Comparative Analysis of Contextual Relation Extraction based on Deep Learning Models

[1]R. Priyadharshini, [2] G. Jeyakodi, [3] P. Shanthi Bala
[1]Department of Computer Science, School of Engineering and Technology, Pondicherry University, Puducherry, India
[1]priyadharsini.r02@gmail.com, [2]rjeyakdi02@gmail.com, [3]shanthibala.cs@gmail.com

ABSTRACT

Contextual Relation Extraction (CRE) is mainly used for constructing a knowledge graph with a help of ontology. It performs various tasks such as semantic search, query answering, and textual entailment. Relation extraction identifies the entities from raw texts and the relations among them. An efficient and accurate CRE system is essential for creating domain knowledge in the biomedical industry. Existing Machine Learning and Natural Language Processing (NLP) techniques are not suitable to predict complex relations from sentences that consist of more than two relations and unspecified entities efficiently. In this work, deep learning techniques have been used to identify the appropriate semantic relation based on the context from multiple sentences. Even though various machine learning models have been used for relation extraction, they provide better results only for binary relations, i.e., relations occurred exactly between the two entities in a sentence. Machine learning models are not suited for complex sentences that consist of the words that have various meanings. To address these issues, hybrid deep learning models have been used to extract the relations from complex sentence effectively. This paper explores the analysis of various deep learning models that are used for relation extraction.

KEYWORDS: Contextual Relation Extraction, Word Embeddings, BERT, Deep Learning Model.

INTRODUCTION

Contextual Relation Extraction (CRE) helps to understand the meaning of the entities and their relationship in a sentence. It can improve the performance of Natural Language Processing tasks such as information retrieval, question answering, and semantic search [1]. Named Entity Recognition aims to automatically identify and classify objects like people, products, organizations, locations, etc. The process of identifying the terms in a text and arranging in an appropriate group is a source for named entity recognition and a key component for text analysis. The analysis of common syntactic patterns is an important factor of NER. Many deep learning models solve entity recognition applications such as indexing documents, finding relationship among entities, and building an ontology [2-4]. The combination of NER and CRE can provide a rich understanding of the text by identifying both the entities and their relationships based on the context. The joint modeling of entity recognition and relation classification

attained more focus recently [5]. Additionally, these end-to-end models have generated massively to improve the results. Information Extraction (IE) begins with the creation of knowledge graphs that transforms unformatted text into formatted data. Entity extraction and Relation extraction are the two subtasks of IE.

Relation extraction is ongoing research for the recent years. Neural networks enabled technology is used to efficiently classify entities and relation. Natural Language Understanding (NLU) represents the associated relationship among the existing objects and a distinct relationship between two or more entities. Entity relationship is the basis for automatically creating a knowledge graph. Relation extraction instantly detect and categorizes the entities from the text during semantic relationship extraction. Example of binary and n-ary relation are shown in Figure.1.
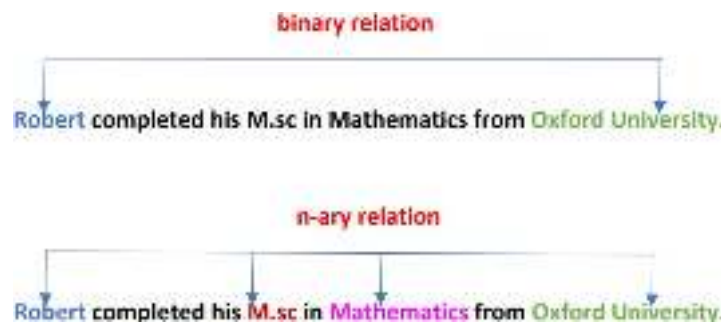

Figure 1. Example of binary and n-ary relation

Binary relation consists of two entities and one relation and n-ary relation consists of more than two entities and many relations. Binary relation extraction models may have trouble in handling larger sentence and take lot of time for processing. Some of the common issues in binary relation extraction are ambiguity, incomplete data and noise in the text. To find and understand the connections between different established categories, RE makes use of a range of technologies. Recent joint extraction models work on fixed word vector format for word embedding that are unsuitable for a word that has multiple semantic meanings. To address this problem, Bo Qiao et al. developed a dynamic fine-tuning method to overcome the issues in static word embedding using the LSTM-LSTM-Bias method proposed by Zheng et al [6].

