# **Document-Level Relation Extraction** with a Dependency Syntax Transformer and Supervised Contrastive Learning

Ming Yang, Yijia Zhang, Santosh Kumar Banbhrani, Hongfei Lin, Mingyu Lu.

## Introduction

Studies have shown that the Transformer architecture models long-distance dependencies without regard to the syntax-level dependen• Gauss Enhancement Layer. Gaussian probability distributions are also designed to capture local information around entities.

misjudgment. Our results show that the paper's approach promotes the progress of document-level relation extraction.

cies between tokens in the sequence, which hinders its ability to model long-range dependencies.

The global information among relational triples and local information around entities is critical.

# Methods



• Supervised Contrastive Losses with **Knowledge Layer.** 

Supervised contrastive learning with domain knowledge captures global information among relational triples.

### Results

#### Table 1. Main results (%) of CDR.

Model	Dev	Test		
Widdei	$\mathbf{F1}$	F1 Precision/Recall		
Without KBs				
CNN+SDP [6]	-	58.02/76.20	65.88	
BRAN [13]	-	55.60/70.80	62.10	
Bio-Seq [7]	-	60.00/58.60	63.50	
LSR [9]	-	-/-	64.80	
SciBERT	66.01	61.00/68.42	64.50	
DSTSC_SciBERT (ours)	68.33	65.80/68.30	67.03	
Without KBs				
CAN(+CTD) [15]	-	60.52/80.48	69.08	
LSTM+CNN(+CTD) [14]	-	65.80/68.30	69.60	
DSTSC_BERT (ours)	70.89	57.60/73.81	68.65	
DSTSC_SciBERT (ours)	<b>72.16</b>	66.24/76.14	70.85	
DSTSC_BioBERT (ours)	71.34	65.54/73.33	69.22	

22

• • •

adriamycin, 30 mg / m2, cyclophosphamide, 300 mg / m2, on day 1; and hexamethylmelamine (HMM), 6 mg / kg daily, for 14 Part 1, days .Each course was repeated monthly

[P3] Hematologic toxicity was moderate and with reversible an emia developing in 71 % of patients al side effects from CPDD were universal. HMM gastrointestinal toxicity necessitated discontinuation of the drug in t



Fig. 2. Relational weights between entities in the baseline model and DSTSC model.

## Conclusion

In this work, we propose the DSTSC model for document-level relation extraction. Dependency syntax Transformer model enhances the Transformer's ability to model dependencies in long-range texts. Entities are also provided with comprehensive information. Experiments on two biomedical datasets, CDR and GDA, demonstrate that the DSTSC model outperforms existing models in document-level relation extraction.

Fig. 1. The overall structure of the methods (DSTSC) model) of the paper.

**Dependency Syntax Transformer Model** Layer.

Dependency syntax information is introduced into the Transformer to enhance the attention between tokens with dependency syntax in a sentence, and further improve the Transformer's ability to model dependencies in long sentences.

Table 2. Ablation Study of The DSTSC Model on CDR.

Model	Dev F1	Test F1	F1_cross/F1_noncros
DSTSC_SciBERT	72.16	70.85	57.34/75.82
- Dependency information	70.36	68.71	54.60/75.15
- Guess enhancement layer	71.22	69.07	55.67/75.20
- Contrast Loss Layer	71.43	69.52	56.93/75.32
- Knowledge	68.33	67.03	55.90/72.29
- all	66.01	64.50	51.25/70.73

## Discussion

In Part4 of Fig. 2, the closer the color in each grid is to red, the greater the weight of the relation between the two entities judged by the model and vice versa. Obviously, the color difference obtained by the DSTSC model is quite distinct, while the color obtained by SciBERT model is not distinct. This demonstrates that SciBERT is prone to

## References

[1]. Zhou, W., Huang, K., Ma, T., Huang, J.: Document-level relation extraction with adaptive thresholding and localized context pooling. In: AAAI 2021. vol. 35, pp. 14612–14620 (2021), https://ojs.aaai.org/index.php/AAAI/article/view/17 717



<sup>[</sup>P1] Treatment of ovarian cancer with a combination of cis - platinum, adriamycin, cyclophosphamide and hexamethylmelamine [P2] During the last 21/2 years, 38 patients with ovarian cancer were treated with a combination of cisplatinum (CPDD), 50 mg/m2