# Revisiting Document-Level Relation Extraction with Context-Guided Link Prediction

**Monika Jain**[1], **Raghava Mutharaju**[1], **Ramakanth Kavuluru**[2], **Kuldeep Singh**[3]
[1]Indraprastha Institute of Information Technology, Delhi, India
[2]University of Kentucky, Lexington, Kentucky, United States
[3]Cerence GmbH and Zerotha Research, Germany
{monikaja, raghava.mutharaju@iiitd.ac.in}, ramakanth.kavuluru@uky.edu, kuldeep.singh1@cerence.com

## Abstract

Document-level relation extraction (DocRE) poses the challenge of identifying relationships between entities within a document as opposed to the traditional RE setting where a single sentence is input. Existing approaches rely on logical reasoning or contextual cues from entities. This paper reframes document-level RE as link prediction over a knowledge graph with distinct benefits: 1) Our approach combines entity context with document-derived logical reasoning, enhancing link prediction quality. 2) Predicted links between entities offer interpretability, elucidating employed reasoning. We evaluate our approach on three benchmark datasets: DocRED, ReDocRED, and DWIE. The results indicate that our proposed method outperforms the state-of-the-art models and suggests that incorporating context-based link prediction techniques can enhance the performance of document-level relation extraction models.

## Introduction

Relation extraction is the task of extracting semantic links or connections between entities from an input text (Baldini Soares et al. 2019). In recent years, document-level relation extraction problem (DocRE) evolved as a new subtopic due to the widespread use of relational knowledge in knowledge graphs (Yu et al. 2017) and the inherent manifestation of cross-sentence relations involving multi-hop reasoning. Thus, as compared to traditional RE, DocRE has two major challenges: subject and object entities in a given triple might be dispersed across distinct sentences, and certain entities may have aliases in the form of distinct entity mentions. Consequently, the signal (hints) needed for DocRE is not confined to a single sentence. A common approach to solve this problem is by taking the input sentences and constructing a structured graph based on syntactic trees, coreferences, or heuristics to represent relation information between all entity pairs (Nan et al. 2020). A graph neural network model is applied to the constructed graph, which performs multi-hop graph convolutions to derive features of the involved entities. A classifier uses these features to make predictions (Zhang et al. 2020). Another approach (Xu, Chen, and Zhao 2021a, 2023) explicitly models the reasoning process of different reasoning skills (e.g., multi-hop, coreference-mediated). However, even after considering features between the entity pairs and executing the reasoning process, DocRE is still hard due to the latent and unspecific/imprecise contexts.

Consider the sentence in Figure 1 and its labeled relation *applies_to_jurisdiction* (Congress, US) from the DocRED dataset (Yao et al. 2019). Even with the inclusion of multi-hop and co-reference reasoning, inferring the correct relation becomes challenging because the relation depends on multiple sentences and cannot be identified based on the language used in the sentence.

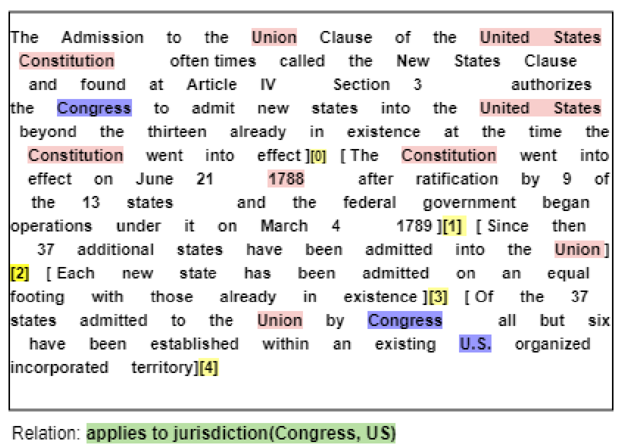


Figure 1: A partial document and labeled relation from DocRED. Blue color represents concerned entities, pink color represents other mentioned entities, and yellow color denotes the sentence number.

For these kinds of sentences, external context (knowledge) can play a vital role in helping the model capture more about the involved entities. For the above example, using the Wikidata knowledge base (Vrandečić and Krötzsch 2014) and WordNet (Miller 1995), we can get details, such as the entity types, synonyms, and other direct and indirect relations between entities (if they exist) on the Web. For these kinds of sentences, external context (knowledge) can play a vital role in helping the model capture more about the involved entities. For the above example, using the Wikidata knowledge base (Vrandečić and Krötzsch 2014), we can get additional details, such as the entity types, synonyms, and other direct and indirect relations between entities (if they exist) on the Web.

Previous research in this domain has underscored the potential of external context to enhance performance in relation extraction, co-reference resolution, and named entity recognition (Shimorina, Heinecke, and Herledan 2022). The distinctive innovation of our work lies in the fusion of context extracted from Wikidata and WordNet with a reasoning framework, enabling the prediction of entity relationships based on input document observations. Given the wide availability of external context, Knowledge Graph (KG) triples can augment training data, thereby casting the DocRE task as a knowledge Graph based link prediction challenge. In other words, given head and tail entities, we address the question of determining the appropriate relation.

We demonstrate that framing DocRE as a link prediction problem, combined with contextual knowledge and reasoning, yields enhanced accuracy in predicting the relation between entities. We furnish traversal paths as compelling justifications for relation predictions, thereby shedding light on why a particular relation is favored over others. Notably, this marks the first instance of presenting a traversal path between entities for each prediction in the context of DocRE. Our contributions in this work are as follows.

- We introduce an innovative approach named DocRE-CLiP (**Doc**ument-level **R**elation **E**xtraction with **C**ontext-guided **Li**nk **P**rediction), which amalgamates external entity context with reasoning via a link prediction algorithm.
- Our empirical analyses encompass three widely-used public document-level relation extraction datasets, showcasing our model's improvement over the recent state-of-the-art methods.
- Significantly, for every prediction, our approach is first in DocRE literature to supply a traversal path as corroborative evidence, bolstering the model's interpretability.

