

# CCL23-Eval 任务1系统报告：基于持续预训练方法与上下文增强策略的古籍命名实体识别

王士权，石玲玲，蒲璐汶，方瑞玉，赵宇，宋双永

中国电信股份有限公司数字智能科技分公司

{wangsq23, pulw, fangry, zhaoy11, songshy}@chinatelecom.cn  
sll0107@sina.com

## 摘要

本文描述了队伍“翼智团”在CCL23古籍命名实体识别评测中提交的参赛系统。该任务旨在自动识别出古籍文本中人名、书名、官职名等事件基本构成要素的重要实体，并根据使用模型参数是否大于10b分为开放赛道和封闭赛道。该任务中，我们首先利用古籍相关的领域数据和任务数据对开源预训练模型进行持续预训练和微调，显著提升了基座模型在古籍命名实体识别任务上的性能表现。其次提出了一种基于pair-wise投票的不置信实体筛选算法用来得到候选实体，并对候选实体利用上下文增强策略进行实体识别修正。在最终的评估中，我们的系统在封闭赛道中排名第二，F1得分为95.8727。

**关键词：** 命名实体识别，持续预训练，实体修正

## System Report for CCL23-Eval Task 1: Named Entity Recognition for Ancient Books based on Continual Pre-training Method and Context Augmentation Strategy

Shiquan Wang, Lingling Shi, Luwen Pu, Ruiyu Fang, Yu Zhao, Shuangyong Song

China Telecom Corporation Ltd. Data&AI Technology Company

Beijing, China

{wangsq23, pulw, fangry, zhaoy11, songshy}@chinatelecom.cn  
sll0107@sina.com

## Abstract

This article describes the entry system submitted by our team in the CCL23 ancient book named entity recognition evaluation. The task aims to automatically identify important entities of the basic elements of events such as names of people, titles of books, and official titles in ancient texts, and divide them into open tracks and closed tracks according to whether the model parameters used are greater than 10b. In this task, we first use the domain data and task data related to ancient books to continuously pre-train and fine-tune the open source pre-training model, which significantly improves the performance of the pedestal model on the task of named entity recognition in ancient books. Secondly, an untrusted entity screening algorithm based on pair-wise voting is proposed to obtain candidate entities, and the context enhancement strategy is used to correct entity recognition for candidate entities. In the final evaluation, our system ranked second in the closed circuit with an F1 score of 95.8727.

**Keywords:** Named Entity Recognition, Continuous Pre-training, Entity Modification

## 1 引言

古籍命名实体识别任务的目标是自动识别古籍文献中人名、地名、机构名等重要实体，这是正确分析古汉语文本、深度挖掘人文知识的基础步骤(苏祺, 2023)。然而，古籍命名实体识别任务缺乏公开的用于模型训练和评测的数据资源，制约了技术的进一步发展。此外，古籍文献命名实体识别任务受古文字词含义多样性、行文结构连续性、繁体字和无句读的影响具有较高的复杂性。

我们提交的系统采用两阶段策略提高模型在古籍命名实体识别任务上的表现，第一阶段通过基于领域和任务知识的持续预训练与微调提升基座模型在古籍命名实体识别任务上的性能表现，第二阶段通过基于上下文信息的实体修正提升模型对于不置信实体的判别准确率。

我们首先基于领域和任务知识对基座模型进行持续预训练与微调，针对古籍文本复杂的语言结构、古老的语义表达和特定的文化背景，我们使用开源的二十四史数据进行领域持续预训练，增加预训练语言模型对于古籍文本的理解能力，使用评测数据集进行任务持续预训练，增加模型对于该数据集文本的理解能力，同时采用W2NER、BERT-CRF和BERT-Span等多种方式针对命名实体识别任务进行模型微调。

我们其次基于实体的上下文信息对其进行修正，我们观察到引入实体所在上下文信息后可以帮助模型更好的对其进行分类，因此针对不同模型的投票结果，我们首先设计了一种基于pair-wise投票的不置信实体筛选算法得到预测不置信的实体候选，然后通过加入实体所在二十四史篇章的上下文对识别结果进行修正。我们的系统在最终线上评测中F1得分为95.8727，获得了封闭赛道第二名的成绩。

