CoReX-GNN: A Co-reference-Aware Hybrid Transformer and Graph Neural Network for Document-Level Relation Extraction

Bruna Vasconcelos
u1488820@utah.edu

November 10, 2024

1 Introduction

Document-level relation extraction (RE) is a critical task in natural language processing (NLP) that involves identifying relationships between entities across sentences in a document. Applications include knowledge graph construction, biomedical research, and legal document analysis. Unlike sentence-level RE, document-level RE requires models to handle long-range dependencies and resolve co-references, which makes the task significantly more complex.

The proposed solution, **CoReX-GNN**, integrates transformer-based architectures with graph neural networks (GNNs) and incorporates a neural co-reference resolution module. This approach addresses the challenges of inter-sentence reasoning and co-reference resolution by explicitly modeling relationships between entities, sentences, and co-references. Progress so far includes implementing and evaluating a baseline Longformer model on the DocRED dataset. The next steps focus on integrating GNNs and co-reference resolution into the pipeline.

2 Problem Statement

The objective is to identify relationships between entities in documents containing multiple sentences. Current transformer-based approaches, such as BERT and Longformer, perform well at capturing local context but struggle with explicit inter-sentence reasoning and co-reference resolution.

Dataset and Expected Results This project uses the **DocRED dataset** [1], which is a benchmark for document-level RE. The dataset contains over 5,000 documents annotated with entity mentions and relationships. Expected results include:

- Improved F1 scores for RE compared to Longformer baselines.
- Enhanced handling of long-range dependencies and co-references.

Importance and Challenges Document-level RE is vital for extracting actionable information from unstructured text. Key challenges include:

- **Long-Range Dependencies**: Relationships may span multiple sentences or paragraphs.
- **Co-reference Resolution**: Entities are often referred to indirectly, requiring robust co-reference handling.

3 Related Work

Document-level RE has seen significant advancements, with key contributions including:

- **DocRED Dataset**: Yao et al. [1] introduced the DocRED dataset, emphasizing the importance of multi-sentence reasoning.
- **Transformer Baselines**: Models like BERT [2] and Longformer [3] provide strong baselines for RE tasks, but lack explicit inter-sentence reasoning.

- **Graph-Based Approaches**: Nan et al. [4] demonstrated the use of GNNs for modeling relationships between entities and sentences.
- **Co-reference Resolution**: Lee et al. [5] proposed a neural co-reference resolution system, which improves contextual understanding in NLP tasks.

CoReX-GNN builds on these works by integrating transformers, GNNs, and co-reference resolution into a unified framework for document-level RE.

4 Dataset

The **DocRED dataset** is the primary benchmark for this project. Key properties include:

- **Source**: Wikipedia articles annotated for entities and relationships.
- **Size**: 5,053 documents with over 96,000 entity mentions and 132,000 relationships.
- **Features**: Relationships requiring multi-sentence reasoning and co-reference resolution.

Preprocessing includes:

- Tokenization using HuggingFace tokenizers.
- Sentence segmentation and co-reference annotation with tools like Stanza and AllenNLP.

5 Technical Approach

The proposed model, **CoReX-GNN**, consists of three key components:

1. **Transformer Encoder**: Longformer captures document-wide context with sparse attention, encoding sentence-level representations.
2. **Graph Neural Network (GNN)**:
    - Nodes represent entities, sentences, and co-references.
    - Edges encode relationships based on syntax, position, and semantics.
3. **Co-reference Resolution Module**: Neural co-reference resolution identifies and resolves references, enhancing both the transformer and GNN components.

The model uses a multi-task framework, jointly optimizing named entity recognition (NER) and RE. Implementation tools include HuggingFace Transformers, PyTorch Geometric, and Stanza.

Progress so far includes the implementation of the Longformer baseline and preprocessing of the DocRED dataset. The next phase involves integrating the GNN and co-reference module.

6 Intermediate/Preliminary Results

Baseline experiments using Longformer on the DocRED dataset produced the following results:

- **Precision**: 74.3%, **Recall**: 71.1%, **F1 Score**: 72.7%.

Observations Errors were primarily due to unresolved co-references and the model's inability to explicitly model inter-sentence relationships. We hypothesize that:

- Integrating GNNs will improve the model's ability to reason across sentences.
- Incorporating a co-reference resolution module will enhance precision and recall for RE tasks.

The next steps include:

- Implementing the GNN component to model explicit relationships between entities and sentences.
- Adding co-reference resolution to improve contextual understanding and long-range dependency modeling.

References

[1] Yao, Y., Ye, D., Li, P., Han, X., Lin, Y., Liu, Z., ... & 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 (ACL)*. 1

[2] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). 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 (NAACL)*. 1

[3] Beltagy, I., Peters, M. E., & Cohan, A. (2020). Longformer: The Long-Document Transformer. In *arXiv preprint arXiv:2004.05150*. 1

[4] Nan, G., Gui, L., Long, F., He, Y. (2020). Heterogeneous Graph Representation Learning for Document-Level Relation Extraction. In *Proceedings of ACL*. 2

[5] Lee, K., He, L., Lewis, M., & Zettlemoyer, L. (2017). End-to-end Neural Coreference Resolution. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2