BILINGUAL PUBLISHING GROUP

Journal of Computer Science Research
https://journals.bilpubgroup.com/index.php/jcsr

ARTICLE

RoBERTa-GCN: A New Method for Relation Extraction in Automobile Accessory Domain

Ruhu Ma [1], *Linsong Hu* [2*], *Ziqi Wang* [1], *Ruhao Ma* [3]

[1] *Hefei Information Technology University, Hefei 230601, China*
[2] *School of Artificial Intelligence and Big Data, Hefei University, Hefei 230601, China*
[3] *Anhuir Yuanhou Zhijia Information Technology Co., Ltd, China*

ABSTRACT

The automotive industry's rapid expansion has sparked increasing interest in the realm of automotive accessories. Navigating vast information landscapes to find accurate matches has become paramount. Leveraging cutting-edge information technologies, such as knowledge graphs and graph database-based question-answering systems, offers a crucial avenue for enhancing search efficiency. Addressing challenges posed by the domain's specialized terminology and intricate relationships, this paper introduces an innovative approach that combines a pre-trained model (RoBERTa) with graph convolutional networks (GCN). Initially, the text undergoes processing through the pre-trained model, yielding semantic feature vectors that enhance comprehension of industry-specific terminology. Subsequently, a graph convolutional network (GCN) is employed to process these semantic vectors, capturing a broader scope of neighboring vector node information. This approach not only strengthens the relationships between semantic information but also captures the intricate interconnections among entities. Ultimately, an automotive accessory query knowledge graph question-answering system is constructed using extracted entity relationship triplets. Experimental results demonstrate that the proposed RoBERTa-GCN model outperforms other baseline models, achieving an impressive F1 score of 83.93%. This research significantly enhances query capabilities and exhibits versatility in handling natural language inputs from diverse users.

Keywords: Knowledge graph; Relation extraction; Pre-trained model; Graph convolutional mode

*CORRESPONDING AUTHOR:
Linsong Hu, School of Artificial Intelligence and Big Data, Hefei University, Hefei 230601, China; Email: lins_hu@163.com

ARTICLE INFO
Received: 28 May 2024 | Revised: 31 May 2024 | Accepted: 18 June 2024 | Published Online: 30 June 2024
DOI: https://doi.org/10.30564/jcsr.v6i3.6695

CITATION
Ma, R.H., Hu, L.S., Wang, Z.Q., et al., 2024. RoBERTa-GCN: A New Method for Relation Extraction in Automobile Accessory Domain. Journal of Computer Science Research. 6(3): 1–9. DOI: https://doi.org/10.30564/jcsr.v6i3.6695

1. Introduction

Building an automobile accessory knowledge graph plays a crucial role in user queries [1]. Unlike traditional relational databases, graph databases offer numerous advantages such as easy modeling, storage, and querying of massive amounts of relationship data, as well as complex relationship queries and analysis. The emergence of websites containing various automotive accessory information, such as YiChe Index and Dataeye, provides a foundation for constructing knowledge graphs. Currently, automotive accessory businesses still rely on searching within relational databases, which not only severely impacts efficiency but also requires specific syntax formats. Therefore, extracting valuable entity relationships from a large amount of unstructured data is essential for building a knowledge graph question-answering system.

Conventional approaches to relation extraction heavily depend on rule-driven matching techniques., using predefined rules to extract structured knowledge from a large volume of unstructured or semi-structured text. There are a lot of pre-trained models proposed in relation extraction in recent years, such as (RoBERTa) [2] and graph convolutional networks (RoBERTa-GCN) [3]. For example, reference [4] proposes a semantic rule-based method for extracting urban fire knowledge, Reference [5] introduces a knowledge extraction approach rooted in semantic and syntactic features within the realm of urban traffic emergencies. While rule-based methods for knowledge extraction offer benefits like interpretability and high controllability, they are also plagued by issues such as low extraction accuracy, constrained scalability, and the intricate nature of defining rules. requiring significant human effort.

