

Sentiment Classification of Chinese Short Text by Global Information Enhancement and Dependency Graph Convolution

Long Jiang¹, Xiaodong Zhu^{1,2*}, Hao Cui², Jue Wang³

¹ Business School, University of Shanghai for Science and Technology, Shanghai, China

² Department of Artificial Intelligence, University of Shanghai for Science & Technology, Shanghai, China

³ University of Southern California, Los Angeles, California, America

Email Address

zhuxd@usst.edu.cn (Xiaodong Zhu)

*Correspondence: zhuxd@usst.edu.cn

Received: 29 December 2025; **Accepted:** 20 January 2026; **Published:** 6 February 2026

Abstract:

Text classification is one of the core problems in text mining, and it is also a basic and important task in the field of natural language processing. However, short text classification is faced with many challenges due to sparse text, non-standard text and limited context information. Aiming at the problems that dependency between words is not fully utilized and global semantic information is not fully mined in short text classification tasks, A Chinese short text sentiment classification model combining Global Information Enhancement and Dependency Graph Convolution (GI-DGC) is proposed. Firstly, the model uses Bidirectional gate recurrent unit (BiGRU) to learn context information in the text and obtain feature vectors containing global information. Secondly, dependency tree structure is used to extract text features, and the contribution weights of embedding representations of different dependency paths to sentiment classification tasks are considered, so as to make better use of the syntactic structure information of text. Finally, the Graph convolutional network (GCN) is used to model both the global information and the dependency syntactic information, so as to obtain a more comprehensive and adequate representation of the emotional features. Experimental results on three Chinese short text data sets show that the GI-DGC model can effectively improve the accuracy of text classification. In addition, the model still shows good performance when processing long text, which reflects the model has strong generalization ability.

Keywords:

Short Text Classification, Global Information, Syntactic Dependency, Graph Convolutional Network (GCN), Bidirectional Gate Recurrent Unit (BiGRU)

1. Introduction

With the rise of e-commerce, online shopping has become a mainstream choice for people. After purchasing products, users will publish personal comments on the platform, which include users' feelings and opinions on the use of products and reflect

their emotional preferences. Therefore, the sentiment analysis of these comment texts can help merchants better understand the preferences and needs of users, and also provide other users with the basis for purchasing decisions [1].

Text sentiment analysis, also known as opinion mining [2], is a key research area in natural language processing (NLP). Product reviews are typically short—ranging from ten to dozens of words—and exhibit randomness and informality, leading to challenges such as semantic ambiguity, sparse features, and limited contextual information. This makes sentiment analysis particularly difficult. In recent years, deep learning has achieved significant success in NLP, with popular architectures including Convolutional Neural Networks (CNN) [3], and Recurrent Neural Networks (RNN) [4]. Compared to traditional machine learning, deep learning enables automatic feature extraction and end-to-end learning, reducing human intervention and associated errors. However, despite their efficiency and automatic feature learning capabilities, deep learning-based sentiment analysis methods often overlook syntactic structures and dependency relations in sentences, and perform poorly on short texts.

Parsing is a natural language processing technique used to analyze the syntactic structure and grammatical relationships within sentences. Common methods include dependency parsing, statistical parsing, deep learning based parsing, and phrase structure parsing [5], each with its own advantages and suitable for different scenarios. Dependency parsing in particular identifies modifiers of sentiment words by analyzing word level dependency relations and has been widely integrated into sentiment analysis tasks. Although it has yielded promising results, it primarily focuses on lexical and grammatical structures while often neglecting deeper semantic information that significantly affects sentiment interpretation. For example, in the sentence “This fresh milk is really delicious”, dependency parsing as shown in Figure 1 captures syntactic links but may fail to uncover nonlinear semantic relationships if used only for rule based feature extraction. Graph Convolutional Networks (GCNs) [6] address this to some extent by processing graph structured data such as dependency trees through convolution operations that generate enriched node representations by aggregating features from neighboring nodes. As a result, GCNs have become a prominent method in NLP and are widely applied in sentiment classification. However, they typically operate on local syntactic contexts and often ignore long range word dependencies and global textual information, which can limit their ability to accurately capture complex sentiment semantics [7].

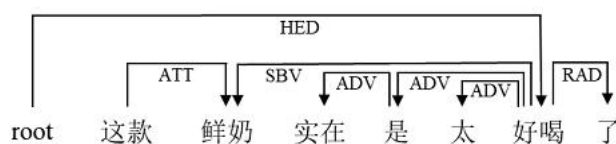


Figure 1. The Image of Dependency Parsing.

To address these limitations, this paper proposes GI-DGC, a Chinese short text sentiment classification model that integrates Global Information Enhancement and Dependency Graph Convolution. It first embeds text using Word2Vec and extracts global semantic features with BiGRU, then fuses them for enhanced representation. Simultaneously, dependency parsing constructs a syntactic tree, which is converted into a graph for GCN processing. GCN jointly models global context and syntactic dependencies, and the combined features are fed into a Softmax classifier for sentiment prediction. The main contributions are summarized in three points.

(1) This paper reframes text classification as graph classification by proposing a novel Chinese short text sentiment model that integrates global information enhancement and dependency graph convolution. The architecture consists of four layers: text representation, global enhancement, dependency graph convolution, and emotion classification. By leveraging GCN, the model jointly encodes global context and syntactic dependencies to produce richer and more comprehensive feature representations.

