Coattention layer:** The Coattention mechanism is also used to fuse the features of the external knowledge and the target sentence. The calculation method is basically the same as that of the Contextual Features Fusion module. The context representation vector H_c is replaced with the knowledge representation vector H_k . After the above calculation, the final fusion of external knowledge and

target sentence information vector U_k is obtained.

$$U_k = \{u_{k1}, u_{k2}, \dots, u_{kn_t}\} \in R^{n_t \times 2d} \quad (11)$$

BiAffine module

To effectively exchange relevant features between the Contextual Features Fusion module and Knowledge Fusion module, we adopt a mutual Bi-Affine (Li et al., 2021a) transformation as a bridge. We formulate the process as follows:

$$U_k' = \text{softmax}(U_k W_1 (U_t)^T) U_t \quad (12)$$

$$U_t' = \text{softmax}(U_t W_2 (U_k)^T) U_k \quad (13)$$

where W_1, W_2 are trainable parameters. Finally, we contact the obtained U_t' and U_k' vectors and input it into the classifier for classification, and obtain the final sentiment polarity of the target sentence.

3.3 KC-ISA-BERT

In natural language processing tasks, the pre-trained language model is a topic that cannot be avoided. In addition to the above-mentioned use of BILSTM to process vectors, we also combine the pre-trained model BERT to propose KC-ISA-BERT, and compare it with the native BERT model, which fully proves the effectiveness of KC-ISA mechanism combining context and external knowledge. The calculation process is mainly similar to

KC-ISA model, replacing the BILSTM layer with the BERT structure. The position of BERT in the KC-ISA-BERT is shown in Figure 2.

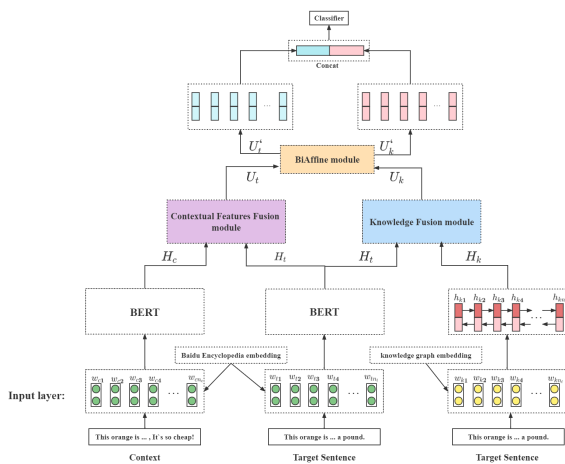


Figure 2: The overall structure of KC-ISA-BERT model

4 Experiment and Analysis

4.1 Datasets and Evaluation Metric

(1) **Datasets:** The main goal of this paper is to improve the accuracy of implicit sentiment analysis. We verify the effectiveness of KC-ISA model on the Chinese implicit sentiment analysis dataset of the "Tuers Cup" held at the 8th Social Media Processing Conference (SMP2019). The data involves Weibo, travel websites, forums, and other fields. Sentiment labels are divided into three types: positive, neutral, and negative. The detailed statistics of the dataset are presented in Table 1.

Subset	Positive	Neutral	Negative	Total
training	3391	6403	3561	13355
development	835	1577	877	3289
testing	835	1577	877	3289

Table 1: The proportion of the implicit sentiment dataset.

In addition, in order to verify the generality and effectiveness of our proposed model, two common datasets are selected for verification. The first is the NLPCC2014 Sentiment Classification with Deep Learning dataset, which is a binary classification without context. The specific division is shown in Table 2. The second is the NLPCC2014 Emotion Analysis in Chinese Weibo Texts dataset. This dataset is a 7-category dataset containing context. In this dataset, some sentences have two labels, which are primary emotion and secondary emotion.

We choose the primary emotion as the final label of the dataset. The division of this dataset is shown in Table 3.

Subset	Positive	Negative	Total
training	3350	3350	6700
development	825	825	1650
testing	825	825	1650

Table 2: The proportion of the Sentiment Classification with Deep Learning dataset.

(2) **Evaluation Metric:** We choose the $F1$ – score as the evaluation metric of the model. The specific calculation method is shown in formulas (14) and (15),

$$F1_{macro} = \frac{1}{N} \sum_{i=1}^N F1_i \quad (14)$$

$$F1_i = \frac{2 * P_i * R_i}{P_i + R_i} \quad (15)$$

where i is the sentiment polarity and P_i and R_i are the precision and recall, respectively, of instances with sentiment polarity i .