As deep learning models have gained prominence, numerous models have been employed in the realm of relation extraction, serving as replacements for traditional rule-based matching techniques. For example, models based on Recurrent Neural Networks (RNN) [6] convert text into vectors to better extract and capture internal entity relationships, thereby alleviating the limitations of rule-based matching methods. A study proposes an attention-based Bidirectional Long Short-Term Memory (BILSTM) model [7], which utilizes bidirectional extraction, incorporating both forward and backward information. The BILSTM model is further extended to a Bidirectional Gated Recurrent Unit (BIGRU) [8]. However, RNN-based models have limitations, especially when dealing with long texts and temporal dependencies. Convolutional Neural Networks (CNN) [9] have strong spatial capturing capabilities, Researchers have proposed various relation extraction methods leveraging CNN. For instance, a study employs the TextCNN model to build a knowledge graph in the power grid domain. Given that knowledge graphs inherently possess a graph-structured format, scholars have introduced the GCN [10] model for relation extraction tasks. Both GCN and CNN models excel at extracting features, but GCN particularly shines in extracting graph-based data objects and demonstrates superior performance in capturing spatial relationships within graph data. Additionally, another study leverages the Graph Convolutional Network (GCN) model to extract crucial knowledge within the oil and gas pipeline emergency domain.

With the advent of pre-trained models such as Bidirectional Encoder Representations from Transformers (BERT) [11] and Generative Pre-trained Transformers (GPT) [12], these models have addressed the limitations and drawbacks of the aforementioned deep learning models in the field of text processing. The current direction of improvement is to combine pre-trained models with deep learning models. For example, the BERT-CNN [13] method combines feature encoding and fusion for joint entity relation extraction. Although the aforementioned research has achieved certain results, it overlooks the unique structure of knowledge graphs, such as the connections between adjacent nodes.

To address the above issues, this paper proposes a method that combines pre-trained models with graph convolutional networks (RoBERTa-GCN) to tackle the challenges faced in the automotive accessory domain, such as the abundance of technical terminology and complex relationships. The main contributions of this paper are as follows:

a. The use of pre-trained models with random masking techniques for professional terminology allows The model has been enhanced to gain a deeper comprehension of the utilization of these specialized terminologies within the automotive accessory domain.

This augmentation in comprehension enhances the model's capacity to extract data features and recognize them with greater efficacy.

b. The GCN model is utilized, incorporating multiple feature adjacency matrices to better capture information between adjacent nodes and enhance the ability to handle complex relationships.

c. A comparison is made between the combination of mainstream pre-trained models and the GCN model in relation extraction. The performance is also compared with traditional deep learning baseline models to construct a knowledge graph in the automotive accessory domain.

2. Related Work

2.1 Pre-trained Models

Pre-trained models are primarily trained on massive text corpora to obtain semantic feature information and are fine-tuned for different downstream tasks. Early models, such as improvements made on word2vec [14], include the BERT model. This model utilizes global information for pre-training and predicts one word based on the semantic context. This improves the model's ability to recognize and predict, while employing self-attention mechanisms for sequence modeling to better capture relationships between different parts of the text. For relation extraction tasks, pre-training is initially performed on a large amount of text data in a semi-supervised manner to establish connections between words. It is then fine-tuned for specific domain extraction to obtain richer feature information. In this paper, As shown in **Table 1**, five models (BERT, RoBERTa, XLNET [15], SpanBERT [16], and ENRIE) are compared.

2.2 RoBERTa

The development of pre-trained models has played a significant role in the field of text processing by reducing the need for extensive human effort and improving accuracy, particularly with the emergence of the BERT model. BERT's architecture is based on Transformers, which have achieved excellent results in various specific tasks. However, since different tasks have their own specific requirements, BERT still needs to be modified accordingly to achieve optimal performance. During the training process of BERT, a certain proportion of the text is masked. However, in the processing phase, different batches of text data may have the same masks. To address this issue, RoBERTa employs dynamic masking instead of static masking, as static masking can lead to unique masks during training. As shown in **Figure 1**.

Table 1. Pre-trained models.

