ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

A Hybrid LLM Based Relationship Extraction Algorithm for Unstructured Data

Sukhvinder Kaur Walia, Dr. Shraddha Masih, Dr. Ugrasen Suman

School of Computer Science & IT, Devi Ahilya Vishwa Vidalaya, Indore, M.P. India

DOI: https://doi.org/10.51583/IJLTEMAS.2026.150700173

Received: 10 August 2026; Accepted: 15 August 2026; Published: 27 August 2026

ABSTRACT

Unstructured text data such as crime reports and witness statements often contain essential connections between crime entities such as suspects, weapons, locations, and crime types. However, it can be difficult to extract and analyze these linkages because they are often inserted within unstructured narratives. This paper presents a hybrid framework that combines crime-specific entity recognition with fine-tuned large language models (LLMs) to extract relationships from crime data. The pipeline initially employs a BERT-based model to identify crime-related entities, followed by a Tiny LLAMA and LoRA-optimized LLM to discern and structure relationships as subject–predicate–object triplets which helps in criminal analytics. Experiments conducted on selected crime dataset showed significant improvements in 85.8 % accuracy with F1-score, recall, and precision over baseline models, especially when handling complicated relationship types. The results show that transformer-based NER combined with targeted LLM fine-tuning offers an effective hybrid approach for extracting useful relations from unstructured crime texts, which is helpful to criminal justice analytics and law enforcement.

Keywords: crime, relationships, triplets, LLM, NER, criminal justice

INTRODUCTION

Crime reports, police records, and witness statements frequently contain crucial details regarding individuals, incidents, locations, and items linked to a crime. These details are typically presented in lengthy paragraphs, are rich in descriptive language, and lack a consistent structure [1]. While the reports are full of information, but it might take a lot of time for law enforcement agencies to make meaningful connections between various factors such as “who committed the crime”, “where it happened“,” what weapon was used”, and “what kind of crime occurred”. In order to provide quicker and more precise analysis, technologies that can automatically recognize and connect these details are required.

In recent advancements machine learning and natural language processing techniques have made it possible to extract structured information from unstructured text, which is a major step forward for crime data analysis [2, 3]. As shown in a figure 1 for any crime sentence “Ramesh used a pistol to commit the robbery” , “Ramesh”, “pistol” and “robbery” are the three extracted crime entities, but these entities are not able to link together to form a meaningful facts, that is why relationship extraction is necessary. Without this linking, the investigators not able to connect the specific evidences, which slow down the investigation process and prone to errors. Relationship extraction identifies hidden patterns, such as recurring connections between specific suspects and crime scenes, in addition to helping to grasp the direct relationships between individuals, locations, and events. This ability to identify connections turns unprocessed text into a powerful investigation tool.

Relationship Extraction tasks can be classified on the basis of document level or sentence level according to the extracted entities [4, 5]. RE tasks are based on NLP techniques where different entities create a links between events and information based on crime text. Generally RE tasks are of many types such as Entity-entity Relationship, Event-Entity relationship, Casual Relationship, Temporal Relationship, and Spatial

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

Relationship. Entity-entity relationship connects two named entities, Event-Entity Relationship link an event to related details like time, location, suspect or victim and Casual Relationship links cause and connection [6].

In this study we focus on n-ary relationship with entity-entity and event-entity relationships in a crime text. Traditional relationship models are limited to acquiring semantic knowledge from their training datasets, which renders them insufficient due to the constraints posed by the availability of labeled data. Thus, augmenting the semantic knowledge embedded in these entities is essential for enhancing the efficacy of relation extraction (RE) tasks. The main aim of this research is to produce semantic links between the extracted crime entities and the relationship patterns which helps in crime lead generation. Relation extraction (RE) problems have seen significant advancements in the past decade because of the popularity of deep learning algorithms. Pre-trained models, such as BERT [2], have been widely used in general relation extraction with impressive results. In more specialized crime domains, models such as Open AI's GPT, DistillBERT, LegalBERT, GPT FinRE [3] use various techniques to extract crime-specific relationships, but due to the limited capability for retrieving internal relationships when crime entities are ambiguous and crime stories contains highly complex sentence structures, fine tune crime specific models are extremely required[4].