## Related Work

**DocRE** Previous efforts (Zhang, Qi, and Manning 2018; Jain, Singh, and Mutharaju 2023) in relation extraction have focused on predicting relationships within a single sentence. Recent years have seen growing interest in relation extraction beyond single sentence (Yao et al. 2019).

**Transformer-Based DocRE** is another interesting approach to tackle the document-level relation extraction problem (Zeng et al. 2020). One primary focus revolves around maximizing the effective utilization of long-distance token dependencies using a transformer. Earlier research considers DocRE as semantic segmentation task, employing the entity matrix, and they utilize a U-Net to capture and model (Zhang et al. 2021). In a separate study, localized contextual pooling was introduced to focus on tokens relevant to individual entity pairs (Zhou et al. 2021). On the other hand, the DocRE challenge is addressed by incorporating explicit supervision for token dependencies, achieved by leveraging evidential information (Ma, Wang, and Okazaki 2023).

**Graph-Based DocRE** is based on a graph constructed with mentions, entities, sentences, or documents. The relations between these nodes are then deduced through reasoning on this constructed graph. Earlier research in this line of work solves DocRE with multi-hop reasoning on a mention level graph for inter-sentential entity pair (Zeng et al. 2020). A discriminative reasoning framework (DRN) is introduced in a different study. This framework involves modeling the pathways of reasoning skills connecting various pairs of entities (Xu, Chen, and Zhao 2021a). DRN is designed to estimate the relation probability distribution of different reasoning paths based on the constructed graph and vectorized document contexts for each entity pair, thereby recognizing their relation. We have used DRN as a base model for implementing reasoning skills.

**Context Knowledge Based RE** study the context integration with model primarily through knowledge base (KB). Past work in this line of work uses entity types and entity aliases to predict the relation (Vashishth et al. 2018; Fernàndez-Cañellas et al. 2020). RECON (Bastos et al. 2021) encoded attribute and relation triples in the Knowledge Graph and combined the embedding with their corresponding sentence embedding. KB-Both (Verlinden et al. 2021) uses entity details from hyperlinked text documents from Wikipedia and Knowledge Graph (KG) from Wikidata to enhance performance. In distinction to that (Wang et al. 2022) integrate knowledge, including co-references, attributes, and relations with different injection methods, for improving the state-of-the-art. In contrast to these approaches, we consider the context of entities and the external relation paths between them. Furthermore, we employ a reasoning model that effectively addresses DocRE and enhances the robustness of the proposed approach.

## Methodology

### Problem Formulation

An unstructured document D consisting of K sentences is represented by $\{S\}_{i=1}^{k}$, where each sentence is a sequence of words and entities $\mathcal{E}$=$\{e_i\}_{i=1}^{P}$(P is the total number of entities). Entities $e_i$ has multiple mentions, $m_i^{s_k}$, scattered across the document D. Entity alias are represented as $\{m_i^{s_k}\}_{k=1}^{Q}$. Our objective is to extract the relation between two entities in $\mathcal{E}$ namely P(r|e$_i$,e$_j$) where $e_i, e_j \in \mathcal{E}, r \in \mathcal{R}$, here $\mathcal{R}$ is a total labeled relation set. The context (background knowledge) of an entity $e_i$ is represented by $C_{e_i}$ and a context path, i.e., a sequence of connected entities and edges from the head entity (e$_i$) to the tail entity (e$_j$) is represented by $CP_{e_i,e_j}$.

### Approach

Our proposed framework, DocRE-CLiP, integrates document-derived reasoning with context knowledge using link prediction. In the first step, we extract triples from the sentences of the given document. In the second step, we extract two types of context: 1) entity context, such as its aliases, and 2) context paths from an external KB (Wikidata, in this case). Using the triples and extracted contexts, we create a context graph to calculate a link prediction score. Then in the third step, we use several reasoning mechanisms such as logical reasoning, intra-sentence reasoning, and co-reference reasoning to calculate relation scores for pairs of entities. In the final step, the aggregation module combines the relation scores from the second and third steps. We have also implemented a path-based beam search in the framework to explain the predicted relation by providing traversal paths based on scores (refer to Figure 3). We now detail the architecture of our proposed framework.

**Triplet Extraction Module.** Document Relation Extraction (DocRE) datasets often contain labeled triplets; however, popular datasets (Yao et al. 2019) have about 64.6% missing triples, yielding an incomplete graph (Tan et al. 2022b). To extract all the triples from the document, we utilize an open-source state-of-the-art method (Huguet Cabot and Navigli 2021) for triplet generation. This model is chosen based on source code availability and run time. This module takes a document $D$ as input and produces triples ($s$, $p$, $o$), where $s$ is the head entity, $p$ is the relation, and $o$ is the tail entity. These extracted triples ($T$) follow the equation below, where $n$ is the total count of triples extracted from document $D$.

$$T(D) = \{s_i, p_i, o_i\}_{i=1}^{n} \tag{1}$$

**Context Module.** Our goal involves extracting two types of contexts – entity context and the contextual path between entity pairs. For entity $e_i$, we generate entity context using entity type and synonyms, which are derived by using WordNet (Miller 1995). We incorporate the entity context $C_{e_i}$ into the triples. Here, $S$ denotes the total number of extracted synonyms. Refer to equations 2 and 3 for more details.

$$C_{e_i} = \{e_i, hasSynonym, Synonym_k\}_{k=1}^{S}, \tag{2}$$

$$C_{e_i} = \{e_i, hasEntityType, EntityType\} \tag{3}$$

Let us consider an entity "USA" as an example. When we utilize WordNet, we discover that synonyms for "US" include, "America", and "United States". Additionally, the entity is categorized as the type "Country". Consequently, the following set of triples is generated through this process:

{*USA, hasSynonym, US*}
{*USA, hasSynonym, America*}
{*USA, hasSynonym, United States*}
{*USA, hasEntityType, Country*}

The second source of contextual information pertains to entity paths. Predicting relations between entity pairs poses challenges stemming from inherent document deficiencies (Tan et al. 2022b). To address these issues, we introduced external context by harnessing insights from Wikidata. The procedure involves extracting paths (direct and indirect) between entity-entity, mention-entity, and entity-mention pairs from Wikidata, provided they exist. The contextual path pertains to an entity pair $e_i, e_j$. We considered context paths spanning an N-hop distance (N being chosen based on experimental findings) between the entity pair. Subsequently, the extracted path is transformed into triples and forwarded to the link prediction model.