Model	Language modeling	Feature extractor	peculiarity
BERT	MLM	Bidirectional Transformer	Ability to obtain context-sensitive bidirectional feature representations
XLNET	PLM	Bidirectional Transformer-XL	Introduce permutation language models
BoBERTa	MLM	Bidirectional Transformer	Dynamic masking policy
SpanBERT	MLM+SBO	Bidirectional Transformer	A random mask strategy is employed
ENRIE [17]	MLM+DEA	Bidirectional Transformer	A new pre-training target is used

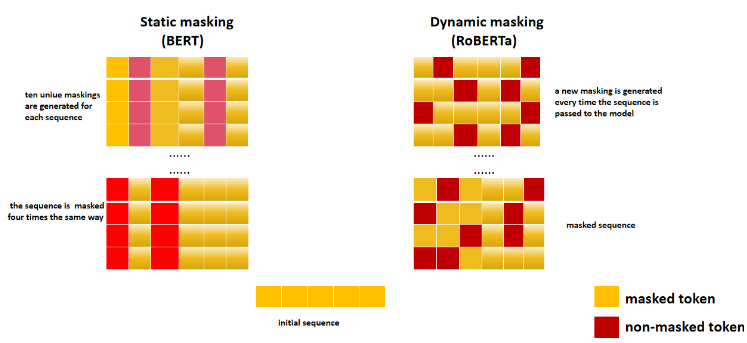

Figure 1. Static shielding and dynamic shielding.

RoBERTa adopts a dynamic masking strategy, which differs from BERT by eliminating the next sentence prediction task. Additionally, RoBERTa is trained on a larger amount of data, which helps improve the model's performance. The input layer of RoBERTa consists of three components: word embeddings, sentence embeddings, and position embeddings. Word embeddings convert textual data into vector representations using a model's embedding table. Sentence embeddings contain information from the entire corpus, while position embeddings provide information about the position of each word, enabling differentiation of semantic information across different positions in the text. The vector processing in the BERT model is illustrated in **Figure 2**.

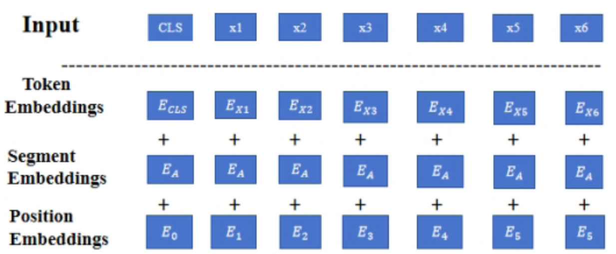

Figure 2. Embedding layer of the RoBERTa model.

As shown in **Figure 3**, the improved model utilizes a bidirectional encoder, which encodes and extracts features from both the forward and backward directions of all words. This model includes a large number of multi-head self-attention mechanisms, which enable the model to attend to different parts of the input sequence. The connections between different parts are established using feed-forward neural networks. The attention mechanism allows for the fusion of information from different sentences based on attention weights. The self-attention process is achieved by mapping the original feature vectors into three branches: Query, Key, and Value. The process involves computing the weight coefficients for Q and K, normalizing the obtained coefficients, and finally applying the weighted matrix coefficients to V to model the overall information of the text. The calculation is as follows:

$$Q_i = QW_i^Q$$
$$K_i = KW_i^K$$
$$V_i = VW_i^V \tag{1}$$

$$head_i = \text{Attention}(Q_i, K_i, V_i) \tag{2}$$

$$\text{MultiHead}(\text{Q, K, V}) = \text{Concact}(head_1, \ldots head_8) \tag{3}$$

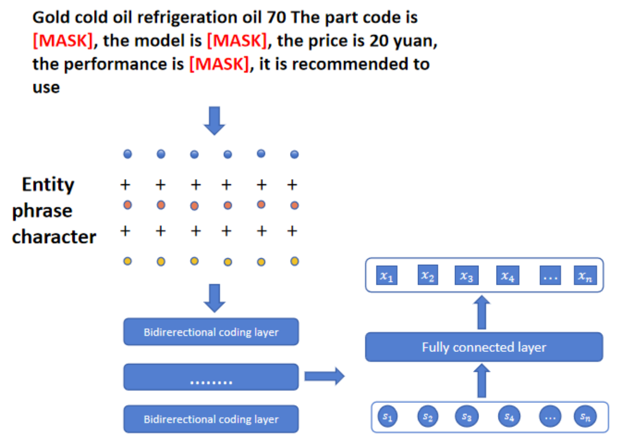

Figure 3. RoBERTa model processing flow.

