

# 基于指针标注的中文医学文本实体关系抽取研究

罗文龙, 王 勇

广东工业大学, 计算机学院, 广东 广州

收稿日期: 2021年12月24日; 录用日期: 2022年1月19日; 发布日期: 2022年1月26日

## 摘 要

随着医学领域科学技术的不断发展, 产生了大量的医学文本数据, 如何从海量的非结构化数据中获取有效的信息成为医学和自然语言处理的研究热点。作为信息抽取的关键一环, 实体关系抽取可以获得自然语言句子中实体对及其之间的语义关系。当前中文医学文本的实体关系抽取方法存在词组信息缺失和关系重叠等问题, 基于此, 本文提出了一个Flat-Lattice-指针标注联合抽取模型。利用相对位置对词组信息进行编码, 增强实体边界, 并通过指针标注框架, 将关系作为一种主实体到客实体的映射函数, 解决了关系重叠的问题。在中文医学文本数据集上与多个基准模型进行对比, 证明了该模型在中文医学文本实体关系抽取上的有效性, 其准确率、召回率和F1值均高于基准模型。

## 关键词

实体关系抽取, 中文医学文本, 关系重叠, 词组信息

# Research on Entity Relation Extraction of Chinese Medical Text Based on Pointer Tagging Framework

Wenlong Luo, Yong Wang

School of Computer, Guangdong University of Technology, Guangzhou Guangdong

Received: Dec. 24<sup>th</sup>, 2021; accepted: Jan. 19<sup>th</sup>, 2022; published: Jan. 26<sup>th</sup>, 2022

## Abstract

With the continuous development of science and technology in the medical field, a large number of

medical text data have been produced. How to obtain effective information from massive unstructured data has become a research hotspot in medicine and natural language processing. As a key part of information extraction, entity relation extraction can obtain entity pairs and their semantic relations in natural language sentences. At present, there are some problems in the entity relation extraction methods of Chinese medical texts, such as lack of word information and overlapping triples. Based on this, this paper proposes a flat-lattice-pointer-tagging joint extraction model. The word information is encoded by relative position to enhance the entity boundary. Through the pointer tagging framework, the relationship is regarded as a mapping function from subject to object, which solves the problem of overlapping. Compared with several benchmark models on Chinese medical text data set, it proves the effectiveness of the model in entity relation extraction of Chinese medical text, and its accuracy, recall and F1 value are higher than those of the benchmark model.

## Keywords

Entity Relation Extraction, Chinese Medical Text, Overlapping Triples, Word Information

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution-NonCommercial International License (CC BY-NC 4.0).

<http://creativecommons.org/licenses/by-nc/4.0/>



Open Access

## 1. 引言

随着医学研究的不断发展, 各种医学文献和医学文本等医学数据的数量不断增加。利用这些医学文本数据, 研究人员和医护人员可以从中发现、分享研究工作中遇到的新问题、新方法, 并获取到多样的医学信息。然而, 由于医学文本的存储方式往往是非结构化的, 并且这些非结构化的文本存储量大, 形式比较复杂。因此, 从海量的数据中自动、便捷的获取到有价值数据的信息抽取(Information Extraction, IE)技术应运而生。而实体关系抽取(Entity Relation Extraction)作为信息抽取中的关键一步, 近年来也受到学术界和工业界的广泛关注。实体关系抽取的目的是抽取自然语言句子中实体对及其之间的语义关系。通过实体关系抽取, 可以检测和表征医学实体之间的语义关系, 为医学知识图谱和知识库, 以及下游任务如医学问答系统等提供支持。

常规的医学实体关系抽取采用流水线的方式, 先进行实体识别抽取实体对, 然后对候选实体对之间的关系进行分类。然而, 实体识别和关系分类两个子任务之间往往是相互依赖的。由于实体识别与关系分类两个子任务之间分离, 忽略了两个子任务之间的交互性与相关性, 导致出现级联错误[1]。