Bidirectional Encoder Representations from Transformers (BERT), is a machine language pre- training model to represent language. BERT uses joint conditions to compare each word context in forward and backward directions. The BERT model can be improved by adding a single additional output layer for tasks such as question answering and language inference. It does not require major changes in the architecture. Devlin et al. proposed the significance of bidirectional pre-training for language representations to eliminate the requirement of multiple task-specific architectures. The BERT model is based on the fine tune representation that outperforms multiple task-specific architectures and reaches cutting-edge performance on a variety of task levels, including token and sentence levels. The pre-training and fine-tuning steps in BERT architecture help to understand the semantic meaning of the words effectively. Before

solving the joint extraction task, it pre-trains the BERT model using another corpus. BERT can be used for a wide range of linguistic activities and primarily adds a thin layer to the basic model [7]. Figure 2 shows the categorization of various BERT (Fine Tuning) based applications.
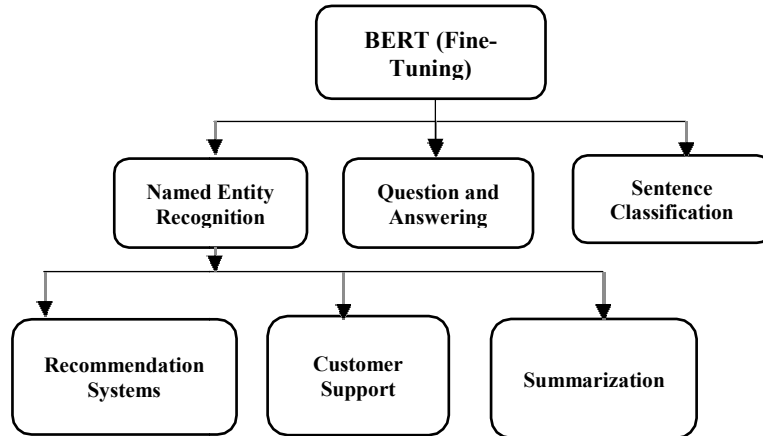

Figure 2. BERT (Fine Tuning) based applications

BERT supports NER model training by providing each tokens output vector into the classification layer for predicting named entities. A Question and Answering model can be trained using BERT by adding two additional vectors to identify the beginning and end sentences. BERT Question-Answering involves feeding a model with a question and context, and the model then generates an answer based on the context. In sentiment analysis, the classification layer is included in the transformer top for performing classifications. As the BERT model overcome the issues of ambiguity, the greater number of relations can be generated that helps to generate many knowledge level questions.

RELATED WORK

In this section, various models for Relation Extraction (RE) are explored. Relation extraction is used to understand the relationships among the various entities in an unlabeled text. There are various methods to perform relation extraction, from a simple string extraction to automated models.

A. Models for Relation Extraction

Recently, many works such as document-level, pipelined and joint model is proposed to solve the Relation Extraction tasks.

1. Pipelined Method: The pipeline method treats NER and Relation Categorization as a distinct operation. Zexuan et al. suggested the new state-of-the-art for entity and RE using a straightforward pipelined strategy, and they obtain a relative improvement over the earlier joint models using a similar pre-trained encoder [8].

2. Joint model: Joint extraction model recognizes entities and relations simultaneously and these models extract entities and relations using a single

task. Feature-based structured systems compose the majority of joint techniques. Zheng et al suggested atagging scheme to convert joint extraction of entities and relations [9].

3. Document-level Relation Extraction Models: When compared to sentence-level Relation Extraction, document level Relation Extraction is a complex process. Because document may contain entity pairs with multiple relationships. The Sentence Importance Estimation and Focusing (SIEF) framework was presented for document-level. In various disciplines, the SIEF framework enhances the performance of basic models [10]. Zeng et al. proposed an architecture to distinguishthe document level based on the intra and inter sequential reasoning techniques [11].

B. Contextual Word Embeddings for Relation Extraction

Word embeddings are a method for finding similarities between words in a corpus by predictingthe co-occurrence of words in a text using some sort of model. When it was proven that word embeddings could be used to find analogies, they became well-known in the field of automated text analysis. Table 1 illustrates various word embedding techniques.

Table 1. Word Embedding Techniques

S. No	Word Embeddings	Explanation	Feature
1	TF-IDF	A statistical technique for determining a word's relevance to the corpus of text. It doesn't record word associations with semantic meaning.	Perform well on retrieving information and extracting keywords from documents.
2	Word2Vec	CBOW and Skip-gram architecturesbased on neural networks are superior at capturing semanti cinformation.	Suitable for smaller and larger datasets.
3	GLoVe	Global word-word co-occurrence-based matrix factorization. It resolves Word2Vec's local context issues.	Better in tasks that involve word analogies and named-entity recognition. Word2Vec is commonly used in semantic analysis tasks.

4	BERT	High-quality contextual information can be captured via a transformer-based attention method.	Translation services and a question-and-answer platform are used in the Google Search engine to interpret search keywords.

Contextual embeddings represent each word based on its context, capture the word usage across a range of situations and encode cross-linguistic knowledge. Contextual embeddings, such as ELMo and BERT, perform significantly better than generic word representations. The ELMo of the bidirectional Language Model combines the representations from its intermediary layer according to the task at hand. When ELMo combines both the representations of forward and backward LSTMs, the interactions between the left and right contexts are not taken into consideration [12]. BERT offers Masked Language Modeling (MLM) that involves randomly masking some of the tokens in input sequence. It employs a Transformer encoder during pre- training to focus on instances involving bi-directional communication and the other one is Next Sentence Prediction (NSP). RE with Distant Supervision and Transformers suggested by Despina et al. predicts better embeddings using fine-tuning BERT [13]. ELMo and BERT perform better than Word2Vec, and offer ground-breaking performance in a range of NLP applications. Using two input sentences, natural language processing (NLP) determines whether the preceding sentence follows the first one. NLP helps to facilitate the tasks which needs sentence pairs analysis.