Illustrated in Figure 2 is the triple generation process employing a contextual path. The entity *Canadian* is two hops away from the entity *Ontario*. *Canada* is an intermediary entity, while *country* and *ethnic group* are intermediary properties. The Contextual Path ($CP_{e_i,e_j}$) are a set of triples formed using the intermediary entities and properties, as shown in Figure 2.

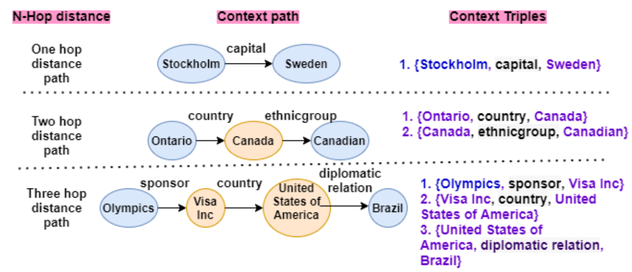


Figure 2: Triples constructed using N-hop path extracted from Wikidata. The head and tail entities are blue in color. Intermediate entities are in peach color.

**Link Prediction Module.** Link prediction is the task of predicting absent or potential connections among nodes within a network (Liben-Nowell and Kleinberg 2003). Given that the document relation extraction (DocRE) task involves constructing a graph that interlinks entities and considering our context, formulated as triples, which can be conceptualized as a Knowledge Graph (KG), we approach the DocRE challenge as a link prediction problem. This approach encompasses both an encoder and a decoder. The encoder maps each entity $e_i \in \mathcal{E}$ to a real-valued vector $v_i \in \mathbb{R}^d$, where $\mathbb{R}$ denotes the set of real-valued relation vectors of dimension $d$. The decoder reconstructs graph edges by leveraging vertex representations, essentially scoring (subject, relation, object) triples using a function: $\mathbb{R}^d \times \mathcal{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$. While prior methods often employ a solitary real-valued vector $e_i \in \mathcal{E}$, our approach computes representations using an R-GCN (Schlichtkrull et al. 2018) encoder, where $h_i^{(l)}$ is the hidden state of node $e_i$ in the l-th layer of the neural network. To compute the forward pass for an entity $e_i$ in a relational multi-graph, the propagation model at layer $l+1$ is computed as follows.

$$h_i^{(l+1)} = \sigma\left(\sum_{r \in R} \sum_{j \in N_i^r} \left(\frac{1}{c_{i,r}} W_r^{(l)} h_j^{(l)} + W_0^{(l)} h_i^{(l)}\right)\right) \tag{4}$$

Here, $h_i^{(l)} \in \mathbb{R}^{d^{(l)}}$ with $d^{(l)}$ being the dimensionality of layer $l$. $W_0^l$ and $W_r^{(l)}$ represent the block diagonal weight matrices of the neural network, and $\sigma$ represents the activation function. $N_i^r$ signifies the set of neighboring indices of node $i$ under relation $r \in R$, and $c_{i,r}$ is a normalization constant.

For training the link prediction model, our dataset comprises a) core training triples from the dataset, b) triplets obtained through the triplet extraction module using Equation 1, c) triplets formulated using the context module guided by

Equations 2 and 3, and d) triplets constructed using context paths connecting entity pairs. We use DistMult (Yang et al. 2015) as the decoder. It performs well on the standard link prediction benchmarks. Every relation r in a triple is scored using equation 5.

$$P(r \mid i, j) = P(e_i^T \times R_r \times e_j) \tag{5}$$

**Reasoning Module.** We consider three types of reasoning in our approach.
1) *Intra-sentence reasoning*, which is a combination of pattern recognition and common sense reasoning. Intra-sentence reasoning path is defined as $PI_{ij} = m_i^{s_1} \circ s_1 \circ m_j^{s_1}$ for entity pair $\{e_i, e_j\}$ insider the sentence $s_1$ in document D. $m_i^{s_1}$ and $m_j^{s_1}$ are mentions and "o" denotes reasoning step on reasoning path from $e_i$ to $e_j$.
2) *Logical reasoning* is where a bridge entity indirectly establishes the relations between two entities. Logical reasoning path is formally denoted as $PL_{ij} = m_i^{s_1} \circ s_1 \circ m_l^{s_1} \circ m_l^{s_2} \circ s_2 \circ m_j^{s_2}$ for entity pair $\{e_i, e_j\}$ from sentence $s_1$ and $s_2$ is directly established by bridge entity $e_l$.
3) *Co-reference reasoning* which is nothing but co-reference resolution. Co-reference reasoning path is defined as $PC_{ij} = m_i^{s_1} \circ s_1 \circ s_2 \circ m_j^{s_2}$ between two entities $e_i$ and $e_j$ which occur in same sentence as other entity. Our implementation of these reasoning skills is inspired by (Xu, Chen, and Zhao 2021a).