在过去的几年中, 人们开始建立联合抽取模型来提取实体关系。最近的研究表明, 联合学习方法可以有效地整合实体和关系的信息, 从而在两个子任务中取得更好的表现[2]。Miwa 等人[3]提出一种端对端模型来联合表示实体和关系, 其中实体识别和关系抽取两个子任务共享 LSTM 编码层序列参数的思想, 虽然考虑了两个子任务间的交互、缓解了错误传播问题, 但模型学习过程仍然类似流水线方法, 不属于真正意义上的联合抽取。Katiyar 等人[4]在 BiLSTM 序列标注方法的基础上引入注意力机制, 使用多层双向 LSTM 将实体识别子任务建模成序列标注任务, 结合序列标注结果和共享编码层表示进行实体关系抽取, 模型可以拓展各种预定义的关系类型, 是真正意义上的第一个基于神经网络的关系联合抽取模型。Zheng 等人[5]提出基于新标注策略的实体关系抽取方法, 将联合抽取任务转化成端到端的序列标注问题而无需分别识别实体和关系, 解决了流水线方法中实体冗余问题。但医学文本实体关系抽取中, 存在如表 1 所示的实体对重叠(Entity Pair Overlap, EPO)和单实体重叠(Single Entity Overlap, SEO)的关系重叠问题。无论

是上述的参数共享还是序列标注的联合抽取模型, 均未有效解决关系重叠问题。

同时, 对于中文医学文本关系抽取, 不同于英文文本每个词组都是独立的, 中文文本的是由每个单独的字符所组成。为了导入词组的信息, 一些研究人员利用 NLP 工具对文本进行分词, 然后利用分词的结果导入词组信息。然而这种做法存在的一个问题, 就是分词的结果有可能的错误的, 而这个错误会在整个过程中不断迭代传播, 最终影响整个模型的有效性。所以目前大部分算法对 Embedding 层进行嵌入时, 通常采用字符级的嵌入方法, 以避免分词造成的错误传播。

因此, 本文提出一种 Flat-Lattice-指针标注联合抽取模型。该模型采用了 Flat-Lattice 结构, 在避免分词错误传播的情况下, 引入中文词组信息, 加强实体边界; 采用指针标注框架, 从先确定实体对再识别关系, 转换为先确定主实体(subject)和关系(relation)再识别对应的客实体(object), 并融合预训练语言模型的先验知识, 捕获医学文本的语句上下文信息, 解决医学实体关系抽取中的关系重叠问题。

**Table 1.** Different entity relation cases

**表 1.** 不同的实体关系案例

类型	文本	三元组(subject, relation, object)
Normal	局部皮肤感染引起的皮疹等。	(局部皮肤感染, 相关(导致), 皮疹)
SEO	早产可以使用倍他米松或地塞米松。	(早产, 药物治疗, 倍他米松) (早产, 药物治疗, 地塞米松)
EPO	肿瘤可能发生或转移到卵巢。	(肿瘤, 发病部位, 卵巢) (肿瘤, 转移部位, 卵巢)

## 2. 相关工作

### 2.1. 词汇增强

在命名实体识别(Named Entity Recognition, NER)任务中, lattice 结构被证明在利用中文词组信息以及避免分词错误传播上有很大的优势[6]。但 LatticeLSTM [6]存在计算性能低下、不能 batch 并行化、信息损失和可迁移性差等问题。Gui 等人[7]针对 RNN 结构无法利用全局信息的问题, 采取 CNN 对字符特征进行编码, 并采用 rethinking 机制调整词汇信息的权重, 解决词汇信息冲突问题。虽然 LR-CNN [7]速度别 LatticeLSTM 快, 但它的计算仍然复杂, 并且不具备可迁移性。Li 等人[8]设计了一种位置编码来融合 lattice 结构(Flat-Lattice), 并与 Transformer 相结合, 在获取全局信息的同时, 捕捉长距离依赖。

