2025 International Conference on Intelligent Computing
July 26-29, Ningbo, China
https://www.ic-icc.cn/2025/index.php

Causality Extraction in Chinese Public Health Events Text

* Shituo Ma[1,2,3], Lingwei Chen[1], and Ran Wang[4,5,6*]

[1] Huazhong University of Science and Technology, Wuhan, Hubei, China
[2] Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
[3] School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China
[4] School of Journalism and Information Communication, Huazhong University of Science and Technology, Hubei, Wuhan; 430074, China
[5] School of Future Technology, Huazhong University of Science and Technology, Hubei, Wuhan; 430074, China
[6] Philosophy and Social Science Laboratory of Big Data and National Communication Strategy, Ministry of Education, China

rex_wang@hust.edu.cn

Abstract. Extracting causality in public health event datasets is crucial, and traditional sentence-level extraction methods have been extensively studied. However, the performance of widely used models re- mains poor, especially for Chinese datasets. One reason is the lack of high-quality labeled Chinese datasets in this field. Additionally, implicit causality, cross-sentence causality, and multiple causalities in Chinese datasets make it difficult for models to fully extract causality. To address these issues, we constructed the first Chinese public health event dataset for causality extraction, containing 33,286 Weibo texts. We propose a model with multi-task learning to provide additional information and an attention mechanism to focus on key context for causality. The model achieved an F1 score of 0.9554 on our dataset and performed well in multiple causalities and cross-sentence causality. Our work focuses on short-text relationship extraction in the context of public health events, addressing the unique challenges of implicit causality and cross-sentence dependencies.

Keywords: Public Health Events, Causality Extraction, BERT- BiLSTM-Attention-CRF, Multi-task Learning

1 Introduction

The sudden outbreak of public health events poses significant challenges and has a substantial social impact [15]. Strong causal relationships exist between these events [11]. Traditional emergency prediction methods are inefficient for early warning and effective response to public health emergencies [14]. With the rapid development of social

* Ran Wang is the corresponding author of this article.

media, predicting public health emergencies through causality analysis based on social media has become feasible [5][29].

Causality is one of the most important relationships between events [25] and has significant applications [3]. Causality includes explicit relationships with clear trigger words (such as “because,” “so that”) and implicit relationships. Most widely used causality extraction models for public health events are limited due to the lack of high-quality labeled Chinese datasets. Previous event relation extraction models perform well on English datasets but poorly on Chinese datasets. How to construct a Chinese dataset of public health events and design related algorithms to make it more accurate in extracting events and causal relationships between events on the Chinese dataset is an important task.

In existing emergency datasets, causality is often expressed without clear cue words; causes and effects are presented in different sentences; and one cause can lead to multiple effects. These characteristics are particularly evident in public health events. Traditional sentence-level causality extraction methods often fail in these scenarios due to the difficulty of efficiently learning complex semantic representations. As a widely used language model, the attention mechanism [21] can allocate different levels of attention to the context, thereby better learning semantic logic.

The essence of causality lies in the logic between events, which requires the model to better extract the events themselves during causality extraction [18]. Multi-task learning integrates different tasks into the same model, enhancing the learning effect of the main task through the learning of sub-tasks [17]. Therefore, joint-learning of event extraction and causality extraction can improve the performance of causality extraction. Our work focuses on short-text relationship extraction in the context of public health events, addressing the unique challenges of implicit causality and cross-sentence dependencies.

Overall, the main contributions of this work include:

1. Based on large-scale social media data and annotaion, the first Chinese public health event dataset was constructed.

2. Implict causality, cross-sentence causality and multiple causalities are complexlogical relationships of public health events which are hard to extract. To solve the problem, our method proposes attention mechanism to ensure that truly crucial semantic information is fully learned when extracting complex causalties.

3. Event extraction plays a crucial role in understanding the causality between events. However, the ignorance of this aspect in past methods has hampered the accuracy of causality extraction. For this reason, our model extract causality through multi-task learning which combines causality extraction and event extraction.