Consider an entity pair $\{e_i, e_j\}$ and its intra sentence reasoning path (PI$_{ij}$), logical reasoning path (PL$_{ij}$) and co-reference reasoning path (PC$_{ij}$) in the sentence. The various reasoning is modeled to recognize the entity pair as intra-sentence reasoning $\mathrm{R}_{PI}(r) = P\left(r \mid e_i, e_j, PI_{ij}, D\right)$, logical reasoning $\mathrm{R}_{PL}(r) = P\left(r \mid e_i, e_j, PL_{ij}, D\right)$ and co-reference reasoning $\mathrm{R}_{PC}(r) = P\left(r \mid e_i, e_j, PC_{ij}, D\right)$. Reasoning type is selected with max probability to recognize the relation between each entity pair using the equation:

$$P\left(r \mid e_i, e_j, D\right) = \max\left[\mathrm{R}_{PI}(r), \mathbf{R}_{PL}(r), \mathbf{R}_{PC}(r)\right] \tag{6}$$

For discerning relations between two entities, we employ two categories of context representation – heterogeneous graph context representation (HGC) and document-level context representation (DLC) to model diverse reasoning paths (Zhou et al. 2021). In heterogeneous graph context representation (HGC), a word is portrayed as a concatenation embedding of its word ($W_e$), entity type ($W_t$), and co-reference embedding ($W_c$). This composite embedding is then input into a BiLSTM to convert the document $D$ into a vectorized form using the equation: BiLSTM([W$_e$:W$_t$:W$_c$]). Following the methodology of (Zeng et al. 2015), a heterogeneous graph is constructed based on sentence and mention nodes. For document-level context representation, following (Eisenbach et al. 2023), a self-attention mechanism is employed to learn document-level context (DLC) for a specific mention based on the vectorized input document $D$.

To model intra-sentence reasoning path ($\alpha_{ij}$), logical reasoning path ($\beta_{ij}$) and co-reference reasoning path ($\gamma_{ij}$), HGC and DLC representation are combined (Xu, Chen, and Zhao 2021b). These reasoning representations are the input to the classifier to compute the probabilities of the relation between $e_i$ and $e_j$ entities by a multi-layer perceptron (MLP) for each path, respectively (equation 7).

$$P\left(r \mid e_i, e_j\right) = \max\begin{bmatrix}\operatorname{sigmoid}\left(\mathrm{MLP}_r\left(\alpha_{ij}\right)\right), \\ \operatorname{sigmoid}\left(\mathrm{MLP}_r\left(\beta_{ij}\right)\right), \\ \operatorname{sigmoid}\left(\mathrm{MLP}_r\left(\gamma_{ij}\right)\right)\end{bmatrix} \tag{7}$$

By the end of this step, we will get the score of each relation ($r$) for a given $\{e_i, e_j\}$.

**Aggregation Module.** In this module, we aggregate the probability score from the reasoning module equation 7 and link prediction probability score using equation 5. Further, the binary cross-entropy is used as a training objective function (Yao et al. 2019) for predicting the final relation.

### Path-Based Beam Search

An essential component of our approach is that it can explain the predicted relation by providing the most relevant path in the graph between the given entity pairs. This represents a notable advancement, as contemporary state-of-the-art models cannot often furnish explanations alongside their predictions. Unlike Greedy search, where each position is assessed in isolation, and the best choice is selected without considering preceding positions, we use beam search. This strategy selects the top "N" sequences thus far and factors in probabilities involving the concatenation of all previous paths and the path in the current position.

Inspired by (Rossi et al. 2022), we used beam search to derive plausible paths leading to the target entity within a graph. This graph (G) is constructed using the triples to train the link prediction module, augmented by test result triplets from the model's predictions. Our objective is to create a comprehensive graph encompassing the maximum available details, to generate substantial explanations for the predictions. Formally, we conceptualize the path-based beam search challenge as follows. Given a structured relational query ($e_i$, $r$, ?), where $e_i$ serves as the head entity, $r$ signifies the query relation, and ($e_i$, $r$, $e_j$) $\in$ G, our objective is to identify a collection of plausible answer entities e$_j$ by navigating paths through the existing entities and relations within G, leading to tail entities. We compile a list of distinct entities reached during the final search step and assign the highest score attained among all paths leading to each entity. Subsequently, we present the top-ranked unique entities. This approach surpasses the direct output of entities ranked at the beam's apex, which often includes duplicates. Upon completing this step, we obtain actual paths (sequences of nodes and edges) for enhanced interpretability.

## Experimental Setup

We conduct our evaluation in response to the following research questions. **RQ1**: What is the effectiveness of DocRE-CLiP that combines context knowledge with reasoning in solving document-level relation extraction tasks? **RQ2**: How does knowledge encoded from external sources impact the performance of DocRE-CLiP? **RQ3**: Does the explanation generated by our approach provide sufficient grounds to support the inferred relation?

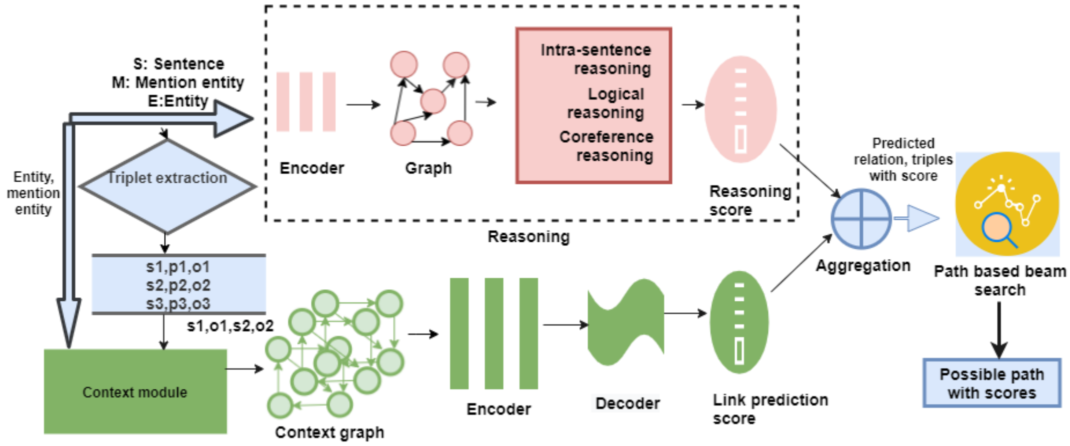


Figure 3: Illustration of proposed framework DocRE-CLiP and its various modules.