(2) On one hand, GI-DGC employs BiGRU to extract global semantic features; on the other hand, it uses dependency parsing to construct a text dependency graph and derive graph convolutional features. By integrating these two types of features, GI-DGC enhances text classification performance.

(3) Based on three short text Chinese review datasets, GI-DGC was compared with multiple baseline models for binary sentiment test, and the results showed the superiority of the model.

2. Related Work

The proposed dependency graph neural network markedly differs from conventional text classification methods by converting text into a graph structure and using a graph neural network for classification, enhanced with syntactic dependency information. This enables more effective capture of textual structure and relations while efficiently handling large-scale data.

2.1. Research on traditional machine learning

In the early research in the field of emotion classification, traditional machine learning methods are usually used to classify emotions. The core of these traditional methods is feature engineering and classifier design. In feature engineering, word bag model and TF-IDF method are usually used to convert text into vector representation, so as to facilitate processing by machine learning algorithm. In terms of classifiers, commonly used algorithms include Naive Bayes[8], Support Vector Machines (SVM), etc.[9]. Zhang et al.[10] proposed a generalized Bayesian fake news detection method that combines web crawling, HITS-based link analysis, data cleaning, and multimodal Transformer features, with greedy selection optimizing the classifier for prior and evidence estimation. Kielbasa et al.[11] employed Word2Vec for text vectorization and an SVM with an RBF kernel to classify Exchange Information Requirements (EIRs) according to ISO 19650 categories.

2.2. Deep learning research

In recent years, deep learning has enabled end-to-end text classification without manual feature engineering by directly taking raw text as input and using neural networks to automatically learn features and classifiers. For instance, convolutional neural networks (CNNs) capture local textual patterns, while recurrent neural networks (RNNs) model contextual and semantic dependencies. Attention-based models like Transformer and BERT further improve classification accuracy and robustness by effectively handling textual ambiguity. Kim et al. [3] demonstrated that a word-level CNN with pre-trained embeddings achieves competitive performance in sentence classification. Alfarhood [12] developed Soutcom, a cloud-based real-time sentiment analysis system using a custom biLSTM trained on a sports-specific Arabic dataset to analyze fan sentiments from X (formerly Twitter) and assess team

satisfaction. Hua et al. [13] proposed SHGCN, a short-text classification method based on heterogeneous graph convolutional networks that jointly model text, entities, and words. While these approaches eliminate the need for feature engineering and scale well to large, complex datasets, they often require substantial data and computational resources and suffer from limited model interpretability.

2.3. Dependency Parsing research

Dependency parsing is a natural language processing technique that analyzes text into a dependency syntax tree to reveal grammatical relationships between words. In recent years, it has become widely adopted across various NLP applications. Zhang et al. [14] proposed a sentence-level emotion analysis model based on dependency graph convolutional networks, leveraging the grammatical roles of words to improve sentiment classification accuracy. Ren et al. [15] introduced MTSA-CG, a Mongolian text sentiment analysis model combining ConvBERT and Graph Attention Networks to address limited annotated data and insufficient modeling of syntactic dependencies and topological structures, thereby enhancing performance. Wang et al. [16] applied dependency parsing to financial news sentiment analysis, effectively identifying emotional polarity and achieving strong results in market price prediction. Overall, dependency parsing enables more accurate capture of emotional, thematic, and semantic information, significantly boosting the effectiveness of NLP tasks.

2.4. Figure Convolutional Neural network research

CNN performs well in processing data such as images and videos, while RNN is widely used for sequential data such as text, speech, and time series. However, both are limited to Euclidean domains and fail to capture implicit dependency relations between sentence components. To address this, Kipf et al. [6] introduced Graph Convolutional Networks (GCN), which extend convolution operations to graph-structured data by leveraging information propagation across nodes and edges, thereby effectively modeling node dependencies and local graph structures. GCNs have since been applied in sentiment analysis to exploit syntactic information: Li et al. [17] proposed AMS-GCN, an adaptive marker segmentation GCN for aspect-level sentiment analysis that learns segmentation patterns to boost accuracy; Sun et al. [18] developed CDT, a CNN model based on dependency syntax trees that integrates contextual and syntactic information; and Yin et al. [19] introduced DPG-LSTM, which uses dependency parse trees to guide a hybrid attention mechanism, re-weighting high-dependency words before feeding them into an LSTM+GCN to strengthen semantic links in social media text. In summary, GCN provides a powerful framework for handling non-Euclidean and graph-structured data, offering novel approaches for advancing natural language processing and related fields.

3. GIE-DGC Model

Input a comment short text statement $S=\{w_1, w_2, \dots, w_n\}$ to be classified, where w_i represents the i -th word in the text sequence S , n represents the number of words contained in S , and the emotional polarity to be judged is $P_s=\{p_1, p_2, \dots, p_m\}$. The core of the task of emotion analysis of short text is to determine the emotion polarity of S to be analyzed. Therefore, the formal definition of this task can be described as follows: for a given short text S and set of emotion categories P , the GI-DGC model is used to calculate the likelihood probability distribution of each category in the set of

predefined emotion categories, and the emotion polarity of the category with the highest probability is found. The specific expression is as follows:

$$\text{GIE-DGC}: (S,P); P_s = \{p_1, p_2, \dots, p_m\} \quad (1)$$

The overall architecture of GI-DGC text classification model is shown in Figure 2, which is mainly composed of four parts, namely text presentation layer, global information enhancement layer, dependency graph convolution layer and emotion classification layer.