2.3 GCN

The pre-trained models only extract entities without considering the relationships between them. In order to match the extracted entities with their relationships, this paper further utilizes a Graph Convolutional Network (GCN) model to extract relationship features between entities.

As shown in **Figure 4**, The GCN model starts from one node and performs convolution on its neighboring nodes to extract their features. The obtained information is then aggregated through the model, resulting in information about the surrounding region for each node. The overall process involves processing one node by transforming it into a matrix, propagating the obtained information to its neighboring nodes, aggregating the information from all nodes (combining the vectors of neighboring nodes), and finally applying the ReLU activation function to the fused vectors for non-linear transformation. The calculation is as follows:

$$h_u^{l+1} = \text{ReLU}(\textstyle\sum_{v \in D(u)})Wh_v^l + b^l \tag{4}$$

u represents the target node, $D(u)$ represents the set of nodes around you and u, indicates that node v hides features, and W, b represents the weight.The

GCN model processing process is shown in **Figure 4**.

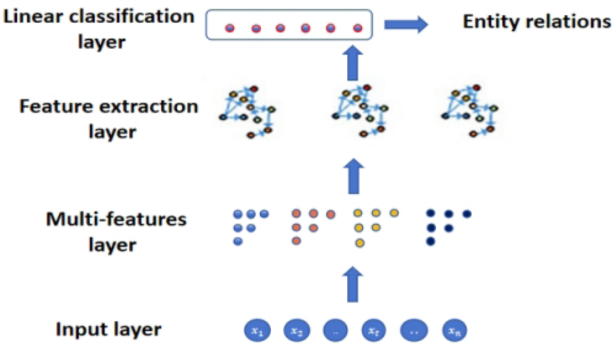

Figure 4. GCN model processing flow.

2.4 Softmax

By employing the Softmax function, the feature vectors outputted by the upper GCN model are classified to determine the connections between individual entities. This crucial step enhances the accuracy of experimental outcomes and enables a better understanding of the relationships among entities. The formula is as follows.

$$P(y = i) = \frac{e^{Z_i}}{\sum_{j=0}^{K-1} e^{Z_j}} \quad i \in \{0,12,...,K\text{-}1\} \tag{5}$$

2.5 RoBERTa-GCN

The model described in this paper consists of three components: a pre-training module, such as a convolutional module, and the Softmax function. The pre-training module primarily processes domain-specific text to enhance its ability to handle proper nouns. The graph convolutional network calculates the relationships between nodes to strengthen the complex connections among entities. The Softmax function classifies the feature vectors obtained from the upper layers to determine the relationships between entities, thereby enhancing the determination of these relationships.

As shown in **Figure 5**,The specific process involves three steps. First, text preprocessing is performed, which includes masking strategies at the character, entity, and text phrase levels, with a particular focus on masking domain-specific terms. A special token, [CLS], is added at the beginning of the text during pre-training to retain the overall information of the text. Second, entity relation extraction is conducted using the RoBERTa model with a dynamic masking strategy, aiming to obtain semantic feature information from the surrounding context of the text. Third, the GCN model processes the input feature vectors from the upper layers to serve as the input for the subsequent Softmax function. Finally, the Softmax function performs classification to determine the relationships between entities. The overall model's final result is computed using the cross-entropy loss function. The formula is as follows.

$$\zeta = -\frac{1}{N} \sum_{i=1}^{n} y^i * \log(\bar{y}^i) + (1 - y^i) * \log(1 - \bar{y}^i) \tag{6}$$

where $\bar{y}$ represents the output of the model, y^i represents the real label, and N represents the number of entity triples.RoBERTa-GCNThe flowchart is as follows.