Nevertheless, studies have demonstrated that employing LLM does not yield appreciable performance improvements when compared to small models, especially when it comes to the task of relationship extraction, which is comparable to a classification problem [5]. Refining them is one technique to enhance LLM results for the RE tasks on particular domains such as crime domain at least two methods fine-tuning and few shot learning were used. Both the methods has its own pros and cons, fine-tuning LLM require an annotated dataset and significant computational resources whereas the few shot only needs a basic set of prompts. The following are the main research questions in this work attempts to answer:

RQ1: How large language models used in domain specific datasets?

RQ2: Is it possible to extract crime-specific relations by fine-tuning LLM?

RQ2: Comparing the proposed fine tune model with the existing models for QoS parameters?

To answer these questions the experiments was performed using Python libraries and transformer based functions.. The structure of the paper is divided into four sections, introduction is provided in section 1. The section 2 gives the literature survey for relationship extraction models. Section 3 and 4 describe the methodology used and experimental setup. Section 5 explains about results, conclusion and future work.

Related Work

Relationship Extraction (RE) used variety of NLP approaches from the last many years [6]. The initial methods viewed relation extraction as a two level process, starting with NER and then relation extraction [7]. Nowadays transformer based models become more powerful approach to represents complex relationships in a datasets [8]. Additionally, sequence-to-sequence (seq2seq) models have emerged as a promising technique for RE, demonstrating significant improvements in task performance [8,9].

Within these approaches, some are tailored towards extracting relationships from short sentences, typically identifying a single relationship between a pair of entities in each sentence. For the longer texts, such as paragraphs or entire documents, the model must extract all possible relationships among multiple pairs of entities [10]. Typically in the field of relation extraction the approaches are divided into three main parts machine learning based models, transformer based models and large language models.

Machine Learning RE models Rink et al. [11] employed Support Vector Machine (SVM) classifiers in conjunction with various features that encapsulate context, semantic role association, and potential pre-existing relationships of nominal for classification purposes. Zhao et al. [12] incorporated cues from

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

multiple syntactic processing levels using kernel methods and experimented with different kernels on both SVM and K-Nearest Neighbors (KNN). Kim et al. [15] introduced a system characterized by both efficiency and scalability, utilizing a linear kernel to detect drug–drug interactions. Bui et al. [17] proposed a feature-based approach to extract drug–drug interactions from text, wherein each candidate drug–drug interaction pair from a dataset was mapped into an appropriate syntactic structure to create feature vectors, which were subsequently used to train an SVM classifier. It is evident that the core of machine learning methods lies in manual feature selection, a process that is both time-consuming and labor-intensive.

Transformer based models: The introduction of the transformer architecture by [5, 13] in their seminal paper "Attention is All You Need" marked a significant paradigm shift in the domain of natural language processing (NLP). In contrast to recurrent neural networks (RNNs) and long short-term memory networks (LSTMs), which rely on sequential data processing, transformers are fundamentally based on self-attention mechanisms. This architecture enables them to effectively capture long-range dependencies, process text concurrently, and offer improved scalability. Since their inception, transformers have become foundational to most state-of-the-art models in NLP, including BERT, GPT, RoBERTa, XLNet, T5, among others [2, 14].

Domain specific Fine tune Models

Pre-trained language models that perform exceptionally well on a variety of natural language processing tasks include BERT [2, 17] , RoBERTa , and GPT . However, their effectiveness is frequently limited in specialist domains such as criminology, law, and medicine because to general-purpose training on large open-domain corpora. Researchers have created domain-specific fine-tuned models to overcome this constraint. These models use additional training on domain-relevant data to adapt general pre-trained architectures to specialized domains. For example BERT that has been trained on extensive legal datasets are Legal-BERT and InLegalBERT. For biological and clinical texts, BioBERT and ClinicalBERT, have led to significant improvements in tasks like entity detection and relation extraction in the criminology and medical literature.