## 2 相关工作

命名实体识别 (Name Entity Recognition) 任务旨在自动识别出文本中人名、地名、机构名等事件基本构成要素的重要实体。因为它具有各种各样的基于知识的应用，例如关系提取(Wei et al., 2020; Li et al., 2021b)、实体链接(Le and Titov, 2018; Hou et al., 2020)等，所以长期以来一直是自然语言处理(NLP)领域的一项基本任务。目前NER任务的主流方法分为四种，分别是 (1) 基于序列标注的NER方法(Lample et al., 2016); (2) 基于超图的NER方法(Lu and Roth, 2015; Katiyar and Cardie, 2018); (3) 基于序列到序列的NER方法(Yan et al., 2021); (4) 基于跨度的NER方法(Luan et al., 2019; Li et al., 2021a)。

基于序列标注的NER方法出现于早期的NER任务中，该方法将文本序列中的每个词都视为一个单独的标注单元，通过为每个词分配一个标签来识别实体。通常主流的研究方法结合了条件随机场 (CRF) (Lafferty et al., 2001)、卷积神经网络 (CNN) (Strubell et al., 2017)、循环神经网络 (RNN) (Tang et al., 2018)和Transformer(Yan et al., 2019)进行模型建模，并结合特征工程来捕捉词的上下文信息、词性等特征。该方法的优势在于其简单性和高效性。

基于超图的NER方法将文本中的实体识别任务建模为在超图上进行推理和解码的问题。超图是一个拓展了传统图模型的概念，它允许节点和边连接形成高阶组合关系。在超图中，每个节点表示一个单词，边表示词之间的关系，而超边则表示更高阶的组合关系，可以跨越多个节点。通过在超图上定义合适的特征表示和推理算法，可以捕捉到更复杂的上下文和语义信息，提高命名实体识别的性能。Lu和Roth等人(Lu and Roth, 2015)率先在NER任务中使用基于超图的方法，随后该方法被Wang和Lu等人(Wang and Lu, 2018)进行扩展和增强。

基于序列到序列的NER方法通常使用序列到序列 (Sequence-to-Sequence) 模型来解决命名实体识别 (NER) 任务。该方法使用编码器-解码器框架，其中编码器将输入序列编码为固定维度的表示，而解码器根据编码器的输出生成目标实体。Gillick等人(Gillick et al., 2016)将Seq2Seq模型应用到NER任务中，通过解码器直接输出对应句子中包含实体的起始位置、跨度长度和标签。

基于跨度的NER方法是一种使用跨度 (Span) 作为基本单位进行命名实体识别 (NER) 的方法。该方法通过枚举所有可能的跨度组合，并对每个跨度进行分类，从而确定文本中的实体位置和类型(Xu et al., 2017; Yamada et al., 2020)。Li等人(Li et al., 2020)对NER进行了重新定义，将其视为机器阅读理解(MRC)任务，并将实体作为答案跨度进行提取。Shen等人(Shen et al., 2021)提出了一种两阶段的标识方法，通过过滤器和回归器生成跨度提案，并将其分类到相应的类别中。

在本次评测中我们主要结合了基于序列标注的NER方法和基于跨度的NER方法提取古籍文本中蕴含的实体，最终取得封闭赛道第二名的成绩。

### 3 模型

在本次古籍命名实体识别评测中我们提出的模型结构如图1所示，该模型由两阶段构成，第一阶段包括基于领域和任务知识的持续预训练以及微调。领域知识持续预训练指的是利用任务相关的领域语料对开源基座进行持续训练，我们采用的领域知识来自于二十四史原文。任务知识持续预训练指利用未标注的训练集进行持续预训练，任务知识相比领域知识更具备领域相关性且训练成本更低。微调阶段我们使用了带标签的训练数据对模型进行精调；第二阶段基于上下文信息对实体识别结果进行修正，该阶段主要利用实体上下文信息对不同结构模型投票产生的不置信实体进行修正。接下来将分别介绍这两阶段的主要内容。