### 2.2. 关系重叠

Zeng [9]是最早在关系三重提取中考虑重叠三重问题的学者之一。他们引入了不同重叠模式的分类, 如表 1 所示, 并提出了一种具有复制机制的序列到序列(sequence-to-sequence, Seq2Seq)模型来提取三元组。在 Seq2Seq 模型的基础上, 他们进一步研究了提取顺序的影响, 并通过强化学习获得了相当大的提升。Fu 等人(2019)也通过使用基于图卷积网络(GCNs)的模型将文本建模为关系图, 研究了关系重叠的问题。尽管这些工作取得了一定的成功, 但它们在提取重叠三元组方面仍有很多不足之处。具体来说, 它们都将关系视为实体对的离散标签, 这种方式存在一定的问题。首先, 类别分布高度不平衡。在所有抽取的实体对中, 大多数没有形成有效的关系, 产生了过多的负例。其次, 当同一个实体参与多个有效关系(重叠三元组)时, 分类器可能会混淆。如果没有足够的训练数据, 分类器很难分辨出实体参与了哪个关系。在此情况下, 提取的三元组通常是不完整和不准确的。基于此, Wei 等人[10]提出一个级联的二元标注框架, 将关系作为 subject 到 object 的映射函数, 来解决关系重叠问题。

### 3. 模型框架

针对医学文本实体关系抽取中存在实体交叉, 中文词组信息缺失, 关系重叠导致实际抽取效果不佳的问题, 本文提出一种基于 Flat-Lattice-指针标注的联合抽取模型, 其框架图如图 1 所示。如图可以看出对于“早产可以使用倍他米松和地塞米松”, 里面有两个重叠的三元组<早产, 药物治疗, 倍他米松>和<早产, 药物治疗, 地塞米松>。在嵌入层会获取到所有可能的词组信息进行嵌入, 然后输出所有可能的 subject, 对于每个特定的 subject, 如“早产”, 提取对应的编码层特征, 与编码层作一个叠加, 输出在特定关系下的 object。如图 1 所示, 在关系“药物治疗”下, 存在两个 object, 分别是“倍他米松”和“地塞米松”。