The BERT model can easily identify domain specific entities but often fail to capture relational structures and hidden patterns. This led to the need of fine tuning domain specific models. Pre-trained model demonstrate how fine-tuning with domain-specific datasets improves performance and produces representations that are more semantically rich and contextually accurate. Several models have been introduced in recent years to address relationship extraction and crime-domain reasoning through domain-specific fine-tuning.

In [7] the author proposed an approach to OIE, called ClausIE which extracts relations and their arguments from text documents. Separating the detection of relevant information pieces, by exploring semantic and syntax knowledge by identifying the type of each sentence according to the grammatical function of its constituents.. In [8 a novel approach was introduces to domain specific relation extraction tasks that was used to extract semantic relationship from Chinese FinRE dataset. The proposed new framework called CRE-LLM was based on fine tuning large language model. The accuracy of the model was measured by F1-score which was calculated as 64.7% much more than other state of art models. Fine-tuning instruction-based LLMs such as LLaMA, GPT and TinyLLAMA on specific relational extraction tasks has emerged as a promising strategy, especially when combined with LoRA (Low-Rank Adaptation) for resource-efficient fine-tuning.

METHODOLOGY

The research methodology proposed in this section is structured into four primary stages: data preparation, entity extraction, triplet generation and fine-tuning of LLM. This hybrid pipeline extracts the crime relationship patterns and enables the model to generate meaningful leads which helps law enforcement agencies to solve crime puzzles. In order to facilitate crime pattern analysis, investigative reasoning, and decision-making, each step is intended to guarantee that unstructured crime reports are gradually converted into structured relationship information.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

The hybrid approach for crime relationship extraction consists of following steps:

Step 1: Crime data collection and preprocessing

Step 2: Crime entity extraction using Fine-tune BERT

Step 3: Entity Pair generation

Step 4: Filtered Entity Pair

Step 5: Triplet Generation for relationship data

Step 6: Instruction-Pair generation

Step 7: Crime Relationship Extraction using Fine-tune LLM

Step 8: Performance analysis of fine-tune LLM

The dataset used in this research was obtained from the Hugging Face repository, which provides publicly accessible datasets for machine learning and natural language processing research. The collected dataset consists of **54,633 crime narratives** containing unstructured textual descriptions of criminal incidents, witness statements, police reports, and crime-related events. Each record in the dataset contains a textual narrative describing various aspects of criminal activities, including suspects, victims, locations, weapons, crime types, dates, and other contextual information.

To ensure data quality, several preprocessing operations were performed, including removal of duplicate records, elimination of special characters, normalization of textual content, tokenization, and sentence segmentation. The resulting corpus was used for crime-specific Named Entity Recognition (NER) and Relationship Extraction (RE) tasks. The complete dataset was divided into training, validation, and testing subsets following an 80:10:10 ratio. Accordingly, 43,706 records were allocated for model training, 5,463 records for validation, and 5,464 records for testing.

Crime - Entity Extraction

The initial phase of the process involves the extraction of crime-specific entities. Crime narratives are typically composed in unstructured, free-text formats, such as police reports, witness statements, or case summaries, which pose challenges for direct computational analysis. To address this issue, we fine-tuned a domain-specific BERT model using a manually annotated crime Named Entity Recognition (NER) dataset[22]. The dataset was formatted in CoNLL style, with tokens appropriately tagged with entity labels. The CoNLL format is consists of a BIO format. In which each crime entities labeled as beginning, inside and outside entity.

The following steps are used for crime entities extraction using fine tune BERT.

Step 1: Data Preparation: Input CSV crime-report data

Step 2: Convert text into **token–label pairs** using hybrid annotation tools

Step 3: Save it in CoNLL format (B-I-O).