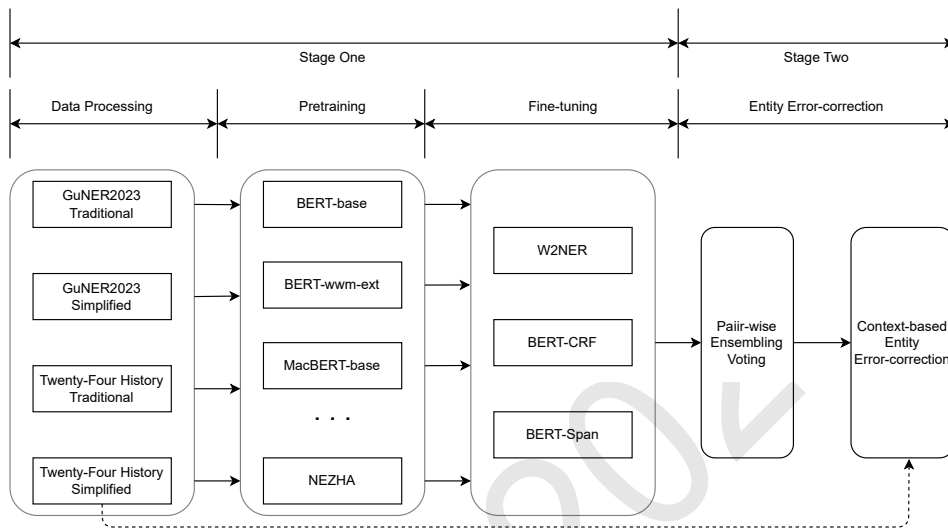


图 1: 模型结构图

#### 3.1 基于领域知识和任务数据的持续预训练与微调

##### 3.1.1 基于领域知识和任务数据的持续预训练

由于开源中文预训练语言模型在初始预训练的过程中使用的古籍数据较少，而本次赛事提供的训练集和测试集均为古籍原文数据，直接使用比赛数据进行微调，首先不能充分的利用开源预训练语言模型的语义表征能力，其次预训练语料的领域分布和任务领域分布存在较大的差异，而领域持续预训练和任务持续预训练可以通过引入领域信息和任务信息来提高模型在下游任务上的表现(Gururangan et al., 2020; Yang et al., 2021)。

首先二十四史是中国古代历史著作的集合，包含了丰富的历史文本，涵盖了各个历史时期和领域的内容。这些古籍数据对于古籍命名实体识别具有重要的参考价值；其次二十四史中包含了大量关于人名、官职名、书名等实体的描述，这些数据与本次比赛任务的领域高度相关。因此我们选择使用网络上可以搜集到的二十四史数据对开源预训练语言模型进行领域持续预训练，经过领域持续预训练之后，再利用未标注的训练集进行任务持续预训练，经过持续预训练模型可以更好地理解古籍文本的语言特点、实体命名规律和上下文关系，从而提升模型在古籍命名实体识别任务上的性能。

##### 3.1.2 基于任务数据的模型微调

在本次评测中，我们选择了三种不同的模型结构进行微调，分别是W2NER(Li et al., 2022)、BERT-CRF和BERT-Span。

其中W2NER是一种基于字级别特征的序列标注模型，它将统一的NER建模为词与词关系分类，通过利用相邻词和尾首词关系对实体词之间的相邻关系进行有效建模，解决了NER的内核瓶颈问题。BERT-CRF方法结合了BERT模型和条件随机场（CRF）来进行命名实体识别。

该模型结构通过BERT获得输入序列的特征向量，然后通过CRF层对这些特征向量进行序列建模，进而识别出命名实体的边界和类别。BERT-Span方法结合了BERT模型和Span方法来进行序列标注，与BERT-CRF相似，该模型结构同样通过BERT模型获取输入序列的特征向量，不同之处在于BERT-Span方法通过Span方法对特征向量进行处理，并基于预测实体的起始位置和结束位置来定位命名实体。该方法通过预测实体的起始和结束位置的概率分布来识别实体的边界，并根据最高概率的起始和结束位置来确定实体的边界。值得注意的是，我们使用的预训练语言模型均为经过领域知识和任务数据持续预训练后的模型。

### 3.2 基于上下文信息的实体修正

在微调训练过程中，我们发现单模型即可充分拟合训练数据集，而且不同模型对验证集的预测结果大致趋同。但是不同结构的模型，甚至同结构的模型在不同训练轮次的预测结果上仍存在稍许差异，这些差异反映了模型对候选实体的不置信程度，对这些不置信样本的精细化处理，是我们在本次评测任务中取得进一步提升的关键，具体算法如算法1所示：