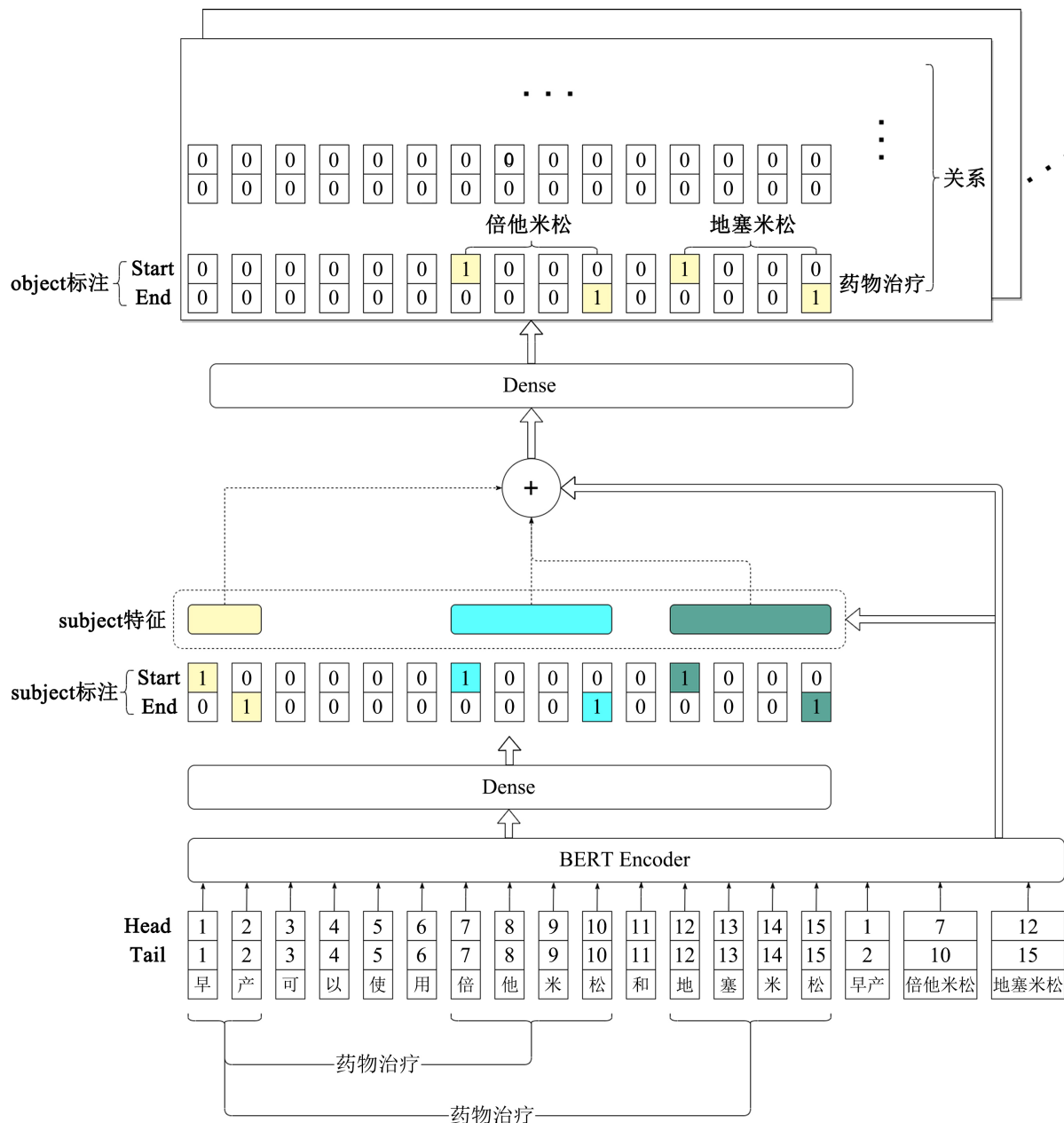


Figure 1. Overall architecture of the model  
图 1. 模型总体架构

模型主要分为以下几个部分:

1) Flat-Lattice-Transformer 编码器: 利用每个文段之间的相互作用进行编码, 采用的基于跨度的相对位置编码, 将词组信息融入到模型中进行训练。

2) 指针标注: 主要分为两个部分, 一个的 subject 的标注, 通过头尾指针标注, 获取到所有的 subject; 另一个是特定关系下的 object 标注, 对每一个 subject, 标注出对应关系下可能的 object。

### 3.1. Flat-Lattice-Transformer 编码器

首先, 从词汇表中获取到所有的词格, 转换到 Flat-Lattice 结构。一个 Flat-Lattice 可以被定义为一组文段, 一个文段对应一个 token, 及其相应的头位置和尾位置。token 代表一个字符或者词组, 头位置和尾位置代表 token 在原始序列中的位置索引。

使用基于相对位置编码对不同的文段之间的相互关系进行编码。对于两个的文段, 它们之间的有交叉、包含和分离三种位置关系。首先, 对于每个字符或者词组, 都有其对应的头、尾位置信息。head[i] 和 tail[i] 分别表示  $x_i$  在句子中的头尾位置。文段跨度信息不仅可以表示两个 token 之间的关系, 还可以表示更详细的信息, 例如一个字符和一个单词之间的位置距离。对于两个输入  $x_i, x_j$ , 他们头尾之间的相对距离表示如下:

$$d_{ij}^{(hh)} = head[i] - head[j] \quad (1)$$

$$d_{ij}^{(ht)} = head[i] - tail[j] \quad (2)$$

$$d_{ij}^{(th)} = tail[i] - head[j] \quad (3)$$

$$d_{ij}^{(tt)} = tail[i] - tail[j] \quad (4)$$

然后通过四个相对距离的简单非线性变化, 得到文段跨度的最终相对位置编码, 如下所示:

$$R_{ij} = ReLU \left( W_r \left( P_{d_{ij}^{(hh)}} \oplus P_{d_{ij}^{(ht)}} \oplus P_{d_{ij}^{(th)}} \oplus P_{d_{ij}^{(tt)}} \right) \right) \quad (5)$$