4.2 Implementation Details

In KC-ISA, the initialization embedding of the neural network selects Baidu Encyclopedia pre-training embedding (Li et al., 2018) with dimension 300. The number of BILSTM layers is 2 and the hidden unit size is 256. The learning rate is 0.001, the maximum length of the context is 300, and the maximum length of the target sentence is 128. The batch size is selected as 128, and the dropout is 0.6. In KC-ISA-BERT, we use the BERT-Adam optimizer, the dropout is set to 0.5, the number of neural units is 768, which are the default configurations of BERT, and the batch size is 16.

4.3 Baselines

We have selected some models that perform well in sentiment analysis as the baselines. The description and implementation details are as follows:

(1) **DPCNN (Johnson and Zhang, 2017):** DPCNN continuously deepens the network by stacking convolution modules and negative sampling layers, which can effectively extract dependencies in long-distance text. The dropout is set to 0.5, the number of convolution channels is 256, the batch size is 128, and the learning rate is 0.001.

(2) **TextCNN (Kim, 2014):** TextCNN has a simple structure, but it is very effective in processing

Subset	Like	Disgust	Fear	Surprise	Sadness	Happiness	Anger	Total
training	2854	2097	200	550	1660	1879	1272	10512
development	702	516	50	135	409	463	314	2589
testing	703	517	49	135	409	463	313	2589

Table 3: The proportion of the Emotion Analysis in Chinese Weibo Texts dataset.

Models	Implicit Sentiment datasets		Sentiment Classification dataset		Chinese Weibo Texts dataset	
	Accuracy	F1	Accuracy	F1	Accuracy	F1
DPCNN	0.759	0.735	0.773	0.773	0.513	0.438
TextCNN	0.762	0.721	0.764	0.764	0.576	0.545
TextRCNN	0.769	0.738	0.756	0.756	0.538	0.474
BILSTM	0.755	0.723	0.750	0.750	0.524	0.454
BILSTM+Attention	0.761	0.734	0.760	0.759	0.547	0.465
Transformer	0.760	0.728	0.740	0.740	0.543	0.450
MPOA(ELMo)*	-	0.680	-	-	-	-
MPOA(random)*	-	0.675	-	-	-	-
MPOA(TENCE)*	-	0.733	-	-	-	-
C-MPOA(ELMo)*	-	0.694	-	-	-	-
C-MPOA(random)*	-	0.690	-	-	-	-
C-MPOA(TENCE)*	-	0.746	-	-	-	-
our KC-ISA	0.786	0.755	0.776	0.775	0.687	0.608
BERT(target sentence only)	0.803	0.787	0.822	0.822	0.648	0.595
MPOA(BERT)*	-	0.773	-	-	-	-
our KC-ISA-BERT	0.836	0.817	0.828	0.828	0.724	0.687

Table 4: The performance of implicit and explicit sentiment analysis..

short text sentiment analysis. The dropout is set to 0.5, the number of convolution channels is 128, the size of the convolution kernel is a mixed convolution of 2 and 3, the batch size is 128, and the learning rate is 0.001.

(3) **TextRCNN (Lai et al., 2015)**: TextRCNN uses a bidirectional recurrent structure to obtain contextual information, which can reduce noise more than traditional neural networks, and can preserve word order in a large range when learning text expressions. Then, TextRCNN uses the maximum pooling layer to obtain important parts of the text, and automatically determine which feature plays a more important role in the text classification process. The dropout is 0.5, the learning rate is 0.001, the number of hidden layer units is 256, and the number of hidden layers is 2.

(4) **BILSTM (Graves and Schmidhuber, 2005)**: BILSTM is widely used in natural language processing tasks. In this paper, the parameters of the BILSTM model are dropout of 0.5, learning rate of 0.001, the number of hidden layer units is 256, and the number of hidden layers is 2.

(5) **BILSTM+Attention (Zhou et al., 2016)**: BILSTM with an attention mechanism has been the most popular and effective model in recent years. The parameters of the BILSTM+Attention model are dropout of 0.5, learning rate of 0.001, the num-

ber of hidden layer units is 256, and the number of hidden layers is 2.