---

#### 算法 1 基于pair-wise投票的不置信实体筛选

---

**关键词：**

定义 $R_i$ 为模型 $i$ 对测试集识别出的实体序列 $[entity_i^1, entity_i^2, \dots, entity_i^m]$ ;

定义修正操作 $op_k$ 为四元组 $\langle op\_type, span(entity_m^j), label(entity_m^j), label(entity_n^k) \rangle$ ，其中 $op\_type \in \{add, delete, change\}$ ,

函数 $span$ 得到实体所在的文本区域位置，

函数 $label$ 得到实体的类型，

定义 $\lambda(op_i)$ 为操作 $op_i$ 的置信度

**Input:** 单模型预测结果和对应的验证集F1分值的集

合 $RS = [\langle R_1, S_1 \rangle, \langle R_2, S_2 \rangle, \dots, \langle R_n, S_n \rangle]$ ，置信度过滤阈值 $\theta$

**Output:**  $[\lambda(op_1), \lambda(op_2), \dots, \lambda(op_n)]$

```

1 result=[]
2 for 采样两个单模型结果( $\langle R_m, S_m \rangle, \langle R_n, S_n \rangle$ ) do
3   for  $j = 1, \dots, len(R_m)$  do
4     for  $k = 1, \dots, len(R_n)$  do
5       if  $span(entity_m^j) == span(entity_n^k)$  and  $label(entity_m^j) != label(entity_n^k)$  then
6         if
7           ( $\langle op\_type, span(entity_m^j), label(entity_m^j), label(entity_n^k) \rangle$ ) not in result
8           then
9             Registering
10             $\lambda(\langle op\_type, span(entity_m^j), label(entity_m^j), label(entity_n^k) \rangle)$  to result;
11            Updating
12             $\lambda(\langle op\_type, span(entity_m^j), label(entity_n^k), label(entity_m^j) \rangle)$  to 0.0;
13          end
14           $op\_type = get\_op\_type(entity_n^k, entity_m^j)$ ;
15          Updating  $\lambda(\langle op\_type, span(entity_m^j), label(entity_m^j), label(entity_n^k) \rangle)$  by
16          adding  $S_n - S_m$ ;
17          Updating  $\lambda(\langle op\_type, span(entity_m^j), label(entity_n^k), label(entity_m^j) \rangle)$  by
18          adding  $S_m - S_n$ ;
19        end
20      end
21    end
22  end
23 end
24 Filtering the result with  $\lambda(op_i)$  smaller than  $\theta$ ;
25 return result;
```

---

我们首先保留多个经过领域和任务知识持续预训练的模型微调结果；其次利用不同模型结构的优势，利用投票方法和基于pair-wise投票的不置信实体筛选算法得到预测不置信的实体候选区域，例如测试集中，“以玉虎吐募兵人下蜀江，代失八都守中、。”的候选提及区域“玉虎吐”在不同模型的预测结果中属于不同的实体类别，因此将其视为一个不置信区域；然后利用二十四史中同段落、同篇章中的上下文对候选区域进行辅助预测，对于命名实体“玉虎吐”，我们补充其出现的段落、篇章级语义上下文“宗王囊加、玉枢虎儿吐华与脱欢悉议加封。”后，可以得到实体类别为PER。最后我们将补充上下文进行修正后的结果和原结果进行融合。

## 4 实验

本小节将介绍我们在本次评测任务中实验所涉及的内容，主要包括数据处理、实验参数设置、实验结果与分析三部分。

### 4.1 数据处理

本次评测任务提供的数据集（GuNER2023）<sup>0</sup>来源于网络上公开的二十四史文本，其中训练集包括2347条文本，测试集包括224条文本，这些文本均是从二十四史基础语料中随机截断的，长度约为100字。其中训练集标注了文本中的人名（PER）、书名（BOOK）和官职名（OFI）三种实体，测试集无标注。

由于目前大部分开源中文预训练语言模型的预训练语料大多以简体中文为主，因此我们通过繁简转换开源工具OPENCC把该数据集转换为简体中文，值得注意的是，在繁体转简体的过程中我们记录了字级别的繁简映射关系，在最终把简体预测结果转换成繁体时可以做到无损转换。除了本次评测任务提供的数据集外，我们收集了网络上公开可用的二十四史文本，包括繁体<sup>1</sup>、简体和白话文版本<sup>2</sup>。我们将其按照句号进行分割，将该部分数据和本次评测任务提供的数据集一起作为模型持续预训练的语料。