The below figure shows the crime entity CoNLL labeled data

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

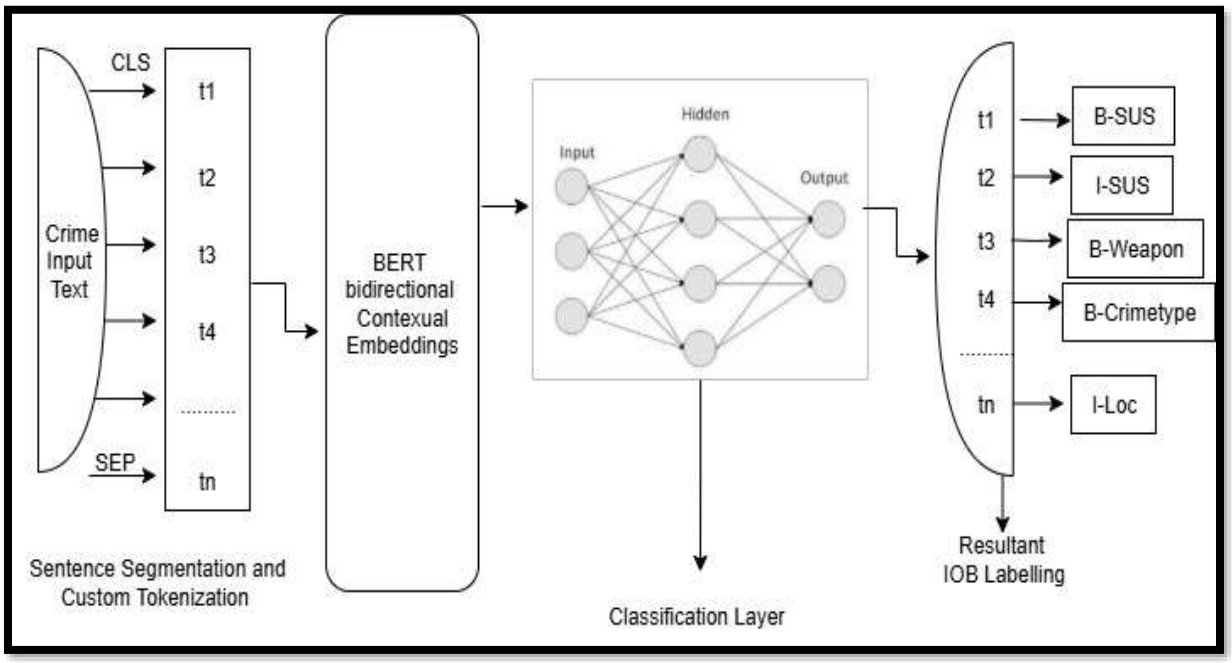

Figure I. B-I-O Format crime entities

Relationship Annotation

The second stage for relationship extraction was triplet generation. Relationship annotation is the primary step in this stage. The triplet generation is a bridge between crime extracted entities and relationship extraction in crime hybrid pipeline. In this stage each extracted entities were converted into subject-relation-object structured format suitable for fine tune LLM. Following steps are used in this stage for triplet generation.

Step 1: Entity Candidate Pairing

After identified crime entities, the model generates all possible entity candidate pairs from each sentence in a paragraph. Each extracted entity candidate pair represents relation as shown in following figure 3.

Step 2: Filtering and context extraction

Semantic meaning was not present in every pair of extracted crime entities. To eliminate unwanted redundant data, a filtering mechanism was employed. The mechanism utilizes two type of process namely dependency parsing and context extraction. Upon identifying valid pairs, the contextual sentence containing these entities is extracted. This method provides the LLM with the necessary information which used for relationship extraction.