其中  $W_r$  是可学习的参数,  $\oplus$  表示拼接操作,  $P_d$  的计算如下所示:

$$P_d^{(2k)} = \sin \left( d / 10000^{2k/d_{model}} \right) \quad (6)$$

$$P_d^{(2k+1)} = \cos \left( d / 10000^{2k/d_{model}} \right) \quad (7)$$

其中,  $k$  表示位置编码的维度下标索引。

最后, 使用一个自注意力机制的变种[11], 来编码相对位置:

$$A_{i,j}^* = W_q^T E_{x_i}^T E_{x_j} W_{k,E} + W_q^T E_{x_i}^T R_{ij} W_{k,R} + u^T E_{x_j} W_{k,E} + v^T R_{ij} W_{k,R} \quad (8)$$

其中,  $W_{k,E}, W_{k,R}, W_q, u, v$  是可学习的参数,  $E_i$  为对应输入的嵌入向量, 其他的计算方式与原始的自注意力机制一致。

### 3.2. 指针标注

采用级联的指针的标注方案。首先, 从输入的句子中检测 subject。然后对于每个候选 subject, 检查所有可能的关系, 看看一个关系是否能将句子中的 object 与那个 subject 关联起来。subject 和 object 都通过起始与结束位置指针标注出来。

### 3.2.1. Subject 标注

通过 Flat-Lattice-Transformer 编码器的输出  $X_i$ , 使用两个相同的二元分类器得出两个位置指针标注, 该指针标注表示当前 subject 的起始与结束位置。指正标注如下所示:

$$p_i^{\text{subject\_start}} = \sigma(W_{\text{start}} X_i + b_{\text{start}}) \quad (9)$$

$$p_i^{\text{subject\_end}} = \sigma(W_{\text{end}} X_i + b_{\text{end}}) \quad (10)$$

其中,  $p_i^{\text{subject\_start}}$  和  $p_i^{\text{subject\_end}}$  表示句子中当前 token 是一个 subject 的起始或者结束位置的概率。如果概率超过某个阈值, 相应的标注将被设置为 1, 否则为 0。 $\sigma$  表示 sigmoid 激活函数。

subject 标注模块用以确定句子  $X$  中 subject 位置范围的优化函数如下:

$$p_\theta(s|X) = \prod_{t \in \{\text{subject\_start}, \text{subject\_end}\}} \prod_{i=1}^L (p_i^t)^{I\{y_i^t=1\}} (1-p_i^t)^{I\{y_i^t=0\}} \quad (11)$$

其中,  $L$  是句子的长度。如果  $z$  为真, 则  $I\{z\}=1$ 。 $y_i^t$  表示前述关于 subject 头尾位置的二分标注。参数  $\theta = \{W_{\text{start}}, b_{\text{start}}, W_{\text{end}}, b_{\text{end}}\}$ 。

对于多 subject 的情况, 采用最近匹配原则, 将距离最近的起始与结束标注对作为一个 subject 标注。不考虑结束标注在起始标注之前的情况, 这种匹配策略能够保持任何实体的完整性, 前提是在给定的句子中, 任何实体都能正确检测到起始位置和结束位置。

### 3.2.2. 特定关系下的 object 标注

在 subject 指针标注模块运行的同时, object 指针标注模块也同时标注出对应 subject 在特定关系下的 object。所有 subject 在各种关系下对应的 object 都会在同一时间被标注出来。与 subject 指针标注不同的是在引入编码器的输出的同时, 还引入了 subject 的特征, 具体如下:

$$p_i^{\text{object\_start}} = \sigma(W_{\text{start}}^r (X_i + V_{\text{sub}}^k) + b_{\text{start}}^r) \quad (12)$$

$$p_i^{\text{object\_end}} = \sigma(W_{\text{end}}^r (X_i + V_{\text{sub}}^k) + b_{\text{end}}^r) \quad (13)$$