### 4.2 实验设置

我们使用的深度学习框架为PyTorch，编程语言为Python。W2NER、BERT-CRF和BERT-Span中使用的开源预训练语言模型有BERT-wwm-ext、MacBERT-base、MacBERT-large、ERNIE-large、GuwenBERT-base和MengziBERT-base等，重要参数设置如表1所示。

name	value	name	value	name	value
emb_dropout	0.5	task_type	crf	task_type	span
conv_dropout	0.5	train_epochs	10	train_epochs	10
out_dropout	0.33	swa_start	5	swa_start	5
max_epochs	500	train_batch_size	24	train_batch_size	24
batch_size	32	dropout_prob	0.1	dropout_prob	0.1
learning_rate	1e-3	max_seq_len	512	max_seq_len	512
weight_decay	0	lr	2e-5	lr	2e-5
clip_grad_norm	5.0	other_lr	2e-3	other_lr	2e-3
bert_learning_rate	5e-6	seed	123	seed	123
warm_factor	0.1	weight_decay	0.01	weight_decay	0.01
seed	13	loss_type	ls_ce	loss_type	ls_ce

(a) W2Ner参数设置

(b) BERT-CRF参数设置

(c) BERT-Span参数设置

表 1: 模型参数设置

### 4.3 实验结果与分析

#### 4.3.1 实验一

为了寻找到最适合处理古籍领域文本的基座模型，我们首先在GuNER2023数据集上基于W2NER架构尝试了若干开源预训练语言模型，实验结果如表2所示。

<sup>0</sup><https://tianchi.aliyun.com/dataset/151499>

<sup>1</sup><http://www.guoxuedashi.net/a/30p/>

<sup>2</sup><https://github.com/maxzxc0110/24-histories>

Model	ERNIE-large	GuwenBERT-base	BERT-base	MacBERT-large	MengziBERT-base	BERT-wwm-ext
GuNER2023_split	81.6057	85.4847	90.1980	90.6496	91.6667	<b>92.2014</b>

表 2: GuNER2023数据集上开源预训练语言模型的实验结果

表2展示了在GuNER2023数据集上基于不同开源预训练语言模型W2NER的性能表现，GuNER2023\_split表示将训练集数据按照9:1的比例随机划分训练集和验证集，其中的F1得分为该开源预训练语言模型在繁体数据集上和简体数据集上的最高F1得分。从实验结果可以看出BERT-wwm-ext在GuNER2023数据集上表现最佳，因此在本次评测中我们选择该模型作为后续实验的基座模型。

#### 4.3.2 实验二

由于BERT-wwm-ext在预训练的过程中使用的大多为简体中文，为了提高BERT-wwm-ext在古籍文本领域中进行命名实体识别任务时的性能表现，我们利用网络上开源的二十四史语料和GuNER2023数据集对该模型进行了领域持续预训练（DAP）和任务持续预训练（TAP），模型结构为W2NER，实验结果如表3所示。

Metric	no_trick	+DAP	+TAP
GuNER2023_split_t	90.9449	91.5805	92.9980
GuNER2023_split_s	92.2014	92.9383	93.3400
GuNER2023_s	\	93.4394	93.8958

表 3: GuNER2023数据集上基于BERT-wwm-ext持续预训练的实验结果

表3展示了在GuNER2023数据集上对BERT-wwm-ext进行持续预训练的实验结果，其中GuNER2023\_split\_t表示繁体的GuNER2023数据集并且随机抽取出十分之一的数据作为验证集，GuNER2023\_split\_s表示简体的GuNER2023数据集并且随机抽取出十分之一的数据作为验证集，GuNER2023\_s表示全部简体的GuNER2023数据集，不抽取部分数据作为验证集。no\_trick表示使用公开的BERT-wwm-ext模型，+DAP表示在公开的模型参数基础上使用获取的24史数据进行基于领域知识的持续预训练，+TAP表示在领域知识持续预训练后使用GuNER2023数据集进行基于任务的持续预训练。实验结果表明加入基于领域知识和基于任务的持续预训练后模型的性能获得了提升，在GuNER2023\_split\_t数据上分别获得了0.6356和2.0531的绝对提升，在GuNER2023\_split\_s数据上分别获得了0.7369和1.1386的绝对提升。这证明在古籍命名实体识别任务上进行领域持续预训练和任务持续预训练是有必要的。除此之外，我们还尝试了使用全部数据不划分验证集的方法进行模型训练，实验结果证明在该评测任务中使用全量数据可以获得更好的性能表现，这是因为在该任务中数据量较少，因此后续实验均使用全量数据进行模型训练。