Step 3: Triplet formation

Filtered pairs, along with their contextual sentences, are passed to the fine-tuned LLM. The LLM predicts the semantic relation between the entities, such as “Attacked”, “Used”, “Occurred_At,” or “Occurred_O.”. The outputs are then structured into **triplets** of the form (Entity1, Relation Entity2). For example, from the earlier sentence, the system produces triplets such as (A-attacked-B), (A-knife-B), (A-Jabalpur-10PM), (attacked -Jabalpur -10PM), (knife-Jabalpur-10PM) as shown in following diagram

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

Sentence#: "X attacked Y with a knife in Jabalpur at 10 PM"

Extracted Crime Entities using NER: A (Suspect), B (Victim), Knife (Weapon), Jabalpur (Location), 10PM (Time)

Extracted Candidate Pairs: (A, B), (A, Knife), (A, Jabalpur), (A, 10PM), (B, Knife),(B, Jabalpur),(B,10PM)

Generated Triplets: (A-attacked-B), (A-knife-B), (A-Jabalpur-10PM) (attacked -Jabalpur -10PM), (knife-Jabalpur-10PM)

Figure 3. Stages to generate triplets

Fine tune Relationship Algorithm

In our fine-tuning design, we integrate three essential components into our prompt: the instruction, the input sentence, and the output format. The instruction is articulated as follows:

Prompt: "What is the relationship between {E1} and {E2} in the context of the input sentence? Choose an answer from: {list_of_relations}}".

This formulation directs the model's focus towards accurately identifying the relationship between the specified entities. To further elucidate the expected response, we append the output format: ((E1, Relation, E2)), thereby ensuring that the model consistently generates relation triplets in a structured manner. This predefined prompt format is systematically applied throughout the fine-tuning dataset to guide the model's training and enhance its performance in relation extraction tasks.

Algorithm

Input: crine_text_data.csv

Output: generates triplets

begin

{

FOR each document IN crime_text_data:

CLEAN text (remove noise, normalize);

SPLIT text into sentences

FOR each sentence IN dataset:

entities = BERT_NER(sentence) ,

STORE entities in structured format

FOR each entity_set IN extracted_entities:

candidate_pairs = GENERATE all possible (Entity1, Entity2) combinations

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

```
            STORE candidate_pairs
LOAD fine_tuned_LLM_RE_model
FOR each candidate_pair IN candidate_pairs:
            context = original_sentence_containing_entities
            relation = LLM_RE(candidate_pair, context)
      IF   relation!= "No-Relation":
            STORE (Entity1, relation, Entity2)
      COMBINE entity outputs (BERT) WITH relation outputs (LLM)
GENERATE triplets
      REMOVE duplicate triplets;
NORMALIZE entity names
   training_args = TrainingArguments(
   per_device_train_batch_size=2, per_device_eval_batch_size=2,
   gradient_accumulation_steps=4,  warmup_steps=50,
   num_train_epochs=2,     learning_rate=2e-4,
   fp16=True,     logging_dir="./logs"    logging_steps=10,     evaluation_strategy="epoch",
   save_strategy="epoch",     output_dir="./lora-re-colab",
   save_total_limit=2,     push_to_hub=False,
 )
      trainer = Trainer ( model=model,          args=training_args,    train_dataset=train_dataset,
eval_dataset=test_dataset,
            tokenizer=tokenizer, )  trainer.train()
COMPARE predicted triplets WITH gold-standard test set
CALCULATE Precision, Recall, F1-score for:  - Entity extraction,  - Relationship extraction
}
 End
```

Experiments and Results

To evaluate the effectiveness of the proposed hybrid relationship extraction pipeline, a series of experiments were conducted on a crime-specific dataset containing unstructured witness statements and incident reports. The dataset was divided into training (70%), validation (15%), and testing (15%) subsets.The experiments

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

were carried out in **Google Colab** with a GPU runtime. The implementation was based on **Hugging Face Transformers** library. For dependency parsing, **spaCy (en_core_web_sm)** was used. The LLM component was implemented using **TinyLLaMA with LoRA fine-tuning** for domain-specific relation classification.

Two baseline models were used for comparison with our hybrid algorithm. The two models was rule based extraction (spaCy) for dependency parsing and transformer based (BERT) for relationship classification. The evaluation metrics used for relationship extraction tasks were precision, recall and F1 score. The precision measures the percentage of correctly extracted relations with all the extracted relations, whereas recall is the percentage of correctly extracted relations with all the actual relations. At last F1 score was calculated as the harmonic mean of precision and recall.