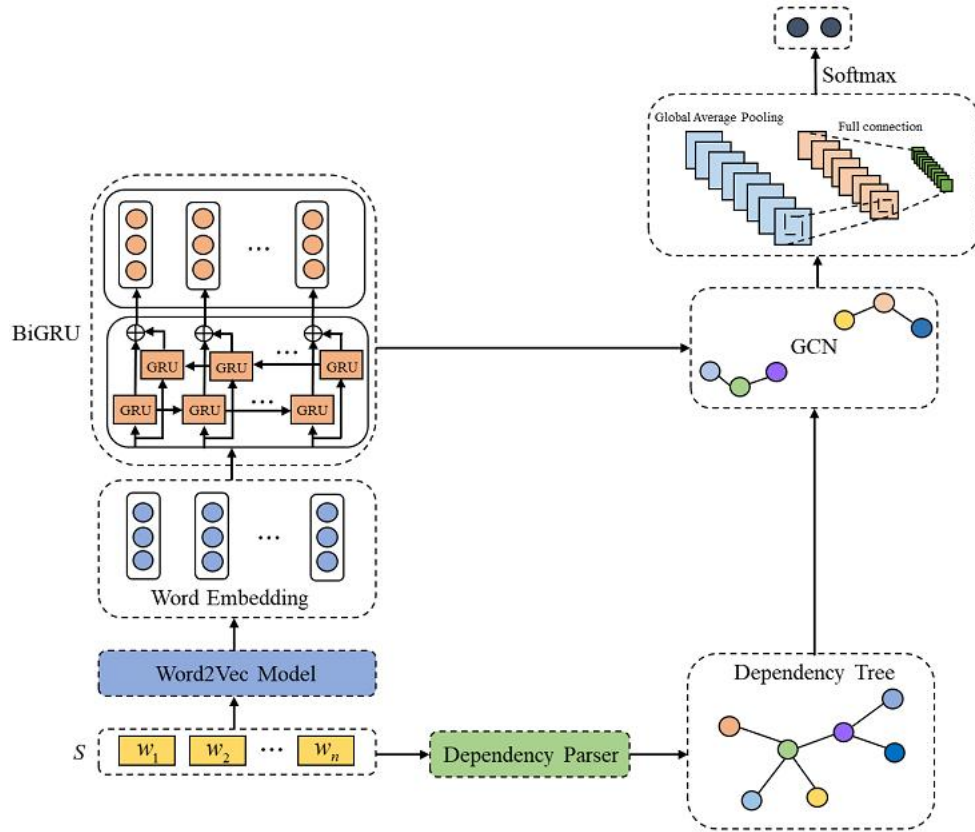


Figure 2. GIE-DGC model framework.

3.1. Text presentation layer

The purpose of the text presentation layer is to embed each word in the input text into a two-dimensional word vector space. In this paper, Word2Vec word embedding algorithm[22] is used to conduct unsupervised training on short review texts $S = \{w_1, w_2, \dots, w_n\}$ and generate feature word vectors $w_j = \{w_j^1, w_j^2, \dots, w_j^i\}$. Where, $w_j^i \in \mathbb{R}^{d_w}$, and i is the text length and d_w is the dimension of the w_j word vector.

3.2. Global Information Enhancement Layer

Gated Recurrent Units (GRUs) are a simplified RNN variant that processes sequential data using gating mechanisms to control information flow and retention. Compared to other RNN variants, GRUs have a simpler memory structure, require fewer computational resources and less time, and effectively mitigate vanishing gradient and long-term dependency issues in backpropagation.

The GRUs consists of two gates, namely, the reset gate and the update gate. The calculated expression is as follows:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (2)$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (3)$$

Where, r_t represents the reset gate, z_t represents the update gate, w_r and w_z represents the weight matrix. x_t represents the input of t , h_{t-1} represents the output of $t-1$, σ represents the sigmoid function. The final output of the GRUs is determined by both reset gate and update gate. The calculated expression is as follows:

$$\tilde{h}_t = \tanh(W \cdot [r_t \times h_{t-1}, x_t]) \quad (4)$$

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times \tilde{h}_t \quad (5)$$

Where, $\tilde{h}_t$ represents the candidate hidden layer, h_t represents the hidden layer, and W represents the weight matrix.

Grus only consider one-way text information, and cannot process the relationship between before and after moments in text classification, which is easy to cause information omission. In this paper, BiGRU is used to learn text context information by considering both positive and reverse order, which can better mine the context global feature information in sequence data, achieve the purpose of global information enhancement, and improve the effect of text classification. The calculated expression is as follows:

$$h_t^{\text{forward}} = \text{GRU}(x_t, h_{t-1}^{\text{forward}}) \quad (6)$$

$$h_t^{\text{backward}} = \text{GRU}(x_t, h_{t-1}^{\text{backward}}) \quad (7)$$

$$H_t = h_t^{\text{forward}} \oplus h_t^{\text{backward}} \quad (8)$$

Where, h_t^{forward} represents the semantic expression obtained by the forward GRUs, h_t^{backward} represents the semantic expression obtained by the backward GRUs, H_t represents the final semantic expression of the output of BiGRU, $\oplus$ is the connection symbol.