#### 4.3.3 实验三

为了弥补单个模型结构的局限性和增加模型投票时的多样性，除W2NER外，我们选取了BERT-CRF和BERT-Span两种额外的模型结构来参与该评测任务，模型微调后实验结果如表4所示。

Metric	Precision	Recall	F1
W2NER	94.9200	94.0770	<b>94.4968</b>
BERT-CRF	94.1942	92.8924	93.5388
BERT-Span	94.6721	91.2142	92.9110

表 4: GuNER2023数据集上模型微调的实验结果

表4展示了W2NER、BERT-CRF和BERT-Span三种不同的模型结构在GuNER2023数据集上微调后的实验结果，实验结果表明基于字级别特征的W2NER可以更好的捕获词与词之间的联系，从而更好的识别出古籍文本中的重要实体，值得注意的是，我们修改了W2NER获取句向量的方法，将其从取最后BERT四层的平均输出改为取最后一层的输出，该方法获得

了0.6010的绝对提升。虽然BERT-CRF和BERT-Span单模型的表现结果不如W2NER，但是在模型投票的过程中我们发现三种不同结构的模型对于某些实体的识别结果存在差异性。

#### 4.3.4 实验四

由于不同结构的模型对于某些实体的判别存在差异，为了利用不同结构模型之间的差异性，我们设计了一种基于上下文信息进行实体修正的方法，实验结果如表5所示。

Metric	Precision	Recall	F1
<b>W2NER</b>	94.9200	94.0770	94.4968
+context_correction	95.1629	96.5932	<b>95.8727</b>

表 5: GuNER2023数据集上基于上下文进行实体修正的实验结果

表5展示了在W2NER上利用上下文信息进行实体修正的实验结果，其中+context\_correction表示在W2NER中加入上下文信息修正方法。该方法首先通过基于pair-wise投票的不置信实体筛选算法共筛选出120个不置信的候选提及区域，然后让模型对加入篇章级别的上下文信息后的实体进行预测，最终共有26个实体的类别发生了变化，我们将发生变化的实体类别和W2NER的结果进行融合，最终获得了1.3759的绝对提升。实验结果证明加入实体所在篇章的上下文信息可以改善模型对于该实体的不置信程度，从而提高模型在古籍命名实体识别任务上的性能表现。

## 5 总结与展望

在本次CCL23古籍命名实体识别评测中，我们观察到古籍文本的特殊性和复杂性以及不同结构的模型对于同一实体预测的差异性，因此提出了基于领域和任务知识的持续预训练与微调的方法和基于上下文信息的实体修正方法，通过领域和任务知识提高模型对于古籍领域文本的理解能力，通过上下文信息提高模型对于不置信实体的判别能力。我们“翼智团”队伍提交的系统在封闭赛道中排名第二，测试集上F1得分为95.8727。然而，我们的挑选不置信实体和上下文的方法还有改进的空间，未来可以通过探索更有效获取不置信实体及上下文的方法来进一步提高模型的性能表现。

## 致谢

我们的队伍在参加本次CCL23古籍命名实体识别评测中受益匪浅，在不断学习的过程中突破了自我，在持续的交流与合作中取得了进步。我们衷心感谢中国电信股份有限公司数字智能科技分公司的领导和同事们对我们技术在技术、时间和硬件方面的大力支持。同时，我们要特别感谢主办方组织本次比赛，提供了跨领域、历时的数据资源，推动了古籍资源的智能开发和利用，促进了技术的突破与发展。

## 参考文献

- Dan Gillick, Cliff Brunk, Oriol Vinyals, and Amarnag Subramanya. 2016. Multilingual language processing from bytes. In *Proceedings of NAACL-HLT*, pages 1296–1306.
- Suchin Gururangan, Ana Marasović, Swabha Swayamdipta, Kyle Lo, Iz Beltagy, Doug Downey, and Noah A. Smith. 2020. Don’t stop pretraining: Adapt language models to domains and tasks. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8342–8360.
- Feng Hou, Ruili Wang, Jun He, and Yi Zhou. 2020. Improving entity linking through semantic reinforced entity embeddings. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6843–6848.
- Arzoo Katiyar and Claire Cardie. 2018. Nested named entity recognition revisited. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 861–871.

