Word Pair Information is Important for Nested Named Entity Recognition

No Author Given

No Institute Given

Abstract. Named Entity Recognition (NER) is the task of detecting and classifying entity spans in texts. When entity spans overlap with each other, the task is called nested NER. The span-based approach is highly efficient for handling nested NER. However, in previous studies, span representations were generated by integrating the endpoint word representations of the span or by integrating all the word representations within the span, without fully considering the local dependencies among different words within the span. Furthermore, when using feature matrices for entity prediction, the dependencies between entities and the contextual information of entities in a sentence were not considered. To address these issues, we propose a **S**pan-**En**hanced **Net**work (**SEnNet**), which utilizes word pair information within the span to construct initial span representations. These span representations are then gradually enriched through interactions between different spans and the introduction of contextual information, effectively capturing both internal and external span information. Experimental results show that the proposed model achieves the best performance compared to 10 advanced baseline models on three nested datasets and one non-nested dataset. Further experimental analysis shows that fully leveraging word pair information within spans helps enhance the recognition performance of entities of different lengths, especially long entities. The source code for our model is available at: https://anonymous.4open.science/r/SEnNet-228C

Keywords: nested named entity recognition· span representation· word pair information· conditional layer normalization· long entity identification· feature matrix.

1 Introduction

Named Entity Recognition (NER) is a fundamental information extraction task in Natural Language Processing (NLP) that identifies spans of entities and their semantic categories in text. Previously, this task was primarily addressed using a sequence labeling approach, whereby each token is assigned a label [3,6]. It cannot be directly applied to nested NER, where one entity is nested within another, since each token might belong to two or more entities.

Various approaches have been proposed to tackle nested NER, including optimization based on sequence labeling [19,20], hypergraphs [18,24], sequence-to-sequence [22,27], and span-based [7,25,26]. The span-based approach have be-