C. Datasets for Relation Extraction

Several datasets for relation extraction have been developed recently to enhance the relation extraction systems. Two examples of RE datasets created through human annotations with relation types are SemEval-2010 Task 8 and ACE05. The crowdsourcing method is used to build TACRED dataset to meet the demands of the large-scale dataset. To enhance document-level RE research, DocRED was developed. Ten thousand annotated examples and more than one hundred relations are included in FewRel. The issues with few-short relation extraction have been addressed with the development of FewRel and FewRel 2.0. HacRED consists of 65,225 relational facts that has been identified from 9,231 documents [14-18].

ANALYSIS OF DEEP LEARNING MODELS

Deep learning uses artificial neural networks using representation learning. It can be supervised, semi-supervised, or unsupervised. The rapid growth and use of Artificial Intelligence based systems have elevated concerns regarding understandability [19]. Rahman et al. constructed artificial neural networks model for effectively forecast solar radiation [20]. Representation learning helps to reduce the data dimension to simplify in identifying patterns and anomalies. A neural network instructs computers to scrutinize data like human brain. The hidden layers are referred as the term "deep." Deep neural networks consist of 150 hidden layers, compared to the two or three layers that traditional neural networks normally have. The structure of deep neural network is depicted in Figure 3.
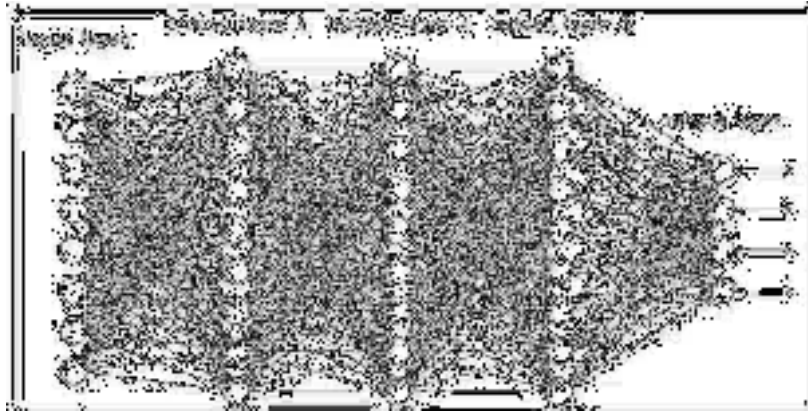
Figure 3. Structure of Deep Neural Network

Deep learning model helps to learn categorization that take input from the various sources such as images, text, and sounds. It can also attain high accuracy, occasionally even superior human performance. Large labeled data and multi-layered architectures are used to train the models to learn data characteristics automatically. Deep learning has the ability to achieve high levels of accuracy when trained on huge amounts of data. There are many complex problems to solve in natural language. In some specific natural language problems, deep learning achieves the best results. Table. 2 illustrates some of the deep learning techniques that are widely used in the task of RE. Survey on existing RE model on Deep learning Techniques using various dataset are mentioned in Table. 3.

Table 2. Deep Learning Techniques

S. No	Models	Working Principle	Benefits	Issues
1	CNN [21-23]	Convolutional Neural Network has Multiple layers to process and extract features.	human supervision is not required for features recognizing.	Overfitting, exploding gradient, and class imbalance.
2	Bi-GRU [24]	Model that combines the Gated Recurrent Unit (GRU) and the bidirectional Recurrent Neural Network.	Simple than LSTM	Only input and forget gate.
3	LSTM [25]	Long Short-Term Memory picks up and remembers enduring addictions and long-term retention of past knowledge.	Offers parameters like learning rates, input, and output biases.	Overfitting
4	CRF [26]	It's a discriminative model to predict contextual information.	Perform well on NLP tasks such as part of speech tagging, NER.	More accurate but difficult to train.