The experiment was performed using 10-fold cross validation applied on the test dataset. A BERT based model was fine- tuned on the annotated data saved in CoNLL file. This file stored all the crime specific extracted entities such as crime_type, suspect, weapon, location and time. To generate the triplets, each crime entity pairs as a candidate pair used dependency parsing and filter out all the irrelevant combinations and for each filtered pair context segments were passed to the LLM for relations classification on each crime sentence. Finally the results were stored as structured triplets in the form Entity1, Relation, Entity2.The following table 4.1 shows the performance comparison of our proposed model and it was found that hybrid model significantly outperformed both baseline models. From the table it was observe that spaCy rule based system achieved moderate precision but suffered from low recall, as handcrafted rules failed to capture implicit relations. The transformer model BERT improved recall but occasionally predicted false relations due to the lack of explicit filtering. Therefore the best balance between precision, recall and F1-score was shown in our hybrid model with the highest accuracy of 85.8% and 86.4% precision , 84.2% recall and 85.3% F1-score.

Table 4.1 Performance Analysis among three baseline models

Model	**Precision (%)**	**Recall (%)**	**F1-Score (%)**	**Accuracy (%)**
Rule-Based spaCy)	68.22	61.54	64.65	65.13
Transformer-Only (BERT)	78.93	75.36	77.03	76.52
BiLSTM+CRF	83.77	81.55	82.75	82.01
Proposed Hybrid Pipeline	**86.41**	**84.21**	**85.32**	**85.84**

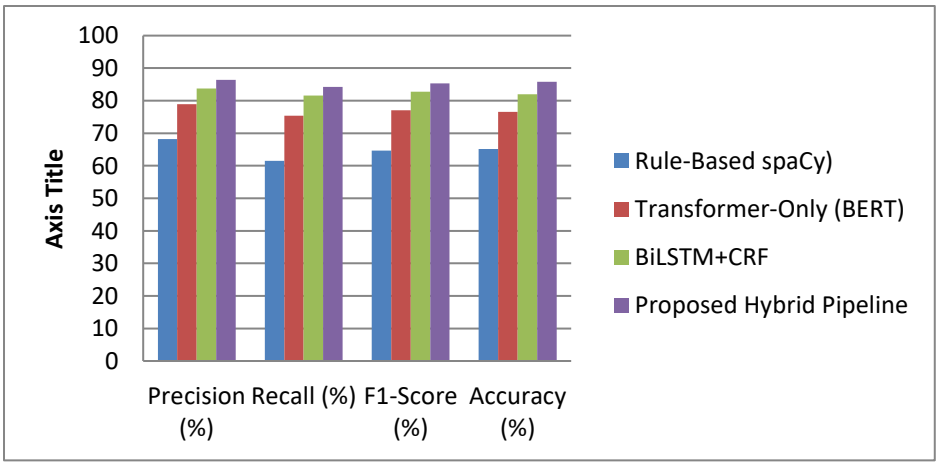

Figure 4: Performance analysis of all the four models

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

CONCLUSION AND FUTURE WORK

This research introduced a hybrid pipeline for relationship extraction with a central focus on crime-specific LLM fine-tuning. The core idea was to combine the dependency parsing rule based method with fine – tuning LLM using TinyLLAMA with LoRA . The experimental results indicate that the hybrid approach consistently surpassed traditional baseline methods. While the rule-based system offered interpretability, it encountered difficulties in identifying context-dependent relationships. The transformer-only classifier demonstrated superior recall however, it lacked the necessary filtering capability to prevent false predictions. Conversely, the domain-specific fine-tuned LLM, when integrated with dependency parsing and entity filtering produced a balanced and accurate output that closely corresponded with gold-standard annotations. The structured triplets extracted through this method can be used to construct crime patterns and support law enforcement decision making which helps in crime lead generation. Despite with the promising result, the creation of high-quality annotated datasets for crime relationship extraction is time-consuming and requires expert knowledge, often leading to limited training data and class imbalance across relation types. Looking ahead, scaling fine-tuning with larger datasets, integrating knowledge graphs, and developing explainable, real-time systems will further strengthen the impact of this research.