## Datasets

The proposed model is evaluated on three widely-used public datasets 1) DocRED (Yao et al. 2019), containing 50,503 triples, 5053 documents and 96 relation 2) ReDocRED (Ma, Wang, and Okazaki 2023) containing 120,664 triples, 5053 documents, 96 relation and 3) DWIE (Zaporojets et al. 2021). ReDocRED is a revised version of handling DocRED issues such as false negatives and incompleteness (Tan et al. 2022b), containing 19465 triples, 777 documents, and 66 relations.

## Baseline Models for Comparison

We used several competitive baselines and a recent state-of-the-art dataset for comparison. For DocRED, we compared our approach with BERT based models such as SIRE (Zeng, Wu, and Chang 2021), HeterGSAN-Rec (Xu, Chen, and Zhao 2021b), ATLOP (Zhou et al. 2021), DRN (Xu, Chen, and Zhao 2021a), and RoBERTa based model such as DREEAM (Ma, Wang, and Okazaki 2023). and ATLOP (Zhou et al. 2021), KD-DocRE (Tan et al. 2022a), DocuNet (Zhang et al. 2021), EIDER (Xie et al. 2022), SAIS (Xiao et al. 2022), (Zhou et al. 2021) are also compared with the proposed DocRE-CLiP.

Similarly, for the ReDocRED dataset, we used ATLOP (Zhou et al. 2021), DRN (Xu, Chen, and Zhao 2021a), DocuNet (Zhang et al. 2021), KD-DocRE (Tan et al. 2022a) and the best baseline DREEAM$_{inference}$ (Ma, Wang, and Okazaki 2023). For the DWIE dataset, we considered the state-of-the-art model DRN (Xu, Chen, and Zhao 2021a). Other than baseline model, we evaluated our model with context-based models such as KIRE (Wang et al. 2022), RESIDE (Vashishth et al. 2018), RECON (Bastos et al. 2021) and KB-graph (Verlinden et al. 2021).

## Hyper-Parameters and Metrics

For the reasoning module, we follow the settings of (Xu, Chen, and Zhao 2021a). We use the word embedding from GloVe (100d) and apply a Bidirectional LSTM (128d) to a word representation for encoding. We employ uncased BERT-Based model (768d) as an encoder with a learning rate 1e-3. We used AdamW as an optimizer, and the learning rate is $1e-3$. R-GCN is used as an encoder with a single encoding layer (200d) embeddings for the link prediction model. We regularize the encoder through edge dropout applied before normalization, with a dropout rate of 0.2 for self-loops and 0.4 for other edges. We apply l2 regularization to the decoder with a penalty of 0.01. Adam (Kingma and Ba 2015) is used as an optimizer, and the model is trained with 100 epochs using a learning rate of 0.01. For extracting the context paths, we use SPARQL queries to retrieve paths between entities. If multiple paths exist between entities, we consider the path with the highest page rank. The N-hop path length of the context varies from 1 to 4. The rationale behind this range is that we found no pertinent information for the context beyond four hops. We have used beam size of 128 for beam search for all three datasets.

We use the evaluation metrics of DocRED (Yao et al. 2019), i.e., F1 and Ign F1 for DocRE-CLiP. Ign F1 is measured by removing relations in the annotated training set from the development and test sets.

# Results

We have compared DocRE-CLiP with various baseline models on DocRED, ReDocRED, and DWIE datasets given in Table 1. The results effectively address our primary research question (**RQ1**). To delve into the specifics of **(RQ1)**, we observe that incorporating context information from Wikidata and WordNet improves the performance compared to the baseline models. Notably, DocRE-CLiP surpasses all graph-based, reasoning-oriented, and transformer-based models by incorporating contextual information.

Examining the results for the DocRED dataset, our DocRE-CLiP model showcases an improvement of approximately 1% compared to top-performing models like KD-DocRE and DREEAM. For the ReDocRED dataset, DocRE-CLiP outperforms baseline models, including DREEAM, KD-DocRE, and DocuNet. Furthermore, in the case of DWIE dataset, our model outperforms all the baseline

| Dataset | Baseline | PLM/GNN | Dev | | Test | |
|---|---|---|---|---|---|---|
| | | | F1 | Ign F1 | F1 | Ign F1 |
| DocRED | SIRE | BERT | 61.6 | 59.82 | 62.05 | 60.18 |
| | HeterGSAN-Rec | BERT | 60.18 | 58.13 | 59.45 | 57.12 |
| | ATLOP | BERT | 61.09 | 59.22 | 61.30 | 59.31 |
| | DRN | BERT | 61.39 | 59.33 | 61.37 | 59.15 |
| | DocuNet | RoBERTa | 64.12 | 62.23 | 64.55 | 62.39 |
| | ATLOP | RoBERTa | 63.18 | 61.32 | 63.40 | 61.39 |
| | KD-DocRE | RoBERTa | 67.12 | 65.27 | 67.28 | 65.24 |
| | SAIS | RoBERTa | 65.17 | 62.23 | 65.11 | 63.44 |
| | DREEAM | RoBERTa | 67.41 | 65.52 | 67.53 | 65.47 |
| | EIDER | RoBERTa | 64.27 | 62.34 | 64.79 | 62.85 |
| | KIRE | - | 52.65 | 50.46 | 51.98 | 49.69 |
| | RESIDE | GNN | 51.59 | 49.64 | 50.71 | 48.62 |
| | RECON | GNN | 52.89 | 50.78 | 52.27 | 49.97 |
| | KB-Graph | - | 52.81 | 50.69 | 52.19 | 49.88 |
| | DocRE-CLiP | BERT | $68.13_{\pm0.15}$ | $66.43_{\pm0.17}$ | 68.51 | 66.31 |
| DWIE | DRN*$_{GloVe}$ | BERT | - | - | 56.04 | 54.22 |
| | RESIDE | GNN | 65.11 | 55.74 | 66.78 | 57.64 |
| | RECON | GNN | 65.48 | 56.12 | 66.94 | 58.02 |
| | KB-Graph | - | 65.39 | 56.03 | 66.89 | 57.94 |
| | DocRE-CLiP | BERT | $66.12_{\pm0.12}$ | $57.11_{\pm0.16}$ | $67.10_{\pm0.11}$ | $58.87_{\pm0.17}$ |
| ReDocRED | ATLOP | BERT | - | - | 77.56 | 76.82 |
| | DRN* | BERT | - | - | 75.6 | 74.3 |
| | KD-DocRE | BERT | - | - | 81.04 | 80.32 |
| | DocuNet | RoBERTa | - | - | 79.46 | 78.52 |
| | DREEAM | RoBERTa | - | - | 81.44 | 80.39 |
| | DocRE-CLiP | BERT | - | - | $81.55_{\pm0.14}$ | $80.57_{\pm0.22}$ |