(6) **Transformer (Vaswani et al., 2017)**: Transformer is a model structure that avoids recurrent and completely relies on the attention mechanism to model the global dependencies of input and output. Because the modeling of dependencies relies entirely on the attention mechanism, the attention mechanism used by the Transformer is called self-attention. The number of hidden layer units in the Transformer model is 1024, the last hidden size is 512, the number of attention heads is 5, and the number of layers of the encoder is 2.

(7) **MPOA (Wei et al., 2020)**: MPOA is a BILSTM model with multi-polarity orthogonal attention. Compared with traditional single attention model, multi-polarity attention can identify the differences between words and sentiment orientation, which can be seen as an important feature of implicit sentiment analysis.

(8) **C-MPOA (Wang et al., 2021)**: C-MPOA is a method for identifying implicit sentiment sentences with contextual information. First, the model embed implicit sentiment sentence representations by using a multi-polar orthogonal attention representation model. Second, multi-polarity attention layers are established by integrating context and modeling the context information to mine the key information

in the context. Finally, the contextual information representation and the implicit sentiment sentence representation are spliced together to make up for the lack of information in the implicit sentiment sentence itself.

(9) **BERT (Devlin et al., 2019)**: BERT is a pre-trained language representation model. It no longer uses the traditional one-way language model or the method of shallow splicing two one-way language models for pre-training as before but uses a new masked language model, so that it can generate deep bidirectional language representations. BERT has an excellent performance in various natural language processing tasks.

4.4 Experiment and Analysis

To evaluate KC-ISA model, we used the $F1 - score$ as the main evaluation metric. The main experimental results are reported in Table 4. The experimental results with * represent references to other papers. Since we could not obtain the test set of the SMP2019 dataset, we divided the training set and the development set in equal proportions, so the dataset is slightly different from that used by MPOA and C-MPOA. Our KC-ISA achieves the best results on implicit sentiment analysis, with an $F1 - score$ of 0.755. At the same time, the two common sentiment datasets have also been significantly improved, especially the Chinese Weibo Texts dataset containing context, which has been improved from 0.545 of TextCNN to 0.608. There are improvements on all three datasets, proving the generality of KC-ISA model.

In the use of the pre-trained model, we directly concatenated the context on both sides of the target sentence, and found that the effect is worse than directly analyzing the target sentence. It is speculated that the direct concatenation of the context will introduce noise. Therefore, we also conducted an experiment on the impact of the size of the concatenation of the context on the analysis of the target sentence. The experimental results are shown in Table 5. It can be seen from Table 5 that if the context is directly concatenated into both sides of the target sentence without processing, noise will be introduced, which affects the analysis of the target sentence, and as the length of contexts increases, the $F1 - score$ is lower. So we choose BERT that only contains the target sentence as the baseline. From the bottom three rows of Table 4, we can see that our KC-ISA-BERT has a significant improve-

ment over native BERT, especially for sentences with context, and the $F1 - score$ is increased from 0.595 to 0.687, indicating that our KC-ISA-BERT can effectively utilize external knowledge and contextual features.

Models	Accuracy	F1
BERT(target sentence only)	0.803	0.787
BERT(contains one sentence of context)	0.785	0.760
BERT(contains two sentences of context)	0.789	0.759
BERT(contains three sentences of context)	0.791	0.756

Table 5: The impact of the size of the splicing context on implicit sentiment dataset.

Models	Accuracy	F1
BILSTM	0.755	0.723
KC-ISA(Knowledge Fusion module only)	0.760	0.730
KC-ISA(Contextual Features Fusion module only)	0.780	0.748
KC-ISA(Without BiAffine module)	0.780	0.753
KC-ISA	0.786	0.755

Table 6: Experimental results of ablation study on implicit sentiment dataset.

4.5 Ablation Study

To further investigate the role of modules in KC-ISA model, we conducted extensive ablation studies on the implicit sentiment dataset. The results are reported in Table 6. KC-ISA (Knowledge Fusion module only) is a combination of BILSTM and Knowledge Fusion module. It can be seen that integrating external knowledge can bring a 0.007 $F1 - score$ improvement. KC-ISA (Contextual Features Fusion module only) is a combination of BILSTM and Contextual Features Fusion module. It can be seen that integrating contextual features can bring an $F1 - score$ improvement of 0.025, which is the largest improvement among all modules. KC-ISA (Without BiAffine module) is a complete KC-ISA model without BiAffine module, which reduces the interaction between external knowledge and context, which will cause a certain decline. Overall, our KC-ISA with all modules achieves the best performance.