3.3. Convolutional layer of dependency graph

The Dependency Graph Convolutional Network (DGNN) converts text into a dependency syntax tree, represented as a graph ($G = (V, E)$), where each node (V) corresponds to a word and each edge (E) denotes a dependency relation between words, encoded as 0 or 1 in an adjacency matrix. In this work, the syntactic dependency tree is generated using Harbin Institute of Technology's Language Technology Platform (LTP) [23]. For example, the sentence "This fresh milk is so delicious" yields the dependency tree and adjacency matrix shown in Figure 3, where directed edges indicate dependencies labeled by type (e.g., root: root node; HED: head relation; ATT: attributive; SBV: subject-predicate; ADV: adverbial; RAD: right adjunct). A graph convolutional network then performs convolution over (G), aggregating each node's features with those of its neighbors to produce word-level embeddings.

$$H^{(l+1)} = \text{Relu}(\tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2} H^l W^l + b^l) \quad (9)$$

Where, Relu represents the activation function; $\tilde{A}$ is the adjacency matrix of the graph; $\tilde{D}$ is the degree matrix of $\tilde{A}$; $D_{ij} = \sum_{j=1}^i A_{ij}$, $H^{(l+1)}$ represents the characteristics of the $l+1$ layer; W^l is the weight matrix of the l -th layer in GCN, and b^l is the bias term of the l -th layer in GCN.

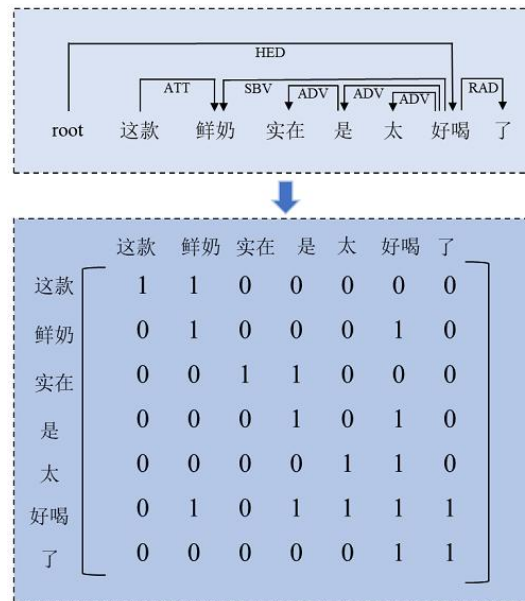


Figure 3. The Dependency Parsing Tree and Its Corresponding Adjacent Matrix.

3.4. Sentiment classification layer

After the averaging pooling of the word sequence vector $H^{(l+1)}$ output by the GCN coding layer, H^* is obtained as the input of the fully connected layer. Softmax was used to predict the emotional outcome of text statements, and the expression was calculated as follows:

$$y = \text{argmax}(\text{Softmax}(WH^* + b)) \quad (10)$$

Where, y is the output text classification prediction result, W is the weight matrix, b is the deviation value. In this paper, cross entropy loss function is used for model training, and the calculated expression is as follows:

$$\text{Loss} = - \sum_{i=1}^{|C|} y_i \log \hat{y}_i + \lambda \|\theta\|^2 \quad (11)$$

Where, C is the number of sentiment types, y_i represents the true sentiment label, $\hat{y}_i$ represents the predicted sentiment label, λ is L2 regularization coefficient, θ is the regularization parameter.

4. Experiment and Analysis

4.1. Experimental Environment

In this paper, Python is used as the model programming language, Keras as the model framework, TensorFlow provides model architecture support, and the environment configuration is shown in Table 1.

Table 1. Experimental environment configuration.

Name	configuration Details
Operating system	Windows 11
Processor	11th Gen Intel(R) Core(TM) i5-1155G7 @ 2.50GHz(8CPUs)
Programming language	Python3.7
Deep learning frameworks	Keras2.4.3 和 TensorFlow2.3.1
Programming tool	JetBrains PyCharm

4.2. Data Set

The GI-DGC model was evaluated on three datasets. Datasets 1 and 2 were collected via web crawling from JD.com, comprising approximately 10,000 Chinese reviews of fresh milk and popular apples posted between January and December 2022. Sentiment labels were assigned based on star ratings: reviews with ≥ 4 stars were labeled positive, and those with ≤ 2 stars negative; intermediate ratings were excluded. The texts underwent cleaning to ensure quality. Dataset 3 is the public “online_shopping_10_cats” dataset, containing over 5,000 “pants” reviews from JD.com (see Table 2 for samples). As typical of e-commerce platforms, most user-generated reviews are short, with lengths predominantly between 10 and 50 words, as shown in Figure 4, confirming their classification as short texts.

Table 2. Part of the sample data.