Research Cell: An International Journal of Engineering Science, Special Issue March 2023, Vol. 35,
A Peer reviewed and refreed journal, UGC Approved Journal (S.No.63019) (till May 2018)
ISSN: 2229-6913(Print), ISSN: 2320-0332(Online), Web Presence: http://ijoe.vidyapublications.com
© 2023 Vidya Publications

5	BiLSTM [27]	Bidirectional recurrent neural networks are the combination of two separate RNNs. The networks access both forward and backward information.	Better predictions compared to Auto Regressive Integrated Moving Average (ARIMA).	Slower and requires more time.
6	RNN [28]	RNN has connections that form directed cycles that allows the current phase to accept the LSTM outputs as inputs. Due to its internal memory, the LSTM's output can recall previous inputs and is used as an input in the current phase.	Remembers every piece of information through time.	Exploding gradient problem, and long-term dependency of words.
7	MLPs [29]	Made up of many layers of perceptron with activation capabilities. Layers of input and output are interconnected and have equal layers for the input and output.	Used to solve complex nonlinear problems.	Feature scaling, and Computational complexity.
8	DBNs [30]	Made up of a lot of latent and random layers. Latent variables, often called hidden units that are characterized by binary values. Boltzmann machines has connections between its layers.	Powerful and learn complex patterns. Process large amounts of data very quickly.	Hardware requirements, expensive to train.
9	RBM [31]	Consists of both visible and hidden components. All hidden units are linked to all visible units.	Computationally efficient and faster than a typical Boltzmann Machine.	Hard to evaluate or simulate.

Table. 3 Comparison of Deep Learning Relation Extraction Models

S. No	Authors & Year	Objective	Techniques	Issues	Dataset
1	Chen Gao et al 2022 [32]	BERT encoder based on the transformer model is used. Extracts the semantic mutuality between entity and relation extraction.	HBT, WDec and CasREL	Overlapping entities in the sentence cannot be resolved by this technique.	New York Times (NYT), WebNLG
2	O.A Tarasova et al 2022 [33]	Method to extract Clinical Named Entities from texts which combines the naive Bayes classifier with specially built filters.	Naïve Bayes classifier	The result of CNER using naive-Bayes method is slightly worse.	CHEMDNER
3	T.Bai et al 2022 [34]	Segment attention method based on CNN to extract local semantic properties through word embedding.	SVM, KNN, CNN, and SEGATT-CNN	This model applies only to supervised methods	Herbal-Disease and Herbal Chemistry, HD-HC
4	Qingbang W et al 2022 [35]	An Entity-relationship extraction model for food public opinion. This model efficiently predicts the information and semantic context of the current text.	BERT-BiLSTM, BiLSTM-ATT	The BERT-BLSTM network does not function well when dealing with the issue of partial entity overlap.	Food public opinion field data
5	Hailin Wang et al 2022 [36]	Supervised and distant supervision methods for Relation Extraction.	DNN, RNN and PCNN	Error propagation in supervised methods.	SemEval 2010-task8, ACE series and NYT+Freebase
6	Yang Yang et al 2022 [37]	Basics of IE and DL, mainly concentrating on DL technologies in the field of IE.	RNN, CNN and BiLSTM	DNN models cannot handle all the knowledge in huge database.	COVID-19 news

7	Zhiyun Z et al 2022 [38]	Distant Supervised Relation Extraction (DSRE) model using residual network.	CNN-ATT, PCNN-ATT, and DSRE	Noise label reducing.	Freebase + NYT
8	Guangyao Wang et al 2022 [39]	Weighted graph convolutional network (WGCN) model to extract the nontaxonomic relationships.	LSTM, CNN and BiLSTM	When the feature graph is used as an input to the GCN, the directed graph's effect is not better.	Human-annotated RE data, NYT data
9	Chantrapor nchai et al 2021 [40]	BERT and spacy model to extract specific information from entire texts based on machine learning.	BERT, Spacy	The Performance of SpaCy is poor.	Tourism Data
10	W. Zhou et al 2021 [41]	The multi-label and multi-entity problems are solved using adaptive thresholding and localized context pooling.	BERT-ATLOP, BERT-E	Adaptive thresholding only work swhen the modelis optimized.	DocRED, CDR
11	Prashant S et al 2021 [42]	Attention Retrieval Model (ARM) framework to improve the applicability of attention-based model for Relation Extraction.	LSTM, RNN, and GRU	The AR Mtechnique must test the model rather than categorize the text for understanding the languag econtext betweenthe repeated words in a text.	Atlas of Inflammatio n Resolution (AIR), BioGRID, and ChemProt.
12	Liu Kang et al 2020 [43]	Neural relation extraction with a specific focus to train neural relation extraction model.	BERT, LSTM, and BiLSTM	Unable to meet demand i npractical applications.	ACE, SemEval 2010, TACRE.