REFERENCES

1. Sboev, A.; Rybka, R.; Selivanov, A.; Moloshnikov, I.; Gryaznov, A.; Naumov, A.; Sboeva, S.; Rylkov, G.; Zakirova, S. Accuracy Analysis of the End-to-End Extraction of Related Named Entities from Russian Drug Review Texts by Modern Approaches Validated on English Biomedical Corpora. Mathematics 2023, 11, 354. [Google Scholar] [CrossRef]
2. Lezama-Sánchez, A.L.; Tovar Vidal, M.; Reyes-Ortiz, J.A. An Approach Based on Semantic Relationship Embeddings for Text Classification. Mathematics 2022, 10, 4161. [Google Scholar] [CrossRef]
3. Wang, H.; Zhang, F.; Xie, X.; Guo, M. DKN: Deep knowledge-aware network for news recommendation. In Proceedings of the 2018 World Wide Web Conference, Lyon, France, 23–27 April 2018; pp. 1835–1844. [Google Scholar]
4. Sun, Q.; Xu, T.; Zhang, K.; Huang, K.; Lv, L.; Li, X.; Zhang, T.; Dore-Natteh, D. Dual-Channel and Hierarchical Graph Convolutional Networks for document-level relation extraction. Expert Syst. Appl. 2022, 205, 117678. [Google Scholar] [CrossRef]
5. Le, H.Q.; Can, D.C.; Collier, N. Exploiting document graphs for inter sentence relation extraction. J. Biomed. Semant. 2022, 13, 1–15. [Google Scholar] [CrossRef] [PubMed]
6. Zhou, D.; Zhong, D.; He, Y. Biomedical relation extraction: From binary to complex. Comput. Math. Methods Med. 2014, 2014, 298473. [Google Scholar] [CrossRef] [Green Version]
7. Luo, H., Xu, C., & Chen, Y. (2024). CRE-LLM: A Domain-Specific Chinese Relation Extraction Framework with Fine-tuned Large Language Model. arXiv. https://arxiv.org/abs/2406.14745 arXiv.
8. Lai, P.T.; Lu, Z. BERT-GT: Cross-sentence n-ary relation extraction with BERT and Graph Transformer. Bioinformatics 2020, 36, 5678–5685. [Google Scholar] [CrossRef]
9. Bui, Q.C.; Sloot, P.M.; Van Mulligen, E.M.; Kors, J.A. A novel feature-based approach to extract drug–drug interactions from biomedical text. Bioinformatics 2014, 30, 3365–3371. [Google Scholar] [CrossRef] [PubMed] [Green Version]
10. Surdeanu, M.; Tibshirani, J.; Nallapati, R.; Manning, C.D. Multi-instance multi-label learning for relation extraction. In Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, Jeju Island, Republic of Korea, 12–14 July 2012; pp. 455–465. [Google Scholar]
11. Zeng, D.; Liu, K.; Lai, S.; Zhou, G.; Zhao, J. Relation classification via convolutional deep neural network. In Proceedings of the COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, Dublin, Ireland, 23–29 August 2014; pp. 2335–2344. [Google Scholar]
12. Xiao, M.; Liu, C. Semantic relation classification via hierarchical recurrent neural network with attention. In Proceedings of the COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, Osaka, Japan, 11–16 December 2016; pp. 1254–1263. [Google Scholar]