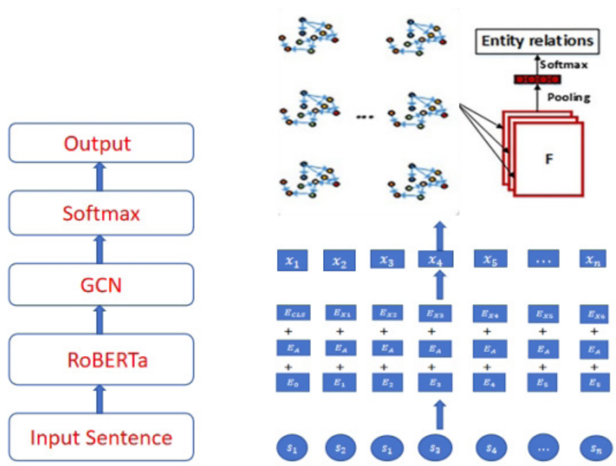

Figure 5. RoBERTa-GCN frame diagram.

3. Experiments and results

Entity relation extraction is a crucial step in constructing domain-specific knowledge graphs. In this paper, we utilized data from the database of Hefei Lianpeitong automotive accessory Company, supplemented by data from Yiche Index and China Auto Parts Network. The focus was primarily on extracting important parameters related to automotive accessory, such as part names, part codes, compatible vehicle models, prices, ratings, and warranty periods.

To validate the effectiveness of the RoBERTa-GCN model, we divided the experiments into two groups. Firstly, we conducted relation extraction experiments on the automotive accessory dataset using the RoBERTa-GCN model and compared its performance with mainstream models like BERT-BILSTM-CRF. Secondly, we performed ablation experiments to assess the impact of dynamic masking and graph convolutional models on the experiments. This chapter consists of three parts: dataset construction, evaluation metrics, baseline models, and results.

3.1 Data set

As shown in **Table 3**,The data used in this paper consists of important parameter information for automotive accessory, collected up until the year 2022. It is divided into seven categories: part code, part name, manufacturer, warranty period, user ratings, compatible vehicle models, and price. The data is further split into training and testing sets, with a ratio of 7:3. **Table 2** displays a partial sample information from the dataset, while **Table 3** provides details about the parameters in the automotive accessory dataset.

Table 2. Sample information of some datasets.

Sample text
1: Gold cold oil refrigeration oil 70 The part code is 05.01.00000-JL, the model is universal, the price is 20 yuan, the performance is OK, it is recommended to use.
2: Special No. 1 refrigeration oil 1 piece/40 parts code is 05.01.00000-MY, the price is 5 yuan, the brand is famous.
3: Special No. 1 refrigeration oil 1 piece/40 Part code is 05.01.00000-MY model is 1 piece/40 bottles Brand is famous The price is 5.
4: Car air conditioning cleaning kit Part code is 05.01.01.2008-01 Applicable model is soft box three bottles 300m The brand is Love Breath The price is 25.

Table 3. Parameters information in automobile accessory dataset.

Parameter	Training Set	Test Set
Auto accessory code	7132	3056
Name of the auto part	7155	3066
Auto parts manufacturer	5688	2438
Shelf life	5632	2414
User reviews	7630	3270
Applicable car models	7460	3197
Price	6355	2723

3.2 Evaluation indicators

The evaluation metrics used in this paper are accuracy (P), recall (R), and F1 score. Accuracy represents the ratio of correctly predicted samples to the total number of samples. Its purpose is to measure the overall prediction accuracy of the model. Recall measures the ratio of correctly predicted samples to the total number of true samples. Its purpose is to evaluate the model's ability to identify true samples correctly. The F1 score combines both accuracy and recall to measure the overall performance of the model. The larger the values of these metrics, the better the performance. The formulas for each metric are as follows:

$$P = \frac{T_p}{T_p + F_p} * 100\% \tag{7}$$

$$R = \frac{T_p}{T_p + F_N} * 100\% \tag{8}$$

$$F = \frac{2P * R}{P + R} \tag{9}$$

In the figure, T_p represents the number of correctly predicted entities, F_p represents the number of predicted entities that are not true entities, and F_N represents the number of entities that were not predicted.