13	Boran Hao et al 2020 [44]	Novel joint training technique is used to develop language model pre-training for clinical corpora using information from the Unified Medical Language System (UMLS) knowledge base.	Clinical BERT + BiLSTM, Clinical KB-ALBERT	The improvement for ALBERT is less significant.	MIMIC-III and UMLS Knowledge Base
14	Diana Sousa et al 2020 [45]	BiOnt uses four different kinds of biomedical ontologies to perform relation extraction.	BO-LSTM, BioBERT	This approach does not allow for the integration of ontological knowledge.	DDI corpus, PGR corpus, BC5CDR corpus
15	Rakesh Patra et al 2019 [46]	A Model for automatic generation of named entity distractor. A combination of statistical and semantic similarity is used.	Frequency based, Co-occurrence based	Existing techniques focus on language elearning and vocabulary testing, these metrics are not applicable for evaluating the named entity ydistractors.	200 cricket-related MCQ-key pairs
16	Veera Ragavendra et al 2018 [47]	The rule-based method for relationship classes.	SVM, BiLSTM	Suitable only for smaller number of samples.	I2b2 2010
17	S. Zeng et al 2018 [48]	Separate Intra- and Inter-sentential Reasoning for Document-level Relation Extraction (SIRE) architecture to effectively represents intra- and inter-sentential relationships in different ways.	BiLSTM, BERT, and SIRE-BERT	This model primarily enhances intra-sentential relations' performance.	DocRED, CDR, and GDA

18	Jing Qiu et al 2018 [49]	SGNRI model to extract non-taxonomic associations using a multi-phase correlation search automated system.	SGNRI(Word2Vec) SGNRI(LDA)	The performance of the word2Vec-based model is poor.	Concept Pairs
19	Henghui et al 2018 [50]	Developed a model for the purpose of clinical feature extraction using a contextual word embedding approach.	BiLSTM-CRF, ELMo	Difficulties in creating a language model on a big corpus of domain- specific data.	I2B2 2010
20	Linfeng Song et al 2018 [51]	Graph-state LSTM model for displaying discourse and relationship structures.	Bidirectional DAG LSTM, GLSTM	Word sens econfusion	Biomedical domain

DISCUSSION

The comparison of existing relation models with various techniques shows that BERT based relation extraction model provides significantly improved performance than other models such as CNN, RNN, KNN, etc. BERT reads text input in both left-to-right and right-to-left directions at once. Using this bidirectional capability, BERT is pretrained on two different NLP tasks such as Masked Language Modeling and Next Sentence Prediction. It is observed that the model can be used for various domains such as clinical, tourism, agriculture, and so on. Table 4 shows the performance evaluation of existing relation extraction models based on deep learning techniques. Table 5 lists the performance accuracy (F1 score) of the BERT, CNN, and RNN based models for the SemEval 2010 dataset. The F1 statistical metric is employed to calculate an estimation of the deep learning model's accuracy. From the literature survey, it has been identified that the BERT-BiLSTM-CRF model achieves better results for breast cancer concepts and their attributes extraction. Even though several BERT based relation extraction for different fields are developed, the overlapping of relation and partial entity overlapping are still in a development state.

Table. 4 Performance evaluation of existing relation extraction models.

Model	Dataset	F1 score	Reference
SIRE-BERT	DocRED	62.05	[9]
RoBERT-ATLOP	DocRED	63.40	[41]
MDL-J3E	COVID-19 News	70.96	[28]
KNN	CDR	71.49	[34]
BiDAG LSTM	Biomedical	75.6	[51]
BERT	Tourism data	77.96	[40]

SGNRI	Concept pairs	81.4	[49]
KB-BERT	I2b2 2010	84.4	[44]
BiLSTM	SemEval 2010	84.7	[6]
WCGN	NYT Data	84.47	[27]
BERT-BiLSTM	Food public opinion field data	87.44	[35]
ELMo+BiLSTM-CRF	I2b2/VA 2010	88.60	[50]
BERT-BiLSTM-CRF	Clinical data	96.73	[52]

Table. 5 Comparison of CNN, RNN, and BERT based RE models (SemEval 2010 dataset).

Model	F1 Score	Reference
Att+CNN	88.0	[53]
CR-CNN	84.1	[54]
MVRNN	82.4	[55]
BRCNN	85.4	[56]
Att+BiLSTM	84.0	[57]
R-BERT	89.2	[58]
BERT-GCN	90.2	[59]

Relation extraction offers a wide range of applications including information retrieval, question answering, and knowledge base construction, etc. Creating models that can extract relationships in a multilingual and cross-lingual situation is another important area of focus. Additionally, the combination of relation extraction with other NLP tasks such as named entity recognition and event extraction is expected to lead to more wide-ranging and sophisticated NLP systems. The contexts such as syntax, pragmatics can be considered for improving the relation prediction accuracy. BERT variants such as RoBERTa, DistilBERT, and XLNet can be incorporated to enhance the contextual relation prediction.

CONCLUSION

This paper provides information on conceptual relation extraction and the various techniques used. It affords information on various deep learning models which are used in different tasks such as building classification models, developing recommendation systems, learning behavior predictions, and so on. It has been identified that BERT- based models can provide better accuracy to identify relations based on their context from multiple sentences. While comparing to other models BERT- BiLSTM-CRF achieved 97% of accuracy with limited information. In future, the overlapping relations problems can be focused to improve prediction accuracy.