2 Related Works

Causality extraction methods can be divided into two types: pattern-based methods and model-based methods. Model-based methods widely use neural networks to extract explicit and implicit causality, outperforming traditional pattern matching.

Pattern-based methods rely more on specific causal syntactic patterns to extract specific causality based on templates [13][6]. [7] designed an event causality extraction pattern rule for different parts of speech of verbs, applied in the field of patent text causality extraction. However, pattern matching is cumbersome, highly dependent on manually designed syntactic patterns, and struggles to extract text without clear causal syntax. Its application is limited to small sample data in specific fields.

Causality extraction based on deep learning requires massive training data, and the demand for high-quality Chinese data annotation poses a significant challenge to the development of these models. Early models based on sequence labeling used LSTM [12][27], CNN [22], attention [26], or CRF models [10]. The DMCNN model [2] uses dynamic pooling to apply CNN in event extraction and performs well. Li et al. proposed the SCITE model of CNN-BiLSTM-CRF architecture, which surpasses the traditional BiLSTM-CRF model in English causality extraction evaluation. Graph-based models are also efficient. Most graph-based models use GCN [4][1], considering the sequential and structural features of the text. Zhu et al. [28] proposed a two-stage GCN method, combining the de- pendency syntax diagram with the causal logic diagram. However, GCN-based models require dependency parsing, making them difficult to train. Feng et al. [3] used GAN to integrate the fused Att-BiGRU model with adversarial learning. But GAN models often struggle with long text extraction since it is hard for them to capture long sequential information.

In the process of causality extraction, there are several typical problems lie in Chinese emergency events datasets.

1. The lack of high-quality labeled Chinese datasets about public health events makes the causality analysis difficult.

2. There is often a contradiction between manually designed syntactic rules and implict causalities, making causalities could not be extracted completely.

3. The cause may be stated in one sentence, while the effect appears in the next. This cross-sentence causality complicates the task, while document-level causality extraction often has difficulty modeling implicit relationships.

4. Another significant issue is the presence of multiple causalities, where a single statement may contain “multiple causes, one effect” or “one cause, multiple effects”. This also cause low recall rate in extraction.

How to deal with these problems is also the core of whether causality can be successfully extracted.

3 Construction of Chinese Public Health Events Dataset

In the current research on causality extraction, a significant issue is the lack of Chinese public health events datasets. Moreover, for existing datasets (such as CEC-2009[25]), issues like insufficient data diversity and polluted annotations hinder the effectiveness of models. With the development of social media, collecting and screening vast amounts of datasets from social media (such as Weibo) is an effective approach.

The text data in dataset is from Weibo. The distributed crawler system crawls the text data published by 135 official Weibo accounts from January 1, 2009 to April 8, 2023, and stores it on the cloud database. The data crawling process is mainly divided into the following four steps:

1. Micro-blog account selection: The white list of 135 Chinese official media can ensure the authority of the data. The language description of the events can better reflect the logic between the events.

2. Weibo data collection: Based on the distributed crawler, text of 135 official media were crawled.

3. Filtering of Weibo data: the texts are classified and labeled what kind of emergencies is reported or it's not a report of emergencies. Then the public health data is selected by data filtering

4. Event causality labeling: Different event clauses and elements (such as time, place, trigger words, etc.) are labeled by particular pattern matching. After manual inspection, the annotation was corrected by LLMs.

Data set events are divided into two categories: public health incidents and animal epidemics. The public health events include 8 types of subordinate events, such as food safety, cholera, iatrogenic infections, etc. Animal epidemics include 6 specific types of events, such as brucellosis. The richness of events is high. The time is more concentrated after 2020, and the outbreak location covers all parts of the country. Through the observation of text data, in the obtained text, the average text length is 426.43 characters, which belongs to medium-sized text. After the subsequent causality labeling, 33,286 texts containing causality are retained.