3.3 Baseline model and results

To validate the effectiveness of the RoBERTa-GCN model proposed in this paper, four baseline models were used: BILSTM, BERT, RoBERTa, and RoBERTa+BILSTM [18]. Additionally, combination models were employed to compare the performance of pre-training models with traditional deep learning models, such as BILSTM+CRF [19] and BILSTM-self-ATT [20]. Furthermore, the performance of pre-training models combined with graph convolutional models was evaluated, specifically comparing BERT+BILSTM [21] with the RoBERTa-GCN model. As shown in **Table 4**, the parameter results of the training model.

Table 4. Parameters description in model training.

Parameter	Values
Batch size	8
Max sentence length	256
Learning rate	1e-5
Epochs	80
Dropout rate	0.2

As shown in **Table 5**,From the experimental results, it can be observed that the BILSTM model performs the worst. The use of bidirectional vectors in processing does not effectively handle the semantic features of domain-specific terms and can lead to gradient explosion when dealing with long texts, limiting its ability to comprehend complex textual semantics. The current mainstream experimental models utilize pre-training models as the underlying processing framework, followed by deep learning models for further extraction to obtain experimental results, such as the BERT-BILSTM-CRF model. The results show a precision (P) value of 80.41%, recall (R) value of 77.68%, and F1 score of 79.02%.

Table 5. The experimental model results are shown.

Models	P	R	F1
BILSTM [22]	69.75	68.42	69.08
BERT	71.45	70.26	70.85
RoBERTa	73.56	71.89	72.71
RoBERTa+BILSTM	77.41	76.52	76.96
BILSTM+CRF	72.71	71.43	72.06
BILSTM-self-ATT	74.38	73.86	74.12
BERT+BILSTM-CRF [23]	80.41	77.68	79.02
RoBERTa-GCN	**83.62**	**84.25**	**83.93**

Compared to traditional deep models, this model first extracts feature vectors using pre-training models and then fine-tunes them to better handle contextual information. Subsequently, bidirectional language models are employed for further processing to enhance semantic understanding. The proposed RoBERTa-GCN model in this paper outperforms the BERT+BILSTM-CRF model with a 4.91% improvement in F1 score. This improvement can be attributed to the use of graph convolutional models, which effectively integrate information from neighboring nodes and handle domain-specific terms in the automotive accessory field. This is particularly beneficial due to the presence of unique entities and complex relationships in the automotive accessory dataset.

Overall, the RoBERTa-GCN model proposed in this paper demonstrates superior performance in relation extraction compared to traditional approaches. As shown in **Figure 6**, this indicates that the RoBERTa-GCN model is effective in handling the relation extraction task in the automotive accessory dataset.

Ablation test

In order to investigate the impact of dynamic masking and graph convolutional models on the extraction performance, this paper utilizes a static masking model, BERT, as the underlying vector processing framework and designs the BERT-GCN model. To evaluate the effectiveness of the graph convolutional model, both the GCN model and the RoBERTa-BILSTM model without GCN are employed. The experimental results are presented in **Figure 7**.