Product categories	Sentiment Polarity	Text Contents
Fresh Milk	Positive	Good milk is a regular purchase, full of milk flavor, continue to order.
	Negative	It's very different from what you buy at the supermarket. It's light as water. Not recommended.
Apple	Positive	I got the apples. Good. Come back after you eat.
	Negative	The apple is mediocre, a little sour, not sweet enough.
Pants	Positive	They have no color difference, they are the right size, and they have elasticity. They are really good!
	Negative	Pants disorderly hair, return and exchange feel too troublesome! The process is complicated and it is a very unsuccessful shopping.

Length distribution of each category data set

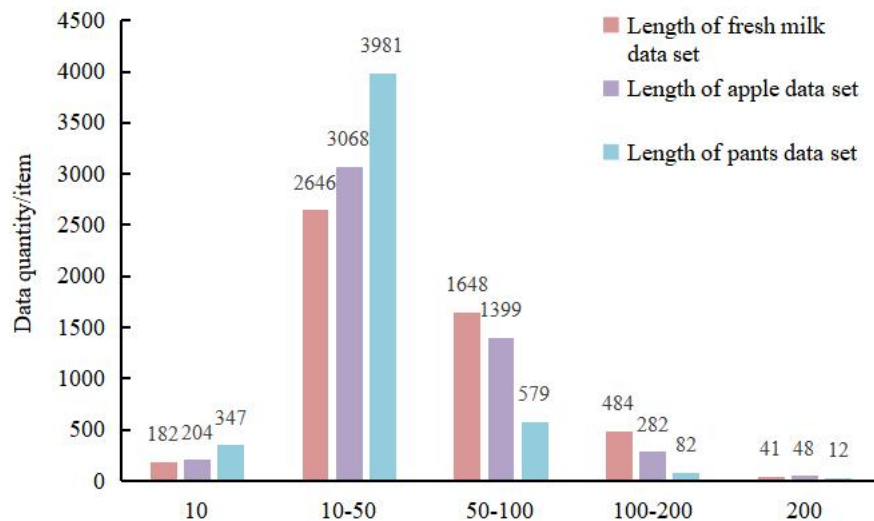


Figure 4. Length distribution of each category data set.

4.3. Evaluation Index

In order to verify the effectiveness of GI-DGC model in text sentiment classification, several evaluation indexes common to classification problems are adopted in the experiment, which are accuracy rate (Accuracy), macro average accuracy rate (Macro_{precision}), macro average recall rate (Macro_{recall}) and macro

average F1 value (Macro_{f1}). Hereinafter referred to as Acc , P_{macro} , R_{macro} and $F1_{\text{macro}}$. The calculation formula is described as follows:

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (12)$$

$$P_{\text{macro}} = \frac{1}{n} \sum_{i=1}^n \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i} \quad (13)$$

$$R_{\text{macro}} = \frac{1}{n} \sum_{i=1}^n \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i} \quad (14)$$

$$F1_{\text{Macro}} = \frac{2 \times P_{\text{Macro}} \times R_{\text{Macro}}}{P_{\text{Macro}} + R_{\text{Macro}}} \quad (15)$$

4.4. Experimental parameter Settings

Experimental parameter settings critically affect result reliability. This paper uses 300-dimensional Word2Vec embeddings, cross-entropy loss, Dropout for regularization, Adam optimizer, and a fixed text length of 50 (truncated or zero-padded as needed). LTP is used to extract dependency syntax trees for deep textual analysis. Detailed model parameters are provided in Table 3.

Table 3. Model parameter setting.

Model parameter	Set value
Batch_size	32
Max_len	50
Dropout	0.5
emb_size	300
BiGRU_units	150
GCN_layers	2
epochs	10

4.5. Experimental results and analysis

4.5.1 Comparison experiment with benchmark model

In order to verify the performance of GI-DGC model, the following 7 baseline models were selected for comparison.

(1) FastText[24]: Based on the embedding of subwords, it uses subwords to represent words and maps each subword to a vector.

(2) TextCNN[11]: Convolution operation is used to capture n-gram features of different lengths, and then these features are summarized into a vector of fixed length. Finally, full connection layer is used for classification.

(3) DPCNN[25]: composed of convolution module and pooling module, the main idea is to use a "pyramid pooling" technology to capture semantic information of different levels.

(4) LSTM[20]: a cyclic neural network (RNN) model commonly used for processing sequence data. By introducing memory unit structure, information in long text can be better obtained.

(5) GRU[21]: Introducing two gated units can better process long sequence data and avoid the problem of disappearing gradient.

(6) BiGRU[26]: Compared with GRU, BiGRU can consider both forward and backward information when processing sequence data, so as to better capture the context information in the sequence.

(7) TextRCNN[27]: a model combining cyclic neural network (RNN) and convolutional neural network (CNN).

In this section, experimental tests were conducted on the above model and GI-DGC model based on three public data sets, and the results were shown in Table 4.

Table 4. Comparison results of different baseline models.