- John D Lafferty, Andrew McCallum, and Fernando CN Pereira. 2001. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the Eighteenth International Conference on Machine Learning*, pages 282–289.
- Guillaume Lample, Miguel Ballesteros, Sandeep Subramanian, Kazuya Kawakami, and Chris Dyer. 2016. Neural architectures for named entity recognition. In *Proceedings of NAACL-HLT*, pages 260–270.
- Phong Le and Ivan Titov. 2018. Improving entity linking by modeling latent relations between mentions. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1595–1604.
- Xiaoya Li, Jingrong Feng, Yuxian Meng, Qinghong Han, Fei Wu, and Jiwei Li. 2020. A unified mrc framework for named entity recognition. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5849–5859.
- Fei Li, ZhiChao Lin, Meishan Zhang, and Donghong Ji. 2021a. A span-based model for joint overlapped and discontinuous named entity recognition. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 4814–4828.
- Jingye Li, Kang Xu, Fei Li, Hao Fei, Yafeng Ren, and Donghong Ji. 2021b. Mrn: A locally and globally mention-based reasoning network for document-level relation extraction. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 1359–1370.
- 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 *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 10965–10973.
- Wei Lu and Dan Roth. 2015. Joint mention extraction and classification with mention hypergraphs. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 857–867.
- Yi Luan, Dave Wadden, Luheng He, Amy Shah, Mari Ostendorf, and Hannaneh Hajishirzi. 2019. A general framework for information extraction using dynamic span graphs. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 3036–3046.
- Yongliang Shen, Xinyin Ma, Zeqi Tan, Shuai Zhang, Wen Wang, and Weiming Lu. 2021. Locate and label: A two-stage identifier for nested named entity recognition. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 2782–2794.
- Emma Strubell, Patrick Verga, David Belanger, and Andrew McCallum. 2017. Fast and accurate entity recognition with iterated dilated convolutions. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2670–2680.
- Buzhou Tang, Jianglu Hu, Xiaolong Wang, and Qingcai Chen. 2018. Recognizing continuous and discontinuous adverse drug reaction mentions from social media using lstm-crf. *Wireless Communications and Mobile Computing*, 2018.
- Bailin Wang and Wei Lu. 2018. Neural segmental hypergraphs for overlapping mention recognition. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 204–214.
- Zhepei Wei, Jianlin Su, Yue Wang, Yuan Tian, and Yi Chang. 2020. A novel cascade binary tagging framework for relational triple extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 1476–1488.
- Mingbin Xu, Hui Jiang, and Sedtawut Watcharawittayakul. 2017. A local detection approach for named entity recognition and mention detection. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1237–1247.
- Ikuya Yamada, Akari Asai, Hiroyuki Shindo, Hideaki Takeda, and Yuji Matsumoto. 2020. Luke: Deep contextualized entity representations with entity-aware self-attention. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 6442–6454.



- Hang Yan, Bocao Deng, Xiaonan Li, and Xipeng Qiu. 2019. Tener: adapting transformer encoder for named entity recognition. *arXiv preprint arXiv:1911.04474*.
- Hang Yan, Tao Gui, Junqi Dai, Qipeng Guo, Zheng Zhang, and Xipeng Qiu. 2021. A unified generative framework for various ner subtasks. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 5808–5822.
- Zinong Yang, Ke-jia Chen, and Jingqiang Chen. 2021. Guwen-unilm: Machine translation between ancient and modern chinese based on pre-trained models. In *Natural Language Processing and Chinese Computing: 10th CCF International Conference, NLPCC 2021, Qingdao, China, October 13–17, 2021, Proceedings, Part I 10*, pages 116–128. Springer.
- 邓泽琨杨浩 王军苏祺, 王莹莹. 2023. Ccl23-eval 任务1总结报告: 古籍命名实体识别(guner2023).