During the annotation process, we followed the BIO tagging rules to la- bel each character as one of the five tags: "O/B-C/I-C/B-E/I-E", representing cause and effect event statements. To accurately mark causal labels, we designed regular expressions to extract explicit causal relationships and a portion of implicit causal relationships based on the characteristics of public health event text descriptions. This operation not only extracted most causal relationships but also captured event elements such as time, location, and participants. For implicit causal relationships that were difficult to describe with rules, we adopted manual annotation and large language model-assisted correction. After multiple rounds of annotation and checks, we ensured the accuracy of the data. Taking a text in the data set as an example, the pattern of labeling causal events and relationships is shown in Figure 1:

In the example shown in Figure 1, irrelevant segments are labeled as "O", while the trigger words of cause event sentences are marked as "B-C", the cause event sentences are labeled as "I-C", the trigger words of effect event sentences are marked as "B-E", and the effect event sentences are labeled as "I-E". As can be seen, the datasets contains various causal relationships, but only some of them have explicit causal cues such as "result in" or "cause". This implicit causality poses significant limitations in traditional methods, making it difficult to fully extract. Additionally, it can be noticed that there are cases of multiple causations in the datasets, for instance, "persistent fever" leads to "being quarantined for treatment" and simultaneously causes the individual to be "diagnosed as a COVID-19 patient." Facing such complex logical relationships, traditional causality extraction models often exhibit low recall rates, prone to missing such causal relationships.

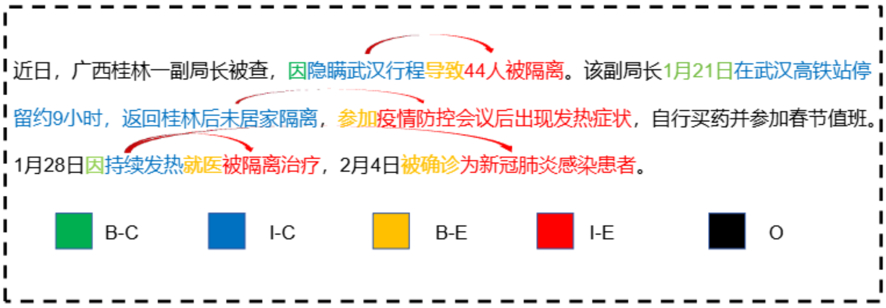

Fig. 1. An example of text annotation. O represents irrelevant text, B-C represents the beginning of the event (begin cause), I-C represents the text in the sentence of the event (in cause), B-E represents the beginning of the event (begin effect), I-E represents the text in the sentence of the event (in effect).

The proposal of this data set will be of great significance for studying the causal relationship model in public health events. The extraction of causality can further help to understand the evolution mechanism of public health events. But in the actual operation process, the following drawbacks often appear:

1. The extraction of events is not completely inadequate, and a large number of events have not been extracted because of the lack of elements;

2. The causal relationship of events is limited to one sentence, and it is difficult to detect the causal relationship of events in adjacent sentences or longer distances by existing deep learning models.

3. The causal correspondence of events is often not one-to-one correspondence, but there are phenomena such as multi-cause and one-result, one-cause and multi-result. How to reasonably match the dependency relationship of similar multi-corresponding phenomena is an important issue.

4. Tasks such as event extraction and event relation extraction can be ab-stracted as sequence labeling problems. From the perspective of model design, two tasks can be performed simultaneously through multi-task learning.

4 Causality Extraction Model via Multi-Task Learning

4.1 Model Design

The Joint-BERT-BiLSTM-Attention-CRF model is shown in Fig. 2.

In method shown in Fig2, causality extraction is modeled as a sequence labeling problem. When viewed as a sequence labeling problem, BERT-BiLSTM- CRF is a widely used baseline model for Chinese datasets. This method can effectively learn contextual semantic information in sequence labeling problems and extract implict causalities, thereby significantly improving the accuracy of implicit causality extraction. Therefore, our method builds upon this backbone and uses it as a testing benchmark.