REFERENCES

[1] X. Chen and R. Badlani, "Proceedings of Deep Learning Inside Out (DeeLIO): The First Workshop on Knowledge Extraction and Integration for Deep Learning Architectures, pages Relation Extraction with Contextualized Relation Embedding (CRE)." [Online]. Available: https://developers.google.com/

[2] P. Li, M. Wang, and J. Wang,"Named entity translation method based on machine-translation lexicon," Neural Comput Appl, vol. 33, no. 9,pp. 3977–3985, May 2021, doi: 10.1007/s00521- 020-05509-y.

[3] C. Gao, X. Zhang, M. Han, and H. Liu, "A review on cyber security named entity recognition,"Frontiers of Information Technology and Electronic Engineering, vol. 22, no. 9. Zhejiang University, pp. 1153–1168, Sep. 01, 2021. doi: 10.1631/FITEE.2000286.

[4] R. Patra and S. K. Saha, "A hybrid approach for automatic generation of named entity distractors formultiple choice questions," Educ Inf Technol(Dordr), vol. 24, no. 2, pp. 973–993, Mar. 2019, doi: 10.1007/s10639- 018-9814-3.

[5] B. Qiao, Z. Zou, Y. Huang, K. Fang, X. Zhu, and Y. Chen, "A joint model for entity and relation extraction based on BERT," Neural Comput Appl, vol. 34, no. 5, pp. 3471–3481, Mar. 2022, doi: 10.1007/s00521- 021-05815-z.

[6] H. Zhu, I. Ch. Paschalidis, and A. Tahmasebi, "Clinical Concept Extraction with Contextual Word Embedding," Oct. 2018, [Online]. Available: http://arxiv.org/abs/1810.10566

[7] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," Oct. 2018, [Online]. Available: http://arxiv.org/abs/1810.04805

[8] Z. Zhong and D. Chen, "A Frustratingly Easy Approach for Entity and Relation Extraction,"
Oct. 2020, [Online]. Available: http://arxiv.org/abs/2010.12812

[9] S. Zheng, F. Wang, H. Bao, Y. Hao, P. Zhou, and B. Xu, "Joint extractionof entities and relations based on a novel tagging scheme," in ACL2017 - 55th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers), 2017, vol.1, pp. 1227– 1236. doi: 10.18653/v1/P17-1113.

[10] W. Xu, K. Chen, L. Mou, and T. Zhao, "Document-Level Relation Extraction with Sentences Importance Estimation and Focusing." [Online].Available: https://github.

[11] S. Zeng, Y. Wu, and B. Chang, "SIRE: Separate Intra- and Inter- sentential Reasoning for Document- level Relation Extraction," Jun. 2021, [Online]. Available: http://arxiv.org/abs/2106.01709

[12] M. E. Peters et al., "Deep contextualized word representations," Feb. 2018, [Online]. Available: http://arxiv.org/abs/1802.05365.

[13] D. Christou and G. Tsoumakas, "Improving Distantly-Supervised Re- lation Extraction through BERT-Based Label and Instance Embed- dings," IEEE Access, vol. 9, pp. 62574–62582, 2021, doi:10.1109/AC- CESS.2021.3073428.

[14] I. Hendrickx et al., "SemEval-2010 Task 8: Multi-Way Classificationof Semantic Relations Between Pairs of Nominals," Association for Computational Linguistics, 2010. [Online]. Available: http://docs.

[15] Y. Yao et al., "DocRED: A Large-Scale Document-Level Relation Extraction

Dataset." [Online].
Available: https://spacy.io

[16] X. Han et al., "FewRel: A Large-Scale Supervised Few-Shot Relation Classification Dataset with State-of-the-Art Evaluation." [Online]. Avail-able: http://zhuhao.me/fewrel

[17] T. Gao et al,"FewRel 2.0: Towards More Challenging Few-Shot Relation Classification." [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/

[18] Q. Cheng et al., "HacRED: A Large-Scale Relation Extraction Dataset Toward Hard Cases in Practical Applications." [Online]. Available: http://lic2019.ccf.org.cn/kg

[19] Haque, A. B., Islam, A. N., & Mikalef, P. (2023). Explainable Artificial Intelligence (XAI)from a user perspective: A synthesis of prior literature and problematizing avenues for future research. Technological Forecasting and Social Change, 186, 122120.

[20] Rahman, S., Rahman, S., & Bahalul Haque, A. K. M. (2022). Automated detection of cardiac arrhythmia based on a hybrid CNN-LSTM network. In Emergent Converging Technologies and Biomedical Systems: Select Proceedings of ETBS 2021 (pp. 395-414). Singapore: Springer Singapore.

[21] G. Kim, C. Lee, J. Jo, and H. Lim, "Automatic extraction of named entities of cyber threats using a deep Bi-LSTM-CRF network," Inter- national Journal of Machine Learning and Cybernetics, vol. 11, no. 10, pp. 2341–2355, Oct. 2020, doi: 10.1007/s13042-020-01122-6.