4.6 Discussion

Now there are some review data that contain both positive and negative sentiments, so there are certain limitations to summarizing the sentiments of a sentence with one label. For example, "The scenery here is like the West Lake, but I waited in line for two hours to buy tickets". The first sentence expresses an implicit positive sentiment, i.e.

the affirmation of the scenery. The second sentence expresses an implicit negative sentiment. For explicit sentiment analysis, there have been many research results. However, for implicit sentiment analysis, it involves different contexts, especially the dependency of different external knowledge. Therefore, our next research endeavour is to extract multiple, more fine-grained implicit sentimental tags in a sentence on the one hand, and extract the aspects corresponding to the implicit sentiments in a sentence on the other hand, so that the model can better identify the various emotions contained in the text.

Secondly, for potential ethical implications and issues, different languages have different expression habits, involving cultural backgrounds, religious beliefs, and even ethics. Especially, when applying the model to individual's requirements such as conversations, chats, and psychological comfort, except for above common factors, there are personal language biases, so the issues need to be considered more.

5 Conclusion

In this paper, we summarized three distinct dependencies of implicit sentiment analysis: context dependency, sentiment target dependency, and knowledge and common sense dependency. And the XLORE bilingual encyclopedia knowledge graph was selected as an external knowledge base to supplement a large amount of common sense omitted in implicit emotional sentences. Finally, we proposed KC-ISA model for the distinct dependencies of implicit sentiment analysis, and combined the pre-trained model BERT to proposed KC-ISA-BERT, which can effectively integrate contextual features and external knowledge. To verify the effectiveness and generality of the model, experiments are conducted on an implicit sentiment analysis dataset and two common sentiment analysis datasets, and the experiments show that our model outperforms baselines.

Acknowledgements

We would like to thank the reviewers for their constructive comments. The work was supported by National Natural Science Foundation of China (62172086, 61872074, 62106039, 62272092)

References

- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. [BERT: pre-training of deep bidirectional transformers for language understanding](#). In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 4171–4186. Association for Computational Linguistics.
- Alex Graves and Jürgen Schmidhuber. 2005. [Frameworkwise phoneme classification with bidirectional LSTM and other neural network architectures](#). *Neural Networks*, 18(5-6):602–610.
- Stephan Greene and Philip Resnik. 2009. [More than words: Syntactic packaging and implicit sentiment](#). In *Human Language Technologies: Conference of the North American Chapter of the Association of Computational Linguistics, Proceedings, May 31 - June 5, 2009, Boulder, Colorado, USA*, pages 503–511. The Association for Computational Linguistics.
- Xu Han, Shulin Cao, Xin Lv, Yankai Lin, Zhiyuan Liu, Maosong Sun, and Juanzi Li. 2018. [Openke: An open toolkit for knowledge embedding](#). In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018: System Demonstrations, Brussels, Belgium, October 31 - November 4, 2018*, pages 139–144. Association for Computational Linguistics.
- Rie Johnson and Tong Zhang. 2017. [Deep pyramid convolutional neural networks for text categorization](#). In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 1: Long Papers*, pages 562–570. Association for Computational Linguistics.
- Yoon Kim. 2014. [Convolutional neural networks for sentence classification](#). In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar; A meeting of SIGDAT, a Special Interest Group of the ACL*, pages 1746–1751. ACL.
- Siwei Lai, Liheng Xu, Kang Liu, and Jun Zhao. 2015. [Recurrent convolutional neural networks for text classification](#). In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA*, pages 2267–2273. AAAI Press.
- Ruifan Li, Hao Chen, Fangxiang Feng, Zhanyu Ma, Xiaojie Wang, and Eduard H. Hovy. 2021a. [Dual graph convolutional networks for aspect-based sentiment analysis](#). In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, (Volume 1:*