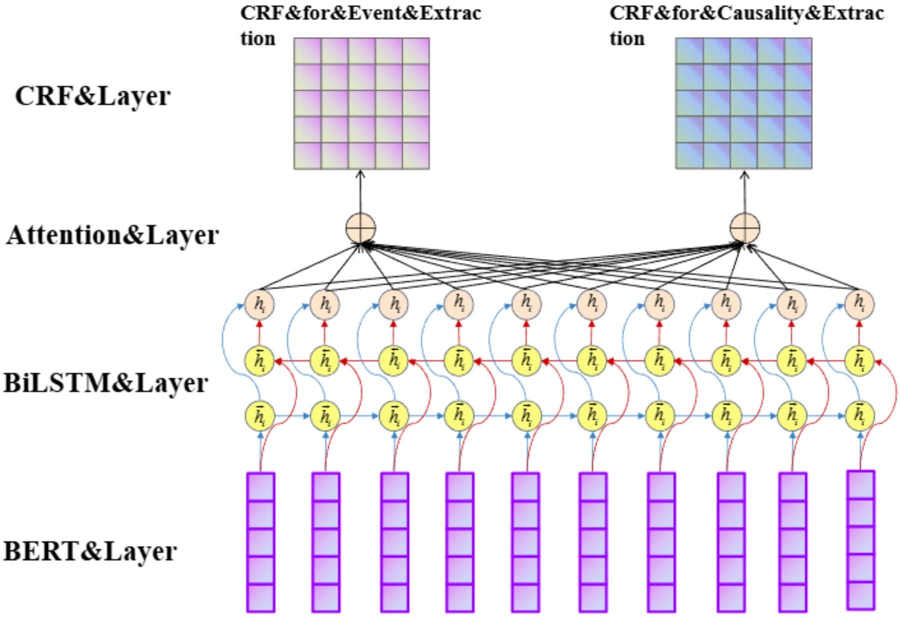

Fig. 2. Joint-BERT-BiLSTM-Attention-CRF model architecture.

However, in specific situations such as cross-sentence causality and multi- plt causalities, most models can't work well. The difficulty in extracting cross- sentence causality and multiple causality is essentially a problem of semantic understanding of complex logic. The understanding of complex logic is inherently caused by the imbalance in the intensity of learning and utilization of different semantic information, which requires the application of an additional attention mechanism. The attention mechanism can assign different weights to different information in the context, thus improving the model's comprehension of complex semantics and enabling more precise extraction of complex causal logic.

To enhance the learning of causal logic, accurate extraction and semantic learning of events themselves are essential. The results of event extraction essentially serve as additional supervisory signals for causality extraction. However, traditional models ignored the rich correlation information between tasks and fails to make full use of the data. Considering that both event extraction and causality extraction are sequence labeling tasks, and the accuracy of event extraction directly affects the extraction effect of event causality, a multi-task learning [16] framework is proposed. The main task can enrich its own semantic feature representation by updating the weight of potential semantic information by other tasks[30]. In the process of event extraction, identifying the event elements contained in different cause and effect events in the data set can enable the task sharing part to capture more accurate semantic information, thereby improving the effect of causality extraction of the main task. Therefore, two CRF layers can be used to BIO label the event elements and causality attributes of Chinese characters simultaneously through the state transition matrix.

4.2 Model Analysis

In the model, the task sharing part of the model includes BERT layer and LSTM layer, and the task-specific part uses two CRF models with self-attention layer for event extraction and causality extraction respectively.

The word embedding representation obtained by BERT integrates more con- textual semantic information, and the dynamic embedding method can also allow words to have different word embedding representations in different con- texts. After obtaining the embedded representation of the upstream text, BERT can fine-tune the network structure for different downstream tasks to obtain a network structure suitable for different tasks.

BiLSTM is widely used in sequence models such as named entity recognition and part-of-speech tagging. Because BiLSTM can learn and model temporal in- formation, it can capture the context semantics of data to form a better semantic representation. Whether the semantic information of the sequence is fully utilized has a significant effect on the improvement of causality extraction. Therefore, the LSTM layer is set as a task sharing module to acquire more efficient semantic representation.