其中,  $p_i^{\text{object\_start}}$  和  $p_i^{\text{object\_end}}$  分别代表句中当前 token 在对应关系下是一个 object 的起始或者结束位置的概率。 $V_{\text{sub}}^k$  表示在底层编码模块中标注到的第  $k$  个 subject 所对应的编码向量。

object 标注模块对于句子  $X$  和 subjects, 在关系  $r$  下的优化函数如下:

$$p_{\phi_r}(o|s, X) = \prod_{t \in \{\text{object\_start}, \text{object\_end}\}} \prod_{i=1}^L (p_i^t)^{I\{y_i^t=1\}} (1-p_i^t)^{I\{y_i^t=0\}} \quad (14)$$

其中, 参数  $\phi_r = \{W_{\text{start}}^r, b_{\text{start}}^r, W_{\text{end}}^r, b_{\text{end}}^r\}$ 。

最后, 损失函数如下, 使用 Adam 随机梯度下降进行训练:

$$\sum_{j=1}^{|D|} \left[ \sum_{s \in T_j} \log p_\theta(s|X_j) + \sum_{r \in T_j|s} \log p_{\phi_r}(o|s, X_j) + \sum_{r \in R \setminus T_j|s} \log p_{\phi_r}(o_\emptyset|s, X_j) \right] \quad (15)$$

## 4. 实验

### 4.1. 数据集

本文采用的数据集是 CHIP2020 评测二: 中文医学文本实体关系抽取使用的数据集, 该数据集是郑

州大学自然语言处理实验室、北京大学计算语言学教育部重点实验室、哈尔滨工业大学(深圳)、鹏城实验室人工智能研究中心智慧医疗课题组联合构建, 基于 schema 的中文医学信息抽取数据集 CMeIE (Chinese Medical Information Extraction)。数据集包含儿科训练语料和百种常见疾病训练语料, 儿科训练语料来源于 518 种儿科疾病, 百种常见疾病训练语料来源于 109 种常见疾病。近 7.5 万三元组数据, 2.8 万疾病语句和 44 种关系。经过数据预处理, 获取到训练数据集、验证数据集和测试数据集, 分别包含 14339 条, 3585 条和 4482 条数据。

## 4.2. 评价指标

当三元组的两个实体以及两者之间的关系都正确时, 实验得出的三元组才被认为是正确的。本文采用准确率(Precision, P)、召回率(Recall, R)和 F1 值(F1-score)作为中文医学文本实体关系抽取的评价指标。具体的计算方式如下:

$$P = \frac{TP}{TP + FP} \quad (16)$$

$$R = \frac{TP}{TP + FN} \quad (17)$$

$$F1 = \frac{2 \times P \times R}{P + R} \quad (18)$$

其中,  $TP$  表示所预测的三元组里正确的个数;  $FP$  表示所预测的错误的三元组个数;  $TP + FP$  表示预测出的三元组的个数;  $TP + FN$  表示实际的三元组的个数。

## 4.3. 实验结果与分析

### 4.3.1. 实验参数

利用验证数据集, 在训练的模型上进行参数调优, 得到模型的超参数如下表 2 所示。

**Table 2.** Hyperparameter of model  
**表 2.** 模型的超参数

超参数	参数值
学习率	1e-5
权重衰减	0.05
batch size	8
字向量维度	100
指针标注阈值	0.5
dropout 率	0.5

### 4.3.2. 基准模型

本文选取以下 4 个医学领域的实体关系抽取模型作为基准模型。

CNN 模型[12]。利用具有最大池化的 CNN 网络对临床医学文本进行关系抽取, 减少了对专家特征知识定义质量的依赖, 验证了深度学习方法在医学领域的适用性。

BiLSTM-Attention 模型[13]。采用 Bi-LSTM 结合注意力机制的序列模型进行实体关系抽取, 利用临

床文本中医学知识信息等进行共同学习从而抽取到最合适的关系。

**NovelTagging** 模型[5]。使用新型的序列标注方式的联合抽取模型, 将实体关系抽取转化成端到端的序列标注。然后, 基于该标签方案研究了不同的端到端模型来抽取实体及实体关系, 而不是分开识别实体和关系。实验表明基于该标签方案的方法比现在大多数流水线和联合抽取方法效果更好。