Data set	Model	$A / \%$	$P_{Macro} / \%$	$R_{Macro} / \%$	$F1_{Macro} / \%$
Dataset 3	FastText	94.01	94.01	94.01	94.12
	TextCNN	94.81	94.86	94.81	94.81
	DPCNN	91.62	92.23	91.60	91.58
	LSTM	93.21	93.22	93.21	93.21
	GRU	93.61	93.90	93.60	93.60
	BiGRU	94.41	94.43	94.41	94.41
	TextRCNN	94.81	94.85	94.81	94.81
	GIE-DGC	95.60	95.60	95.62	95.60

As shown in Table 4, the proposed GI-DGC model outperforms all seven baseline models. On Dataset 3, it achieves 95.60% accuracy, 95.60% macro-averaged precision, 95.62% macro F1, and 95.60% macro recall—improving by 0.79%, 0.74%, 0.81%, and 0.79%, respectively, over the best baseline. BiGRU, which processes text bidirectionally to capture richer global context, contributes a 0.8% accuracy gain over unidirectional GRU. Furthermore, GI-DGC surpasses FastText, TextCNN, DPCNN, LSTM, GRU, BiGRU, and TextRCNN by 1.59%, 0.79%, 3.98%, 2.39%, 1.99%, 1.19%, and 0.79% in accuracy, respectively, demonstrating its superior performance in sentiment classification.

4.5.2. Parameter sensitivity experiment of GIE-DGC

During the experiment, it is found that the emotion classification results of the Gi-DGC model are easily affected by the three superparameters of BiGRU's hidden layer unit number, GCN layer number and Dropout ratio. Dataset 1 and 2 are selected for experimental verification.

(1) Analysis of the hidden layer unit number of BiGRU

The number of hidden units in BiGRU significantly affects model performance. Experiments with 100~200 units (Figure 5) show that accuracy improves as the count increases from 100 to 150, peaks at 150, and then declines, indicating that 150 hidden units yield optimal performance.

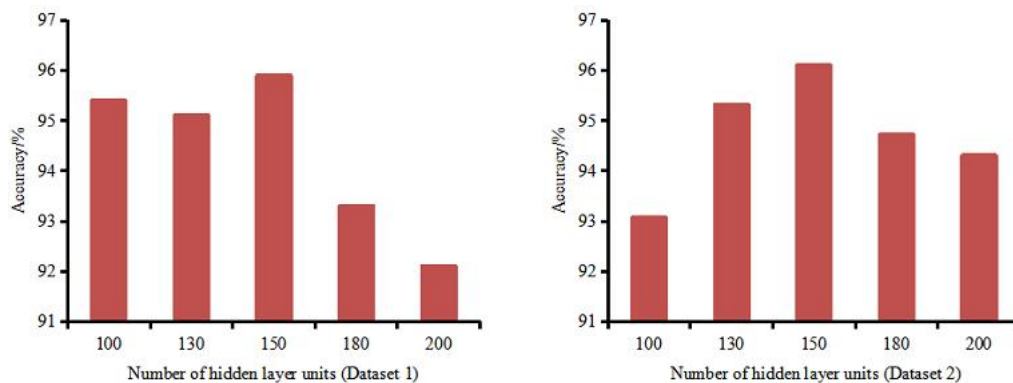


Figure 5. Comparison results of different Hidden Units.

(2) Analysis of the number of GCN layers

In the GCN model, the number of layers significantly impacts performance. Experiments with 2 to 6 layers (Figure 6) show that accuracy peaks at 2 layers and declines thereafter. As layer count increases, model parameters grow and training becomes more difficult; at 6 layers, excessive parameters hinder effective training, causing a sharp drop in accuracy. Thus, this paper sets the GCN depth to 2 for optimal performance.

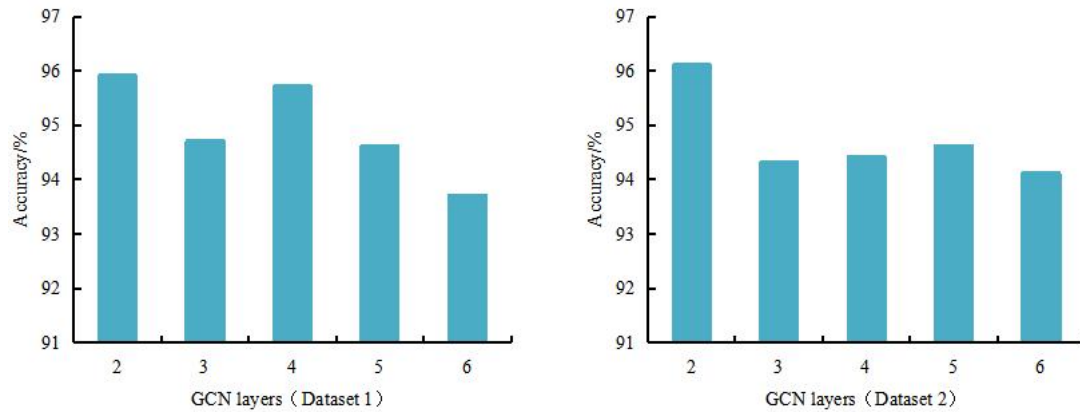


Figure 6. Comparison results of different GCN Layers.

(3) Dropout ratio analysis

A Dropout layer was added to the GI-DGC model to mitigate overfitting and enhance generalization. Experiments with Dropout rates of 0.1–0.5 (Figure 7) showed optimal performance at 0.5, which was thus selected for the final model.