[22] Navid, S. M. A., Priya, S. H., Khandakar, N. H., Ferdous, Z., & Haque, A. B. (2019). Signature verification using convolutional neural network. In 2019 IEEE International Conference on Robotics, Automation, Artificial-intelligence and Internet-of-Things (RAAICON) (pp. 35-39). IEEE.

[23] Siam, S. C., Faisal, A., Mahrab, N., Haque, A. B., & Suvon, M. N. I. (2021, February). Automated student review system with computer vision and convolutional neural network. In 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS) (pp. 493-497). IEEE.

[24] P. Srivastava, S. Bej, K. Schultz, K. Yordanova, and O. Wolkenhauer, "Attention Retrieval Model for Entity Relation Extraction from Biolog- ical Literature," IEEE Access, vol. 10, pp. 22429–22440,2022, doi: 10.1109/ACCESS.2022.3154820.

[25] S. Banerjee and K. Tsioutsiouliklis, "Relation Extraction Using Multi- Encoder LSTM Network on a Distant Supervised Dataset," in Proceedings - 12th IEEE International Conference on Semantic Computing, ICSC 2018, Apr. 2018, vol. 2018-January, pp. 235–238. doi: 10.1109/ICSC.2018.00040.

[26] H. Zhu, I. Ch. Paschalidis, and A. Tahmasebi, "Clinical Concept Extraction with Contextual Word Embedding," Oct. 2018, [Online]. Available: http://arxiv.org/abs/1810.10566

[27] G. Wang, S. Liu, and F. Wei, "Weighted graph convolution over dependency trees for nontaxonomic relation extraction on public opinioninformation," Applied Intelligence, vol. 52, no. 3, pp. 3403– 3417, Feb. 2022, doi: 10.1007/s10489-021-

02596-9.

[28] Y. Yang, Z. Wu, Y. Yang, S. Lian, F. Guo, and Z. Wang, "A Survey of Information Extraction Based on Deep Learning," Applied Sciences (Switzerland), vol. 12, no. 19. MDPI, Oct. 01, 2022. doi:10.3390/app12199691.

[29] S. Mitra, S. Saha, and M. Hasanuzzaman, "A multi-view deep neural network model for chemical- disease relation extraction from imbalanceddatasets," IEEE J Biomed Health Inform, vol. 24, no. 11, pp. 3315–3325,Nov. 2020, doi: 10.1109/JBHI.2020.2983365.

[30] Y. Chen, W. Li, Y. Liu, D. Zheng, and T. Zhao, "Exploring DeepBelief Network for Chinese Relation Extraction." [Online]. Available: http://www.nist.gov/speech/tests/ace/.

[31] T. M. Alam and M. J. Awan, "Domain Analysis of Information Ex- traction Techniques," INTERNATIONAL JOURNAL OF MULTIDISCI-PLINARY SCIENCES AND ENGINEERING,
vol. 9, no. 6, 2018, [On-line]. Available: https://www.researchgate.net/publication/326463350.

[32] C. Gao, X. Zhang, H. Liu, W. Yun, and J. Jiang, "A joint extraction model of entities and relations based on relation decomposition," Inter- national Journal of Machine Learning and Cybernetics, vol. 13, no. 7,pp. 1833–1845, Jul. 2022, doi: 10.1007/s13042-021-01491-6.

[33] O. A. Tarasova, A. v. Rudik, N. Y. Biziukova, D. A. Filimonov, and V. v.Poroikov, "Chemical named entity recognition in the texts of scientific publications using the na¨ıve Bayes classifier approach," J Cheminform,vol. 14, no. 1, Dec. 2022, doi: 10.1186/s13321-022-00633-4.

[34] T. Bai, H. Guan, S. Wang, Y. Wang, and L. Huang, "Traditional Chinese medicine entity relation extraction based on CNN with segmentattention," Neural Comput Appl, vol. 34, no. 4, pp. 2739– 2748, Feb. 2022, doi: 10.1007/s00521-021-05897-9.

[35] Q. Wang, Q. Zhang, M. Zuo, S. He, and B. Zhang, "An Entity Relation Extraction Model with Enhanced Position Attention in Food Domain," Neural Process Lett, vol. 54, no. 2, pp. 1449–1464, Apr. 2022, doi: 10.1007/s11063-021-10690-9.

[36] H. Wang, K. Qin, R. Y. Zakari, G. Lu, and J. Yin, "Deep neural network-based relation extraction: an overview," Neural Comput Appl, vol. 34, no. 6, pp. 4781–4801, Mar. 2022, doi: 10.1007/s00521-021-06667-3.

[37] Y. Yang, Z. Wu, Y. Yang, S. Lian, F. Guo, and Z. Wang, "A Survey of Information Extraction Basedon Deep Learning," Applied Sciences (Switzerland), vol. 12, no. 19. MDPI, Oct. 01, 2022. doi: 10.3390/app12199691.