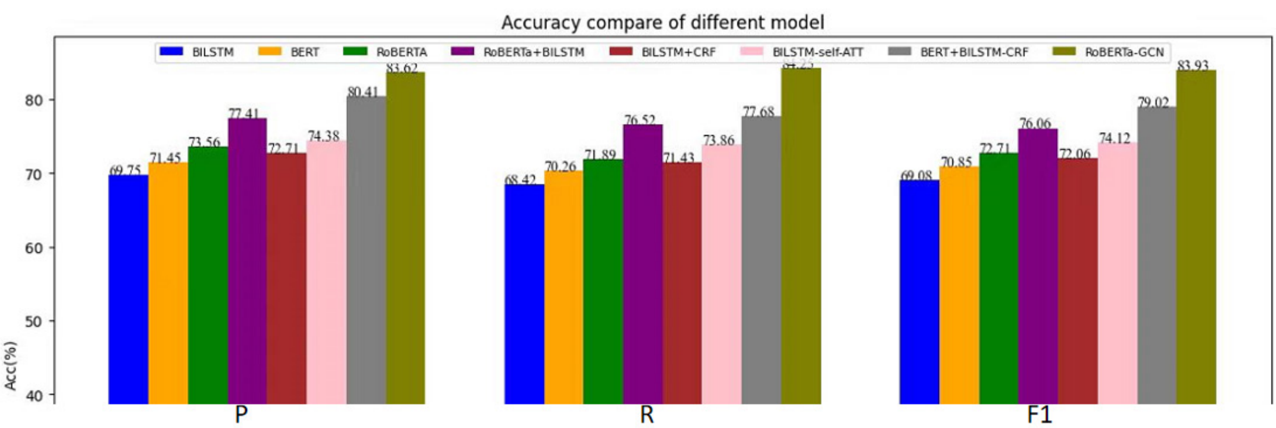

Figure 6. Experimental results of each model.

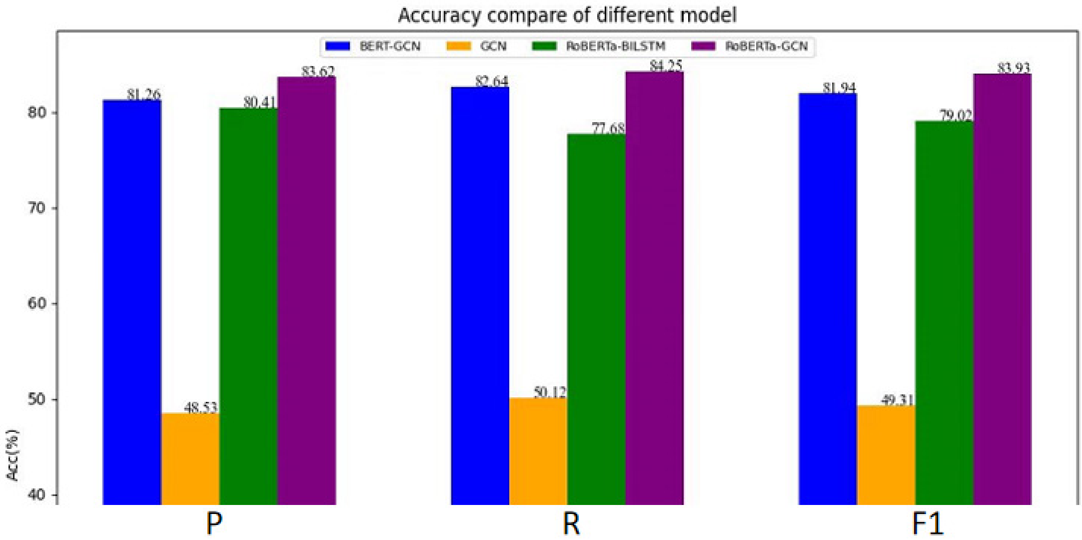

Figure 7. Ablation model experimental results.

As shown in **Table 6**, based on the experimental data, it can be observed that using a dynamic masking model for feature processing yields better results. Compared to the BERT-GCN model, the proposed model in this paper shows a 1.99% improvement in F1 score. The advantage of using the dynamic masking approach in the RoBERTa model, as opposed to the static masking approach in BERT, is that it masks the data differently at different training steps, increasing the diversity of the model's data without expanding the training set. Compared to the RoBERTa-BILSTM model, the proposed RoBERTa-GCN model demonstrates a 4.91% improvement in F1 score, highlighting the ability of the graph convolutional model to handle domain-specific terminology. The graph convolutional model calculates the weighted sum of the corresponding node vector and the neighboring node vectors using the adjacency matrix. This advantage is well-suited for domain-specific extraction tasks.

Table 6. Ablation test results.