Above the BiLSTM layer, we employed an attention mechanism to fuse the sequential information. Given that different positions in the text sequence have varying degrees of influence on the labeling of a particular token, the model opted to utilize a self-attention mechanism to capture the weights of different contextual information, thereby better describing the contextual semantic in- formation. Testing has revealed that the model with the attention mechanism indeed performs better compared to the one without it. Since the judgment of causal relationships and events are distinct tasks, the same context text can feedback with different weights in different tasks. Therefore, two separate self- attention layers were used as task-specific layers rather than shared components.

Employing CRF for causality extraction is analogous to NER. CRF is ap- plied to annotate characters or words with labels such as “O/B-C/B-E/I-C/I-E”, which is not confined to explicit causal trigger words, enabling the capture of causal event semantic features within the sentence structure. In the CRF for causality extraction, our primary task annotation involves identifying the five BIO-formatted labels in the preprocessed dataset, conducting sequential annota- tion at the character level. Conversely, in another CRF used for event extraction, the focus is on identifying elements such as time, location, participants, and trigger words within the datasets. While the two CRFs serve distinct purposes, with a primary emphasis on causality extraction, experiments have shown that the auxiliary tasks plays a crucial role in enabling the model to better understand semantic information.

5 Experiment and Evaluation

The sequence labeling model is completed using Pytorch 1.0.0 with CUDA 10.8 version under Python 3.8.6 version. The hardware environment is Ubuntu server + NVIDIA GTX2080 graphics card running code. We selected several models designed with BERT as the backbone, as well as SOTA methods for causality ex- traction tasks in the past three years, including DeepStruct [24] and DeepEX [23]. The comparative experimental results of the model on the CEC dataset and the public health event dataset are shown in Table 1:

Table 1. Comparative experimental results of causality extraction.

Dataset	CEC-2009[25]			Public health events		
	F1	P	R	F1	P	R
BiLSTM	0.6660	0.7812	0.5804	0.7598	0.7049	0.824
BILSTM-Attention[8]	0.7240	0.8684	0.6211	0.7864	0.8325	0.7452
BiLSTM-Attention-CRF	0.7284	0.7739	0.6881	0.7866	0.8706	0.7173
CNN-BiLSTM-Attention-CRF	0.6588	0.7826	0.5688	0.5637	0.6048	0.5278
DMCNN[2]	0.7383	0.8011	0.6847	0.5829	0.6063	0.5612
KLG[9]	0.7497	0.8162	0.6933	0.8268	0.8966	0.7671
XLNet[20]	0.797	0.8128	0.7820	0.8108	0.8788	0.7526
BERT-GCN[19]	0.8045	0.8269	0.7833	0.7800	0.8674	0.7087
DeepStruct-multitask[24]	0.7830	0.843	0.731	0.8998	0.902	0.8977
DeepEX-Zeroshot[23]	0.8306	0.8512	0.8110	0.8618	0.8412	0.8836
BERT-BiLSTM-CRF	0.8236	0.8216	0.8257	0.8467	0.9012	0.7984
Joint-BERT-BiLSTM-CRF	0.8495	0.8782	0.8227	0.9366	0.932	0.9413
Ours	0.8514	0.8788	0.8257	0.9554	0.9402	0.9711

From Table 1, it can be found that the Joint-BERT-BiLSTM-Attention-CRF model outperforms other methods on the F1 score, significantly exceeding the previous baseline test. This model can greatly improve the precision of the label- ing task, but it has a certain sacrifice in the recall rate compared to the method without BERT. It is worth noting that the method of introducing convolutional neural network is generally not as good as the LSTM model in Chinese causality extraction, which is different from the test results of previous academic circles on English evaluation tasks. The core problem may still be the syntactic logic difference of Chinese implicit causality, and the difficulty in forming Chinese efficient word embedding representation, resulting in the failure of features at different semantic levels.