[38] Z. Zheng, Y. Liu, D. Li, and X. Zhang, "Distant supervised relation extraction based on residual attention," Frontiers of Computer Science, vol. 16, no. 6. Higher Education Press Limited Company,
Dec. 01, 2022. doi: 10.1007/s11704-021-0474-x

[39] G. Wang, S. Liu, and F. Wei, "Weighted graph convolution over dependency trees for nontaxonomic relation extraction on public opinion information," Applied

Intelligence, vol. 52, no. 3, pp. 3403– 3417, Feb. 2022, doi: 10.1007/s10489-021-02596-9.

[40] C. Chantrapornchai and A. Tunsakul, "Information extraction on tourism domain using SpaCy and BERT," *ECTI Transactions on Computer and Information Technology*, vol. 15, no. 1, pp. 108–122, Apr. 2021, doi: 10.37936/ecti-cit.2021151.228621.

[41] W. Zhou, K. Huang, T. Ma, and J. Huang, "Document-Level Relation Extraction with Adaptive Thresholding and Localized Context Pooling," 2021. [Online]. Available: www.aaai.org

[42] P. Srivastava, S. Bej, K. Schultz, K. Yordanova, and O. Wolkenhauer, "Attention Retrieval Model for Entity Relation Extraction from Biolog- ical Literature," IEEE Access, vol. 10, pp. 22429–22440,2021, doi: 10.1109/ACCESS.2022.3154820.

[43] K. Liu, "A survey on neural relation extraction," Science China Technological Sciences, vol. 63, no. 10. Springer Verlag, pp. 1971–1989, Oct. 01, 2020. doi: 10.1007/s11431-020-1673-6.

[44] B. Hao, H. Zhu, and I. Ch Paschalidis, "Enhancing Clinical BERT Embedding using a Biomedical Knowledge Base," Online, 2020. [Online]. Available: https://github.com/noc-lab/clinical-kb-bert

[45] D.Sousa and F. M. Couto, "BiOnt: Deep learning using multiple biomed-ical ontologies for relation extraction," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2020, vol. 12036 LNCS, pp. 367–374. doi: 10.1007/978-3-030-45442-546.

[46] R. Patra and S. K. Saha, "A hybrid approach for automatic generation of named entity distractors formultiple choice questions," Educ Inf Technol (Dordr), vol. 24, no. 2, pp. 973–993, Mar. 2019, doi: 10.1007/s10639-018-9814-3.

[47] Chikka, Veera Raghavendra, and Kamalakar Karlapalem. "A hybrid deep learning approach for medical relation extraction." arXiv preprint arXiv:1806.11189 (2018).

[48] S. Zeng, Y. Wu, and B. Chang, "SIRE: Separate Intra- and Inter- sentential Reasoning for Document- level Relation Extraction," Jun. 2021, [Online]. Available: http://arxiv.org/abs/2106.01709

[49] J. Qiu, Y. Chai, Y. Liu, Z. Gu, S. Li, and Z. Tian, "Automatic Non-Taxonomic Relation Extraction from Big Data in Smart City," IEEE Access, vol. 6, pp. 74854–74864, 2018, doi: 10.1109/AC-CESS.2018.2881422.

[50] H. Zhu, I. Ch. Paschalidis, and A. Tahmasebi, "Clinical Concept Extraction with Contextual Word Embedding," Oct. 2018, [Online]. Available: http://arxiv.org/abs/1810.10566

[51] L. Song, Y. Zhang, Z. Wang, and D. Gildea, "N-ary Relation Extraction using Graph State LSTM."[Online]. Available: https://github.com/

[52] Zhang X, Zhang Y, Zhang Q, Ren Y, Qiu T, Ma J, Sun Q, " Extracting comprehensive clinical information for breast cancer using deep learning methods." Int J Med Inform. 2019 Dec;132:103985. doi:

10.1016/j.ijmedinf.2019.103985. Epub 2019 Oct 2. PMID: 31627032.

[53] L. Wang, Z. Cao, G. de Melo, and Z. Liu, "Relation Classification via Multi-Level Attention CNNs."

[54] S. Wu and Y. He, "Enriching Pre-trained Language Model with Entity Information for Relation Classification," May 2019, [Online]. Available: http://arxiv.org/abs/1905.08284

[55] D. Zeng, K. Liu, S. Lai, G. Zhou, and J. Zhao, "Relation Classification via Convolutional Deep Neural Network." [Online]. Available: http://en.wikipedia.org/wiki/Bag-of-words

[56] R. Cai, X. Zhang, and H. Wang, "Bidirectional Recurrent Convolutional Neural Network forRelation Classification."

[57] P. Zhou et al., "Attention-Based Bidirectional Long Short-Term Memory Networks for RelationClassification."

[58] S. Wu and Y. He, "Enriching Pre-trained Language Model with Entity Information for Relation Classification," May 2019, [Online]. Available: http://arxiv.org/abs/1905.08284.

[59] Y. Zhao, H. Wan, J. Gao, and Y. Lin, "Improving Relation Classification by Entity Pair Graph,"2019.