All for One and One for All: A Multi-Task Framework for Various Information Extraction Tasks

Anonymous EMNLP submission

Abstract

1 Introduction

? develop the W2NER for various NER tasks.

2 Related Work

3 Mirror

3.1 Universal Data Interface for NLU Tasks

3.2 Model Framework

Similar to **?**, we apply consecutive span link prediction. Differently, we use a seperate split matrix to represent skip-link.

3.3 Training

pre-training

4 Experiments

Limitations

Content input length and model compatibility. Multi-turn result modification. Laborious data cleaning and format unification.

Ethics Statement

All datasets are publicly available without further annotation, and the NLU tasks are traditional tasks in natural language processing communities. So there are no ethical issues.

References

Jingye Li, Hao Fei, Jiang Liu, Shengqiong Wu, Meishan Zhang, Chong Teng, Donghong Ji, and Fei Li. 2022. Unified named entity recognition as word-word relation classification. In *Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022, Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence, IAAI 2022, The Twelveth Symposium on Educational Advances in Artificial Intelligence, EAAI 2022 Virtual Event, February 22 - March 1, 2022*, pages 10965–10973. AAAI Press.

A Dataset Statistics

test

This is a section in the appendix.

Model	TANL	UIE	DeepStruct	InstructUIE	USM	UniEX	RexUIE	Mirror
PLM #Params	T5-base 220M	T5-large 770M	GLM 10B	FlanT5 11B	RoBERTa large 372M	RoBERTa large 372M	DeBERTa-v3 large 434M	DeBERTa-v3 large 434M
NAR	✗	✗	✗	✗	✓	✓	✓	✓
Single-step	✓	✓	✓	✓	✓	✓	✗	✓
Multi-span	✗	○	○	○	✗	✗	✗	✓
N-ary	✗	✗	✗	✗	✗	✗	✓	✓

Table 1: Comparisons with other systems. NAR denotes the non-autoregressive decoding strategy. Single-step represents that the model predicts results in a single step. Multi-span means the model supports multi-span extraction, e.g. the discontinuous named entity recognition task. N-ary denotes the ability of n-ary tuple extraction. ○ means the model supports the task theoretically, but the implementation is not available.