We have selected and compared the annotation results of different models on a specific case, as shown in the Fig3.

From the Fig3, it can be observed that traditional baselines such as BERT- BiLSTM-CRF tend to make incorrect annotations for events, mislabeling unrelated text segments as causal events, and exhibit low recall rates for B-C and B-E. However, after introducing multi-task learning, the experimental results indeed show a significant improvement in F1 score, resulting in more accurate annotations for events. Nevertheless, there may still be deviations in the under- standing of causal relationships. Our method proposed an additional attention mechanism based on the previous approaches, and from the ablation experiments and annotation results, it truly achieves accurate annotations.

Fig. 3. An Example of extraction results of different models.

Since the dataset was collaboratively curated by LLMs and human annota- tors, we did not include the F1 scores of LLMs such as Llama in this phase of our evaluation. However, we were surprised to find that our model achieved a remarkably close level of accuracy in causality extraction, surpassing even some LLMs in the identification of implicit causal relationships that they could not fully comprehend. On the one hand, the human annotation process helped correct any biases in the data labeling by LLMs. On the other hand, our model was able to capture semantic information from different perspectives, leading to more precise annotations of implicit causal relationships. Another crucial issue is that the number of parameters in LLMs is extremely large, resulting in significantly slower inference speeds compared to our method. However, we also observed an interesting phenomenon: when modifying the original five label types, our method struggled to distinguish between explicit and implicit causal relation- ships described in the datasets. This indicates that while our model is proficient in identifying causal relationships, its comprehension abilities are still weaker than LLMs. The error curve and confusion metrics of Joint-BERT-BiLSTM- Attention-CRF during training are shown in Fig. 4.

It can be found from Fig. 4 that the joint model can finally converge the loss function well, and the confusion matrix is nearly a diagonal matrix. The results show that 24612 causal event pairs are extracted by model, and 4106 have implict causal syntactic cues. At the same time, we surprisingly found that this serialized multi-task learning approach made it easier for the model to identify cross-sentence causal relationships and multiple causalities. Compared to traditional structured extraction methods, this approach reduced the reliance on causal pairings, making it simpler to extract logical relationships such as “one cause, multi results.” Furthermore, compared to graph-based analysis, this method significantly reduced training costs.

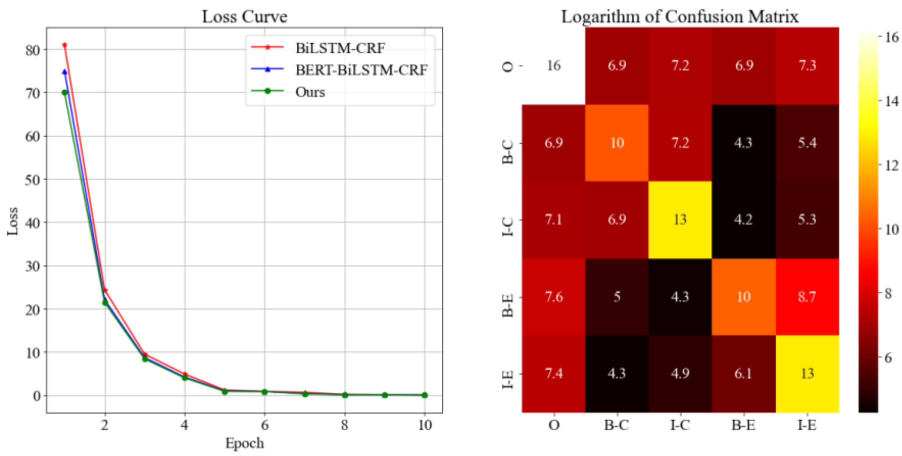

Fig. 4. The performance of Joint-BERT-BiLSTM-Attention-CRF model.

The utilization of various contextual semantic information also enabled more accurate annotation of implicit causal relationships present in different sentences. Fig.5 shows the extraction results of cross-sentence dependence and multiple causal causality.