Table 1: Results on DocRED, ReDocRED, and DWIE datasets, including the baseline models. The precision column is blank (-) for baselines that do not report it. * denotes results obtained after modifying their code as the dataset necessitates. The mean and standard deviation of F1 and IgnF1 on the dev set are reported for three training runs. We report the official test score for DocRED on the best checkpoint on the dev set.

models, such as DRN, GAIN, and ATLOP, KIRE. Notably, the ReDocRED dataset exhibits only slight improvement over the recent state-of-the-art DREEAM. This could be attributed to ReDocRED being an enhanced version of DocRED, which already tackled the issue of dataset incompleteness. On the other hand, both DocRED and DWIE demonstrate significant improvements by incorporating REBEL triplets, setting them apart from ReDocRED in this regard. The unique aspect of DocRE-CLiP lies in its incorporation of contextual paths between entities in addition to entity context, contributing to its superior performance compared to context-aware approaches like KIRE, KB-both, and RESIDE, which leverage knowledge bases. This advantage is rooted in the reasoning framework that underlies DocRE-CLiP.

## Ablation Study

**Study of the path for an explanation on DocRE-CLiP.** Since explanation is an important part of our model, we analyze the performance of the traversal path. Table 2 provides a few examples of explanation with respect to the hops for document-level relation extraction. This answers **RQ3**.

**Effectiveness of different contexts on DocRE-CLiP.** Figure 4 offers an overview of our findings into DocRE-CLiP's performance under varying context conditions. Initially, we gauged its performance without context and documented the outcomes. Subsequently, we introduced document triplets along with the dataset labels and determined the corresponding F1 scores. Additionally, we measured the impact of entity context and documented the ensuing performance. Furthermore, we scrutinized the effect of context path details on DocRE-CLiP by considering only those paths that notably enhanced performance. Our analysis establishes that the incorporation of context enhances performance across all datasets. Thus, we effectively address our second research question, **(RQ2)**. This study's findings lead us to conclude that DocRE-CLiP benefits the most from the context path compared to the other contexts.

## Case Study

We discuss one successful and one failed case of DocRE-CLiP and compare them with the baseline model, DRN (Table 3). **Case1**: To identify the relation between Manche and France in sentence 0, we use external knowledge about France and Mancy. We get connecting context path using DocRE-CLiP: {*France, con-*

**Query**: (IBM research Brazil, parent_organization, ?x)
**Answer**: IBM research
**Explanation**:
*IBM research Brazil, part_of, IBM research*

**Query**: (Piraeus, country, ?x)
**Answer**: Greece
**Explanation**:
{*Piraeus,located_in_the_administrative _territorial_entity, Kiato*}
{*Kiato, country, Greece*}

**Query**: (Quincy, country, ?x)
**Answer**: United States
**Explanation**:
{*Quincy, country, American*}
{*American, country_of_citizenship, America*}
{*America, synonym, United States*}

Table 2: Example queries and results on DocRED dataset

*tains_the_administrative_territorial_entity, Normandy*} {*Normandy, contains_the_administrative_territorial_entity, Manche*}. Following the context path directly leads to relation *contains_the_administrative_territorial* between France and Mance. **Case2**: Using pattern recognition, DocRE-CLiP identifies the publication date as a relation. However, the inclusion of the entity context, such as "Troublemaker", "description", and "song", does not contribute significantly to the accuracy of the prediction. Consequently, DocRE-CLiP encounters difficulties in correctly predicting the relation.

**Case1**:
**Sentence 0**: The Château. de Pirou is a castle in the commune of Pirou in the département of **Manche** ( Normandy ) **France**]
**Correct answer**: contains_administrative_territorial entity
**Baseline**: country
**DocRE-CLiP**: contains_administrative_territorial_entity

**Case2**:
**Sentence 1**:[ Taking a more electronic music sound than his previous releases TY.O was released in **December 2011** by Universal Island Records but for reasons unknown to Cruz its British and American release were held off.] **Sentence 3**:[ TY.O features a range of top - twenty and top - thirty singles including "Hangover" (featuring Florida) **"Troublemaker"** "There She Goes" (sometimes featuring Pitbull) the limited release "World in Our Hands" and "Fast Car" which features on the Special Edition and Fast Hits versions of the album]
**Correct answer**: Inception
**Baseline**: publication_date
**DocRE-CLiP:** publication_date

Table 3: Case study with DocRE-CLiP prediction. Underline text represents entities in the sentence

## Conclusion

This paper introduces DocRE-CLiP, a context-driven approach for document-level relation extraction (DocRE).