- Long Papers*), *Virtual Event, August 1-6, 2021*, pages 6319–6329. Association for Computational Linguistics.
- Shen Li, Zhe Zhao, Renfen Hu, Wensi Li, Tao Liu, and Xiaoyong Du. 2018. [Analogical reasoning on chinese morphological and semantic relations](#). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 2: Short Papers*, pages 138–143. Association for Computational Linguistics.
- Zhengyan Li, Yicheng Zou, Chong Zhang, Qi Zhang, and Zhongyu Wei. 2021b. [Learning implicit sentiment in aspect-based sentiment analysis with supervised contrastive pre-training](#). In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*, pages 246–256. Association for Computational Linguistics.
- Jian Liao, Suge Wang, and Deyu Li. 2019. [Identification of fact-implied implicit sentiment based on multi-level semantic fused representation](#). *Knowl. Based Syst.*, 165:197–207.
- Bing Liu. 2012. *Sentiment Analysis and Opinion Mining*. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers.
- Zhenghao Liu, Chenyan Xiong, Maosong Sun, and Zhiyuan Liu. 2018. [Entity-duet neural ranking: Understanding the role of knowledge graph semantics in neural information retrieval](#). In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers*, pages 2395–2405. Association for Computational Linguistics.
- Yu Sun, Shuohuan Wang, Yu-Kun Li, Shikun Feng, Xuyi Chen, Han Zhang, Xin Tian, Danxiang Zhu, Hao Tian, and Hua Wu. 2019. [ERNIE: enhanced representation through knowledge integration](#). *CoRR*, abs/1904.09223.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. [Attention is all you need](#). In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pages 5998–6008.
- Hongwei Wang, Fuzheng Zhang, Jialin Wang, Miao Zhao, Wenjie Li, Xing Xie, and Minyi Guo. 2018. [Ripplenet: Propagating user preferences on the knowledge graph for recommender systems](#). In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management, CIKM 2018, Torino, Italy, October 22-26, 2018*, pages 417–426. ACM.
- S. Wang, M. Wang, J. Liao, and X. Chen. 2021. [An implicit sentiment sentence identification method base on context information](#). *School of Computer Information Technology, Shanxi University, Taiyuan 030006, China*, pages 1–7.
- Zhigang Wang, Juanzi Li, Zhichun Wang, Shuangjie Li, Mingyang Li, Dongsheng Zhang, Yao Shi, Yongbin Liu, Peng Zhang, and Jie Tang. 2013. [Xlore: A large-scale english-chinese bilingual knowledge graph](#). In *Proceedings of the ISWC 2013 Posters & Demonstrations Track, Sydney, Australia, October 23, 2013*, volume 1035 of *CEUR Workshop Proceedings*, pages 121–124. CEUR-WS.org.
- Zikang Wang, Linjing Li, and Daniel Zeng. 2020. [Knowledge-enhanced natural language inference based on knowledge graphs](#). In *Proceedings of the 28th International Conference on Computational Linguistics, COLING 2020, Barcelona, Spain (Online), December 8-13, 2020*, pages 6498–6508. International Committee on Computational Linguistics.
- Jiyao Wei, Jian Liao, Zhenfei Yang, Suge Wang, and Qiang Zhao. 2020. [Bilstm with multi-polarity orthogonal attention for implicit sentiment analysis](#). *Neurocomputing*, 383:165–173.
- Caiming Xiong, Victor Zhong, and Richard Socher. 2017. [Dynamic coattention networks for question answering](#). In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net.
- Zhengyan Zhang, Xu Han, Zhiyuan Liu, Xin Jiang, Maosong Sun, and Qun Liu. 2019. [ERNIE: enhanced language representation with informative entities](#). In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, pages 1441–1451. Association for Computational Linguistics.
- Deyu Zhou, Jianan Wang, Linhai Zhang, and Yulan He. 2021. [Implicit sentiment analysis with event-centered text representation](#). In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*, pages 6884–6893. Association for Computational Linguistics.
- Peng Zhou, Wei Shi, Jun Tian, Zhenyu Qi, Bingchen Li, Hongwei Hao, and Bo Xu. 2016. [Attention-based bidirectional long short-term memory networks for relation classification](#). In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, ACL 2016, August 7-12, 2016, Berlin, Germany, Volume 2: Short Papers*. The Association for Computer Linguistics.
- Enguang Zuo, Hui Zhao, Bo Chen, and Qiuchang Chen. 2020. [Context-specific heterogeneous graph convolutional network for implicit sentiment analysis](#). *IEEE Access*, 8:37967–37975.