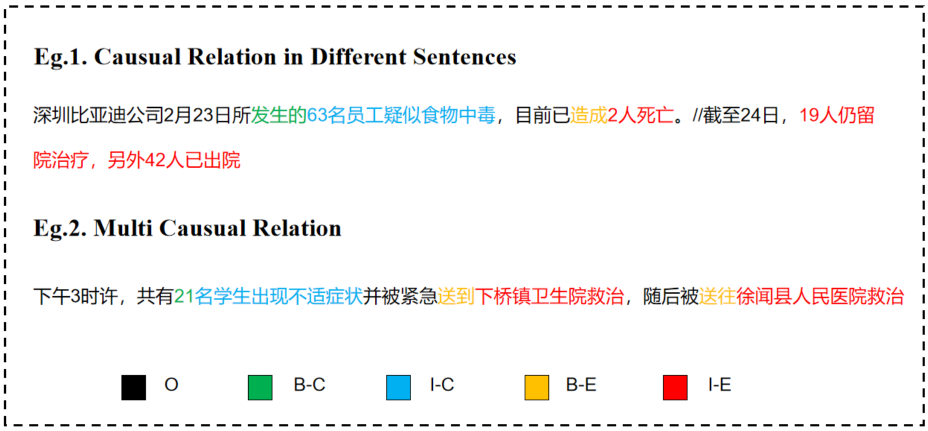

Fig. 3. Testing Cross-sentence Dependence and Multiple Causality.

From Fig.5, it can be found that since the model only focuses on which type of character-level annotation belongs, it is not constrained by the syntactic pattern within or between sentences, and there is no explicit causal syntactic pattern limitation, so it can more accurately identify the causal relationship with cross-sentence and the sentences with multiple causal correspondences.

6 Conclusion

This study addresses the issue of causality extraction in Chinese public health emergencies. We establish the first Chinese public health event dataset and an- notate causality using sequence labeling. We propose the Joint-BERT-BiLSTM- Attention-CRF model, which utilizes the attention mechanism to enhance the performance of cross-sentence and multiple causality extraction. Furthermore, we utilize multi-task learning to further improve the performance of causal extraction. Experimental results show that our method outperforms the benchmark.

Acknowledgments. This work was supported in part by the Taihu Lake Innovation Fund for Future Technology, Huazhong University of Science and Technology (HUST), under Grant 2023-B-8; in part by the Fundamental Research Funds for the Central Universities, HUST,under Grant 82400049. Ran Wang is the corresponding author (rex_wang@hust.edu.cn). The computation is completed in the HPC platform of Huazhong University of Science and Technology.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Balali, A., Asadpour, M., Campos, R., Jatowt, A.: Joint event extraction along shortest dependency paths using graph convolutional networks. Knowledge-Based Systems (2020)
2. Chen, Y., Xu, L., Liu, K., Zeng, D., Zhao, J.: Event extraction via dynamic multi-pooling convolutional neural networks pp. 167–176 (Jul 2015). https://doi.org/10.3115/v1/P15-1017, https://aclanthology.org/P15-1017
3. Feng, C., Kang, L., Shi, g., Huang, H.: Causality extraction with gan. Journal of Automation 44(5), 811–818 (2018)
4. Fu, T.J., Li, P.H., Ma, W.Y.: Graphrel: Modeling text as relational graphs for joint entity and relation extraction. In: Proceedings of the 57th annual meeting of the association for computational linguistics. pp. 1409–1418 (2019)
5. Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S., Bril- liant, L.: Detecting influenza epidemics using search engine query data. Nature 457(7232), 1012–1014 (2009)
6. Hashimoto, C., Torisawa, K., Kloetzer, J., Oh, J.H.: Generating event causality hypotheses through semantic relations. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 29 (2015)
7. Kim, H., Joung, J., Kim, K.: Semi-automatic extraction of technological causality from patents. Computers & Industrial Engineering 115, 532–542 (2018)
8. Lee, J., Seo, S., Choi, Y.S.: Semantic relation classification via bidirectional lstm networks with entity-aware attention using latent entity typing (2019)
9. Li, B., Ye, W., Zhang, J., Zhang, S.: Reviewing labels: Label graph network with top-k prediction set for relation extraction (2022)
10. Li, Z., Li, Q., Zou, X., Ren, J.: Causality extraction based on self-attentive bilstm- crf with transferred embeddings. Neurocomputing 423, 207–219 (2021)
11. Luo, C., Li, X.: Analysis of sudden group events and their evolutionary mecha- nisms. China Soft Science (6), 163–171 (2009)