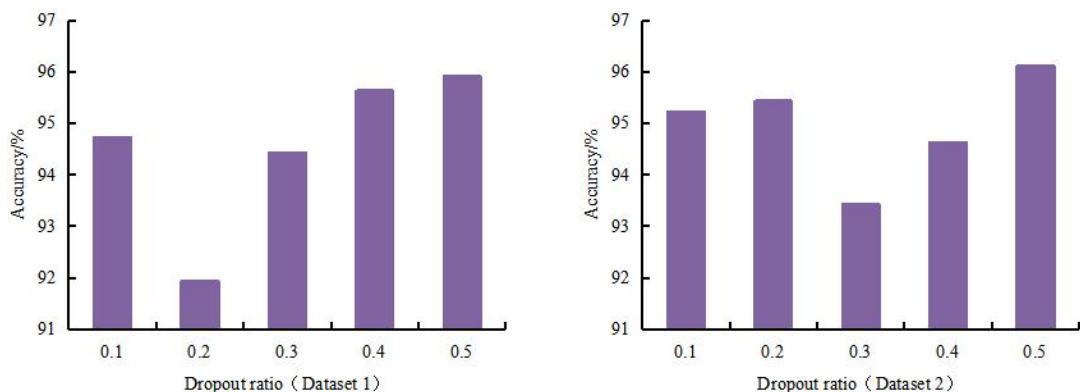


Figure 7. Comparison results of different dropout ratio.

4.5.3. Experimental analysis of the influence of different text lengths on model performance

In the GI-DGC model, both the global information enhancement layer and the dependency graph convolution layer effectively capture long-range dependencies, contextual information, and affective elements (e.g., sentiment words, expressions, and patterns) through linguistic analysis. To evaluate this capability, experiments were conducted on Dataset 1 and 2 with text lengths of 50, 70, 100, 180, and 200. The results are shown in Figure 8.

As shown in Figure 8, the model performs above average at text lengths of 50, 150, 180, and 200, with performance gradually improving beyond 150 characters. This indicates that GI-DGC effectively leverages dependency syntactic information to

capture semantic relationships in longer texts and integrates global contextual features, enabling robust sentiment analysis even on long sentences.

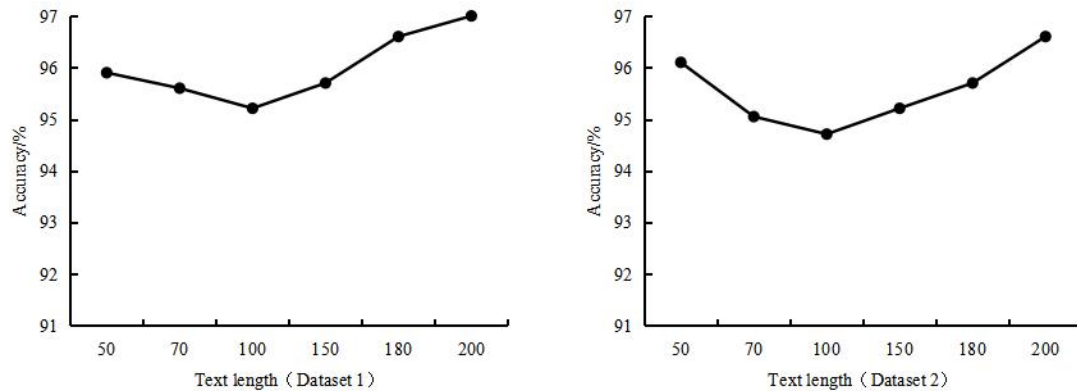


Figure 8. Comparison results of different text length.

5. Conclusions

To address the challenges of Chinese short text sentiment classification, this paper proposes GI-DGC, a model integrating global information enhancement and dependency graph convolution. Its core innovation lies in feature extraction via two components: (1) a BiGRU-based global information enhancement layer that captures contextual semantics to enrich text representation and mitigate semantic sparsity in short texts; and (2) a dependency graph convolution layer that encodes syntactic dependencies—derived from dependency parsing—into a graph structure for grammatical analysis. These features are fused via GCN to yield more comprehensive representations. Experiments on three short-text datasets demonstrate that GI-DGC effectively improves classification by jointly modeling global context and syntactic structure. However, limitations remain: (1) the model only handles binary (positive/negative) sentiment, ignoring fine-grained or multimodal classification; and (2) it uses static Word2Vec embeddings, which cannot resolve polysemy or adapt dynamically to tasks. Future work will explore dynamic word representations, aspect-level sentiment analysis, and multimodal extensions to enhance accuracy and generalization.

Conflicts of Interest

The author declare that there is no conflict of interest regarding the publication of this article.

Funding

This article is supported by the 2025 Graduate Teaching Project of University of Shanghai for Science and Technology (Undergraduate-Graduate Integrated Curriculum on “Object Engineering”).

References

- [1] Chen, L.; Zhang, Y.; Du, X. A novel product review sentiment analysis model based on deep learning. *Journal of Ambient Intelligence and Humanized Computing*, 2021, 12(3), 3053-3067.