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

13. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Minneapolis, MN, USA, 2–7 June 2019; Volume 1, pp. 4171–4186. [Google Scholar]
14. Sun, Y.; Wang, S.; Li, Y.; Feng, S.; Tian, H.; Wu, H.; Wang, H. Ernie 2.0: A continual pre-training framework for language understanding. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; pp. 8968–8975. [Google Scholar]
15. Wang, Y.; Sun, Y.; Ma, Z.; Gao, L.; Xu, Y.; Wu, Y. A method of relation extraction using pre-training models. In Proceedings of the 2020 13th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, 12–13 December 2020; pp. 176–179. [Google Scholar]
16. Wei, Q.; Ji, Z.; Si, Y.; Du, J.; Wang, J.; Tiryaki, F.; Wu, S.; Tao, C.; Roberts, K.; Xu, H. Relation extraction from clinical narratives using pre-trained language models. In Proceedings of the AMIA annual symposium proceedings. American Medical Informatics Association, Washington, DC, USA, 16–20 November 2019; Volume 2019, p. 1236. [Google Scholar]
17. Luciano Del Corro and Rainer Gemulla. Clausie: clause-based open information extraction. In Proceedings of the 22nd international conference on World Wide Web, pages 355–366. ACM, 2013.
18. Han, X.; Gao, T.; Yao, Y.; Ye, D.; Liu, Z.; Sun, M. OpenNRE: An Open and Extensible Toolkit for Neural Relation Extraction. 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): System Demonstrations, Hong Kong, China, 3–9 November 2019; pp. 169–174. [Google Scholar]
19. Cho, C.; Choi, Y.S. Dependency tree positional encoding method for relation extraction. In Proceedings of the 36th Annual ACM Symposium on Applied Computing, Virtual, 22–26 March 2021; pp. 1012–1020. [Google Scholar]
20. Rink, B.; Harabagiu, S. Utd: Classifying semantic relations by combining lexical and semantic resources. In Proceedings of the 5th International Workshop on Semantic Evaluation, Los Angeles, CA, USA, 15–16 July 2010; pp. 256–259. [Google Scholar]
21. Zhao, S.; Grishman, R. Extracting relations with integrated information using kernel methods. In Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05), Ann Arbor, MI, USA, 25–30 June 2005; pp. 419–426. [Google Scholar]
22. Kim, S.; Liu, H.; Yeganova, L.; Wilbur, W.J. Extracting drug–drug interactions from literature using a rich feature-based linear kernel approach. J. Biomed. Inform. 2015, 55, 23–30. [Google Scholar] [CrossRef] [Green Version]
23. Choi, S.P. Extraction of protein–protein interactions (PPIs) from the literature by deep convolutional neural networks with various feature embeddings. J. Inf. Sci. 2018, 44, 60–73. [Google Scholar] [CrossRef] [Green Version]
24. Liu, S.; Tang, B.; Chen, Q.; Wang, X. Drug-drug interaction extraction via convolutional neural networks. Comput. Math. Methods Med. 2016, 2016, 6918381. [Google Scholar] [CrossRef] [Green Version]
25. Peng, Y.; Lu, Z. Deep learning for extracting protein-protein interactions from biomedical literature. In Proceedings of the BioNLP 2017, Vancouver, BC, Canada, 4 August 2017; pp. 29–38. [Google Scholar]
26. Zhou, H.; Ning, S.; Yang, Y.; Liu, Z.; Lang, C.; Lin, Y. Chemical-induced disease relation extraction with dependency information and prior knowledge. J. Biomed. Inform. 2018, 84, 171–178. [Google Scholar] [CrossRef] [PubMed]
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. In Proceedings of the COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, Osaka, Japan, 11–16 December 2016; pp. 1461–1470. [Google Scholar]
28. Li, F.; Zhang, M.; Fu, G.; Ji, D. A neural joint model for entity and relation extraction from biomedical text. BMC Bioinform. 2017, 18, 198. [Google Scholar] [CrossRef] [PubMed] [Green Version]
29. Wang, W.; Yang, X.; Yang, C.; Guo, X.; Zhang, X.; Wu, C. Dependency-based long short term memory network for drug-drug interaction extraction. BMC Bioinform. 2017, 18, 99–109. [Google Scholar] [CrossRef] [PubMed] [Green Version]