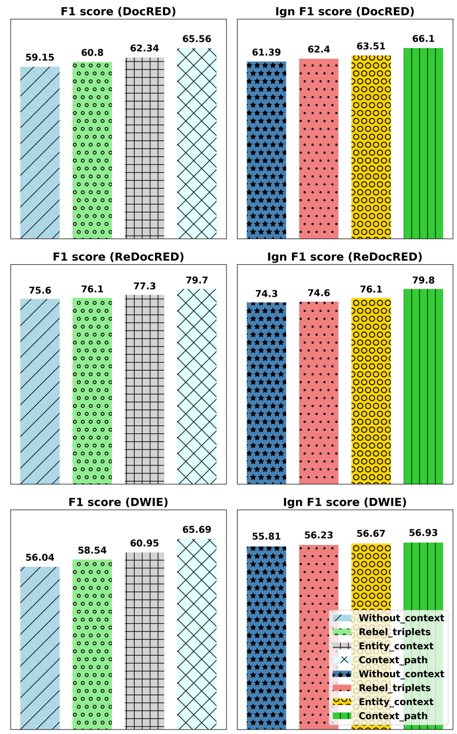


Figure 4: Performance of DocRE-CLiP across various contexts using the DocRED, ReDocRED, and DWIE datasets

Our results suggest that integrating diverse context types into the link prediction module enriches relation prediction within the DocRE framework providing interpretability. As future work, researchers can extend our efforts by crafting a versatile model capable of traversing diverse document types, thereby significantly amplifying its aptitude for assimilating knowledge. Furthermore, with conclusive evidence provided in our work as the first step, the document RE and KG link prediction research findings will interchangeably benefit. The code and the models are available at https://github.com/kracr/document-level-relation-extraction.

## Acknowledgements

We express our sincere gratitude to the Infosys Centre for Artificial Intelligence (CAI) at IIIT-Delhi for their support. RK's effort has been supported by the U.S. National Library of Medicine (through grant R01LM013240)

## References

Baldini Soares, L.; FitzGerald, N.; Ling, J.; and Kwiatkowski, T. 2019. Matching the Blanks: Distributional Similarity for Relation Learning. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 2895–2905. Florence, Italy: Association for Computational Linguistics.

Bastos, A.; Nadgeri, A.; Singh, K.; Mulang', I. O.; Shekarpour, S.; Hoffart, J.; and Kaul, M. 2021. RECON: Relation

Extraction using Knowledge Graph Context in a Graph Neural Network. In Leskovec, J.; Grobelnik, M.; Najork, M.; Tang, J.; and Zia, L., eds., *WWW '21: The Web Conference 2021, Virtual Event / Ljubljana, Slovenia, April 19-23, 2021*, 1673–1685. ACM / IW3C2.

Eisenbach, M.; Lübberstedt, J.; Aganian, D.; and Gross, H. 2023. A Little Bit Attention Is All You Need for Person Re-Identification. In *IEEE International Conference on Robotics and Automation, ICRA 2023, London, UK, May 29 - June 2, 2023*, 7598–7605. IEEE.

Fernàndez-Cañellas, D.; Marco Rimmek, J.; Espadaler, J.; Garolera, B.; Barja, A.; Codina, M.; Sastre, M.; Giro-i Nieto, X.; Riveiro, J. C.; and Bou-Balust, E. 2020. Enhancing Online Knowledge Graph Population with Semantic Knowledge. In Pan, J. Z.; Tamma, V.; d'Amato, C.; Janowicz, K.; Fu, B.; Polleres, A.; Seneviratne, O.; and Kagal, L., eds., *The Semantic Web – ISWC 2020*, 183–200. Cham: Springer International Publishing. ISBN 978-3-030-62419-4.

Huguet Cabot, P.-L.; and Navigli, R. 2021. REBEL: Relation Extraction By End-to-end Language generation. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, 2370–2381. Punta Cana, Dominican Republic: Association for Computational Linguistics.

Jain, M.; Singh, K.; and Mutharaju, R. 2023. ReOnto: A Neuro-Symbolic Approach for Biomedical Relation Extraction. In *Machine Learning and Knowledge Discovery in Databases: Research Track: European Conference, ECML PKDD 2023, Turin, Italy, September 18–22, 2023, Proceedings, Part IV*, 230–247. Berlin, Heidelberg: Springer-Verlag. ISBN 978-3-031-43420-4.

Kingma, D. P.; and Ba, J. 2015. Adam: A Method for Stochastic Optimization. In Bengio, Y.; and LeCun, Y., eds., *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

Liben-Nowell, D.; and Kleinberg, J. 2003. The Link Prediction Problem for Social Networks. In *Proceedings of the Twelfth International Conference on Information and Knowledge Management*, CIKM '03, 556–559. New York, NY, USA: Association for Computing Machinery. ISBN 1581137230.

Ma, Y.; Wang, A.; and Okazaki, N. 2023. DREEAM: Guiding Attention with Evidence for Improving Document-Level Relation Extraction. In Vlachos, A.; and Augenstein, I., eds., *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2023, Dubrovnik, Croatia, May 2-6, 2023*, 1963–1975. Association for Computational Linguistics.

Miller, G. A. 1995. WordNet: A Lexical Database for English. *Commun. ACM*, 38(11): 39–41.

Nan, G.; Guo, Z.; Sekulic, I.; and Lu, W. 2020. Reasoning with Latent Structure Refinement for Document-Level Relation Extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 1546–1557. Online: Association for Computational Linguistics.

Rossi, A.; Firmani, D.; Merialdo, P.; and Teofili, T. 2022. Explaining Link Prediction Systems Based on Knowledge Graph Embeddings. In *Proceedings of the 2022 International Conference on Management of Data*, SIGMOD '22, 2062–2075. New York, NY, USA: Association for Computing Machinery. ISBN 9781450392495.

Schlichtkrull, M. S.; Kipf, T. N.; Bloem, P.; van den Berg, R.; Titov, I.; and Welling, M. 2018. Modeling Relational Data with Graph Convolutional Networks. In Gangemi, A.; Navigli, R.; Vidal, M.; Hitzler, P.; Troncy, R.; Hollink, L.; Tordai, A.; and Alam, M., eds., *The Semantic Web - 15th International Conference, ESWC 2018, Heraklion, Crete, Greece, June 3-7, 2018, Proceedings*, volume 10843 of *Lecture Notes in Computer Science*, 593–607. Springer.

Shimorina, A.; Heinecke, J.; and Herledan, F. 2022. Knowledge Extraction From Texts Based on Wikidata. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Industry Track*, 297–304. Hybrid: Seattle, Washington + Online: Association for Computational Linguistics.