- [2] Pang, B.; Lee, L. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2008, 2(1-2), 1-135.
- [3] Kim, Y. Convolutional neural networks for sentence classification. 2014. arXiv preprint arXiv:1408.5882.
- [4] Mikolov, T.; Karafiát, M.; Burget, L., et al. Recurrent neural network based language model. *In Proceedings of the Annual Conference of the International Speech Communication Association*, 2010; pp. 1045-1048.
- [5] Kiperwasser, E.; Goldberg, Y. Simple and accurate dependency parsing using bidirectional LSTM feature representations. *Transactions of the Association for Computational Linguistics*, 2016, 4, 313-327.
- [6] Kipf, T.N.; Welling, M. Semi-Supervised Classification with Graph Convolutional Networks. *International Conference on Learning Representations (ICLR)*, 2017; pp. 1-14.
- [7] Li, X.; Wang, X.; Liu, Z. et al. Integration of Global and Local Information for Text Classification. *Neural Computing and Applications*, 2022, 35(3), 2471-2486.
- [8] Zhang, H.; Lee, W.S. Naive Bayes classification algorithm for text classification. *Journal of Electronic Imaging*, 2006, 15(4), 043002.
- [9] Li, Z.; Zhang, Y.; Xu, X.; Li, Z. A comparative study on document classification using SVM and KNN. *In Proceedings of the 2015 7th International Conference on Information Technology in Medicine and Education*. 2015; pp. 66-70.
- [10] Zhang, W.; Alzahrani, I.A.; Lee, Y.M. Fake news detection in large-scale social network with generalized Bayesian classification. *Mobile Networks and Applications*, 2024; pp. 1-17.
- [11] Kielbasa, M.E.; Zima, K. Automated classification of exchange information requirements for construction projects using Word2Vec and SVM. *Infrastructures*, 2024, 9(11), 194.
- [12] Alfarhood, S. Soutcom: Real-time sentiment analysis of Arabic text for football fan satisfaction using a bidirectional LSTM. *Expert Systems*, 2024, 42(2), e13641.
- [13] Hua, J.; Sun, D.; Hu, Y. et al. Heterogeneous Graph-Convolution-Network-Based Short-Text Classification. *Applied Sciences*, 2024, 14(6).
- [14] Zhang, B.; Zhang, Y.; Wang, R. et al. Syntax-Aware Opinion Role Labeling with Dependency Graph Convolutional Networks. *ACL*, 2020; pp. 3249-3258.
- [15] Ren, Q.; Wang, Q.; Lu, Y. et al. MTSA-CG: Mongolian Text Sentiment Analysis Based on ConvBERT and Graph Attention Network. *Electronics*, 2025, 14(23), 4581.
- [16] Wang, J.; Wang, H.H.; Wang, L. Dependency Parsing of Financial News to Improve Sentiment Analysis for Predicting Market Prices. *In International Conference on Technologies and Applications of Artificial Intelligence*. IEEE, 2020; pp. 1-7.
- [17] Li, Y.; Zhang, X.; Chen, J. et al. A novel adaptive marker segmentation graph convolutional network for aspect-level sentiment analysis. *In Proceedings of the 2021 Conference of the Association for Computational Linguistics (ACL)*, 2021; pp. 1-12.

- [18]Sun, K.; Zhang, R.C.; Mensah, S. et al. Aspect-Level Sentiment Analysis Via Convolution over Dependency Tree. *In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019; pp. 1-11.
- [19]Yin, Z.; Su, J.; Hafeez, M.J. et al. DPG-LSTM: An Enhanced LSTM Framework for Sentiment Analysis in Social Media Text Based on Dependency Parsing and GCN. *Applied Sciences*, 2022, 13(1), 354.
- [20]Zhu, X.; Sobihani, P.; Guo, H. Long short-term memory over recursive structures. *Proceedings of the 32nd International Conference on Machine Learning*, Lille, France, 2015; pp. 1604-1612.
- [21]Dey, R.; Salrmt, F.M. Gate-variants of gated recurrent unit (GRU) neural networks. *Proceedings of the 60th IEEE international midwest symposium on circuits and systems*. Piscataway: IEEE, 2017; pp. 1597-1600.
- [22]Mikolov, T.; Sutskever, I.; Chen, K. et al. Distributed representations of words and phrases and their com-positionality. *Advances in Neural Information Processing Systems*, 2013; pp. 3111-3119.
- [23]Che, W.X.; Li, Z.H.; Liu, T. LTP: A Chinese Language Technology Platform. *Proceedings of the 23rd International Conference on Computational Linguistics: Demonstrations*. 2010; pp. 13-16.
- [24]Joulin, A.; Grave, E.; Bojanowski, P. et al. Bag of tricks for efficient text classification. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, Valencia, Spain, 2017; pp. 427-431.
- [25]Johnson, R.; Zhang, T. Deep pyramid convolutional neural networks for text categorization// *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Vancouver, Canada. Stroudsburg: Association for Computational Linguistics, 2017; pp. 562-570.
- [26]Cho, K.; Bart, V.M.; Caglar, G. et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, 2014; pp. 1724-1734.
- [27]Lai, S.W.; Xu, L.H.; Liu, K. et al. Recurrent convolutional neural networks for text classification. *Proceedings of the AAAI Conference on Artificial Intelligence*, Austin, Texas, USA, 2015, 29(1), 2267-2273.



© 2026 by the author(s); licensee International Technology and Science Publications (ITS), this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (<http://creativecommons.org/licenses/by/4.0/>)