Models	P	R	F1
BERT-GCN	81.26	82.64	81.94
GCN	48.53	50.12	49.31
RoBERTa-BILSTM	80.41	77.68	79.02
RoBERTa-GCN	**83.62**	**84.25**	**83.93**

4. Conclusions

This paper focuses on relation extraction in the automotive accessory domain and proposes the RoBERTa-GCN model, taking into account the specific characteristics of the domain's data. The experimental results demonstrate that the proposed model performs well compared to mainstream extraction models. By combining the dynamic masking model and the graph convolutional model, the RoBERTa model effectively captures contextual semantic information at the lower layers and the graph convolutional model further processes the upper layers, capturing implicit relationships between all entities and addressing the issue of overlapping relationships.

Data Availability Statement

Some or all of the data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements

The author would like to sincerely thank the reviewers, who give help to improve the quality of this paper. This research is also supported by natural science research project of Anhui Universities in 2021,Anhui,province, No. KJ2021A0990, No.2021hfuhhkc18 and No.2022hfujyzd04.

References

[1] Chen, X., Jia, S., Xiang, Y.J.E.S.w.A., 2020. A review: Knowledge reasoning over knowledge graph. 141, 112948.

[2] Liu, Y., et al., 2019. Roberta: A robustly optimized bert pretraining approach.

[3] Pei, H., et al.,2020. Geom-gcn: Geometric graph convolutional networks.

[4] Wang, X., et al., 2023. Knowledge Graph of Urban Firefighting with Rule-Based Entity Extraction. in International Conference on Engineering Applications of Neural Networks. Springer.

[5] Wang, L.H., et al., 2023. Emergency entity relationship extraction for water diversion project based on pre-trained model and multi-featured graph convolutional network. 18(10), e0292004.

[6] Sherstinsky, A.J.P.D.N.P., 2020. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. 404, 132306.

[7] Sarker, I.H.J.S.C.S., 2021. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. 2(6), 420.

[8] She, D., Jia, M.J.M., 2021. A BiGRU method for remaining useful life prediction of machinery. 167, 108277.

[9] Kattenborn, T., et al., 2021. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. 173, 24–49.

[10] Barthelmy, S., et al., 1998. The GRB coordinates network (GCN): A status report. in AIP Conference Proceedings. American Institute of Physics.

[11] Devlin, J., et al., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding.

[12] Liu, X., et al., 2023. GPT understands, too.

[13] Kaur, K., Kaur, P.J.P.C.S., 2023. BERT-CNN: Improving BERT for Requirements Classification using CNN. 218, 2604–2611.

[14] Di Gennaro, G., Buonanno, A., Palmieri, F.A.J.T.J.o.S., 2021. Considerations about learning Word2Vec. 1–16.

[15] Yang, Z., et al., 2019. Xlnet: Generalized autoregressive pretraining for language understanding. 32.

[16] Wu, C., et al., 2022. Machine Reading Comprehension Based on SpanBERT and Dynamic Convolutional Attention. in Proceedings of the 4th International Conference on Advanced Information Science and System.

[17] Zhang, Z., et al., 2019. ERNIE: Enhanced language representation with informative entities.

[18] Shi, W.,. Song, M., Wang, Y., 2022. Perturbation-enhanced-based RoBERTa combined with BiLSTM model for Text classification. in ICETIS 2022; 7th International Conference on Electronic Technology and Information Science. VDE.

[19] Li, Z., et al., 2021. Causality extraction based on self-attentive BiLSTM-CRF with transferred embeddings. 423, 207–219.

[20] Jin, J., Zhao, Y., Cui, R., 2021. A Comparative Study of Korean Feature Granularity Based on Hybrid Neural Network. in 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA). IEEE.

[21] Meng, F., et al., 2022. Creating knowledge graph of electric power equipment faults based on BERT–BiLSTM–CRF model. 17(4), 2507–2516.

[22] Aslan, M.F., et al., 2021. CNN-based transfer learning–BiLSTM network: A novel approach for COVID-19 infection detection. 98, 106912.

[23] Li, W., et al., 2022. UD_BBC: Named entity recognition in social network combined BERT-BiLSTM-CRF with active learning. 116, 105460.