Tan, Q.; He, R.; Bing, L.; and Ng, H. T. 2022a. Document-Level Relation Extraction with Adaptive Focal Loss and Knowledge Distillation. In *Findings of the Association for Computational Linguistics: ACL 2022*, 1672–1681. Dublin, Ireland: Association for Computational Linguistics.

Tan, Q.; Xu, L.; Bing, L.; Ng, H. T.; and Aljunied, S. M. 2022b. Revisiting DocRED - Addressing the False Negative Problem in Relation Extraction. In Goldberg, Y.; Kozareva, Z.; and Zhang, Y., eds., *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, 8472–8487. Association for Computational Linguistics.

Vashishth, S.; Joshi, R.; Prayaga, S. S.; Bhattacharyya, C.; and Talukdar, P. 2018. RESIDE: Improving Distantly-Supervised Neural Relation Extraction using Side Information. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 1257–1266. Brussels, Belgium: Association for Computational Linguistics.

Verlinden, S.; Zaporojets, K.; Deleu, J.; Demeester, T.; and Develder, C. 2021. Injecting Knowledge Base Information into End-to-End Joint Entity and Relation Extraction and Coreference Resolution. In Zong, C.; Xia, F.; Li, W.; and Navigli, R., eds., *Findings of the Association for Computational Linguistics: ACL/IJCNLP 2021, Online Event, August 1-6, 2021*, volume ACL/IJCNLP 2021 of *Findings of ACL*, 1952–1957. Association for Computational Linguistics.

Vrandečić, D.; and Krötzsch, M. 2014. Wikidata: A Free Collaborative Knowledgebase. *Commun. ACM*, 57(10): 78–85.

Wang, X.; Wang, Z.; Sun, W.; and Hu, W. 2022. Enhancing Document-Level Relation Extraction by Entity Knowledge Injection. In Sattler, U.; Hogan, A.; Keet, C. M.; Presutti, V.; Almeida, J. P. A.; Takeda, H.; Monnin, P.; Pirrò, G.; and d'Amato, C., eds., *The Semantic Web - ISWC 2022 - 21st International Semantic Web Conference, Virtual Event, October 23-27, 2022, Proceedings*, volume 13489 of *Lecture Notes in Computer Science*, 39–56. Springer.

Xiao, Y.; Zhang, Z.; Mao, Y.; Yang, C.; and Han, J. 2022. SAIS: Supervising and Augmenting Intermediate Steps for Document-Level Relation Extraction. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2395–2409. Seattle, United States: Association for Computational Linguistics.

Xie, Y.; Shen, J.; Li, S.; Mao, Y.; and Han, J. 2022. Eider: Empowering Document-level Relation Extraction with Efficient Evidence Extraction and Inference-stage Fusion. In *Findings of the Association for Computational Linguistics: ACL 2022*, 257–268. Dublin, Ireland: Association for Computational Linguistics.

Xu, W.; Chen, K.; and Zhao, T. 2021a. Discriminative Reasoning for Document-level Relation Extraction. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 1653–1663. Online: Association for Computational Linguistics.

Xu, W.; Chen, K.; and Zhao, T. 2021b. Document-Level Relation Extraction with Reconstruction. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, 14167–14175. AAAI Press.

Xu, W.; Chen, K.; and Zhao, T. 2023. Document-Level Relation Extraction with Path Reasoning. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, 22(4).

Yang, B.; Yih, W.; He, X.; Gao, J.; and Deng, L. 2015. Embedding Entities and Relations for Learning and Inference in Knowledge Bases. In Bengio, Y.; and LeCun, Y., eds., *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

Yao, Y.; Ye, D.; Li, P.; Han, X.; Lin, Y.; Liu, Z.; Liu, Z.; Huang, L.; Zhou, J.; and Sun, M. 2019. DocRED: A Large-Scale Document-Level Relation Extraction Dataset. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 764–777. Florence, Italy: Association for Computational Linguistics.

Yu, M.; Yin, W.; Hasan, K. S.; dos Santos, C.; Xiang, B.; and Zhou, B. 2017. Improved Neural Relation Detection for Knowledge Base Question Answering. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 571–581. Vancouver, Canada: Association for Computational Linguistics.

Zaporojets, K.; Deleu, J.; Develder, C.; and Demeester, T. 2021. DWIE: An entity-centric dataset for multi-task document-level information extraction. *Inf. Process. Manag.*, 58(4): 102563.

Zeng, D.; Liu, K.; Chen, Y.; and Zhao, J. 2015. Distant Supervision for Relation Extraction via Piecewise Convolutional Neural Networks. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 1753–1762. Lisbon, Portugal: Association for Computational Linguistics.

Zeng, S.; Wu, Y.; and Chang, B. 2021. SIRE: Separate Intra- and Inter-sentential Reasoning for Document-level Relation Extraction. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 524–534. Online: Association for Computational Linguistics.

Zeng, S.; Xu, R.; Chang, B.; and Li, L. 2020. Double Graph Based Reasoning for Document-level Relation Extraction. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1630–1640. Online: Association for Computational Linguistics.

Zhang, N.; Chen, X.; Xie, X.; Deng, S.; Tan, C.; Chen, M.; Huang, F.; Si, L.; and Chen, H. 2021. Document-level Relation Extraction as Semantic Segmentation. In Zhou, Z.-H., ed., *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, 3999–4006. International Joint Conferences on Artificial Intelligence Organization. Main Track.

Zhang, Y.; Qi, P.; and Manning, C. D. 2018. Graph Convolution over Pruned Dependency Trees Improves Relation Extraction. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2205–2215. Brussels, Belgium: Association for Computational Linguistics.

Zhang, Z.; Cai, J.; Zhang, Y.; and Wang, J. 2020. Learning Hierarchy-Aware Knowledge Graph Embeddings for Link Prediction. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, 3065–3072. AAAI Press.

Zhou, W.; Huang, K.; Ma, T.; and Huang, J. 2021. Document-Level Relation Extraction with Adaptive Thresholding and Localized Context Pooling. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, 14612–14620. AAAI Press.