12. Martínez-Cámara, E., Shwartz, V., Gurevych, I., Dagan, I.: Neural disambiguation of causal lexical markers based on context. In: IWCS 2017 12th International Conference on Computational Semantics—Short papers (2017)
13. Mirza, P., Sprugnoli, R., Tonelli, S., Speranza, M.: Annotating causality in the tempeval-3 corpus. In: Proceedings of the EACL 2014 Workshop on Computational Approaches to Causality in Language (CAtoCL). pp. 10–19 (2014)
14. Peng, Z., Huang, H., Wu, H., Xie, Q.: Big data analysis of ”five conditions” for emergency prevention and control in the early stage of covid-19. Governance Re- search 36(2), 5–20 (2020)
15. Qu, S., Yang, D.: Intelligent response to public health emergencies: Theoretical tracing and trend analysis. Revolution (3), 14–21 (2020)
16. Ruder, S.: An overview of multi-task learning in deep neural networks (2017)
17. Ruder, S.: An overview of multi-task learning in deep neural networks (2017)
18. Sun, Y., Chao, Q., Li, B.: Event causality is key to computational story under- standing (2024)
19. Tian, Y., Chen, G., Song, Y., Wan, X.: Dependency-driven relation extraction with attentive graph convolutional networks pp. 4458–4471 (Aug 2021). https://doi.org/10.18653/v1/2021.acl-long.344, https://aclanthology.org/2021.acl-long.344
20. Tian, Y., Song, Y., Xia, F.: Improving relation extraction through syntax-induced pre-training with dependency masking pp. 1875– 1886 (May 2022). https://doi.org/10.18653/v1/2022.findings-acl.147, https://aclanthology.org/2022.findings-acl.147
21. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is all you need (2023)
22. Vu, N.T., Adel, H., Gupta, P., Schütze, H.: Combining recurrent and convolutional neural networks for relation classification. arXiv preprint arXiv:1605.07333 (2016)
23. Wang, C., Liu, X., Chen, Z., Hong, H., Tang, J., Song, D.: Zero-shot information extraction as a unified text-to-triple translation (2021)
24. Wang, C., Liu, X., Chen, Z., Hong, H., Tang, J., Song, D.: Deepstruct: Pretraining of language models for structure prediction (2023)
25. Xiang, W., Wang, B.: A survey of event extraction from text. IEEE Access 7, 173111–173137 (2019)
26. Xiao, M., Liu, C.: Semantic relation classification via hierarchical recurrent neural network with attention. In: Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. pp. 1254–1263 (2016)
27. Xu, Y., Jia, R., Mou, L., Li, G., Chen, Y., Lu, Y., Jin, Z.: Improved relation classification by deep recurrent neural networks with data augmentation. arXiv preprint arXiv:1601.03651 (2016)
28. Zhu, G., Sun, Z., Zhang, S., Wei, S., Li, K.C.: Causality extraction model based on two-stage gcn. Soft Computing 26(24), 13815–13828 (2022)
29. Zhu, H.: Research on Causality of aviation safety accident based on Event Evolu- tionary Graph. Ph.D. thesis, Tianjin: Civil Aviation University of China (2019)
30. Zhu, Y.: Sequence tagging causal relation extraction based onmulti-task learning. Jilin University 01 (2022)