**MultiHead** 模型[14]。将实体关系抽取联合抽取看作是一个多头选择的问题, 从而使得每个实体能够与其他所有实体判断关系, 不再将关系抽取任务当作一个每个关系互斥的多分类任务, 而是看作每个关系独立的多个二分类任务, 从而能够判断每一对实体是否可能有多个关系存在, 以此解决关系重叠问题。

### 4.3.3. 实验结果

如下表 3 所示, 为实验对比结果。

**Table 3.** Results of different models

**表 3.** 不同模型的实验结果

模型	P%	R%	F1%
CNN	49.16	37.13	42.31
BiLSTM-Attention	52.77	39.28	45.04
NovelTagging	60.23	40.88	48.70
MultHead	62.50	58.20	60.27
指针标注	62.58	62.47	62.52
Flat-Lattice-指针标注	<b>63.40</b>	<b>63.45</b>	<b>63.42</b>

### 4.3.4. 结果分析

表 3 展示了不同模型下医学实体关系抽取的结果。可以从表中得知, 本文提出的 Flat-Lattice-指针标注模型在准确率、召回率和 F1 值上优于所有的基准模型, 说明本文提出模型可以有效的提升医学文本实体关系抽取的效果。其中采用了 Flat-Lattice 的指针标注模型在准确率、召回率和 F1 值上都取得了最优的效果, F1 值达到了 63.42%, 比只使用指针标注的模型高 0.9%, 也可以说明采用 Flat-Lattice 结构作为中文文本的词组信息嵌入在关系抽取任务中起到了一定的作用, 可以加深实体边界, 增强医学文本的语义特征, 解决了中文医学文本实体关系抽取中存在词组信息缺失的问题。同时, 采用的指针标注的模型对比其他的基准模型也有一定的提升, 特别是在召回率上, 反应出指针标注模型在处理关系重叠的问题上的效果, 说明了指针标注模型可以有效的解决中文医学文本实体关系抽取算法中关系重叠的问题, 增强了三元组识别的正确性和完整性。

## 5. 结语

本文针对中文医学文本实体关系抽取任务中存在词组信息缺失、关系重叠等问题, 提出了一种的基于 Flat-Lattice-指针标注的联合抽取模型。该模型通过引入基于跨度的相对位置编码, 嵌入中文词组信息, 加强实体边界; 采用指针标注框架, 从不同的角度进行实体关系抽取, 并融合预训练语言模型的先验知识, 捕获医学文本的语句上下文信息, 解决医学实体关系抽取中的关系重叠问题。在中文医学文本数据集上, 通过与各种基准模型的对比, 验证了该模型的有效性。下一步工作可以考虑引入外部知识系统来提供额外的医学信息, 使得模型具有更好医学语义特征表达能力。



## 参考文献

- [1] Li, Q. and Ji, H. (2014) Incremental Joint Extraction of Entity Mentions and Relations. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, **1**, 402-412. <https://doi.org/10.3115/v1/P14-1038>
- [2] Dai, D., Xiao, X., Lyu, Y., Dou, S. and Wang, H. (2019) Joint Extraction of Entities and Overlapping Relations Using Position-Attentive Sequence Labeling. *Proceedings of the AAAI Conference on Artificial Intelligence*, **33**, 6300-6308. <https://doi.org/10.1609/aaai.v33i01.33016300>
- [3] Miwa, M. and Bansal, M. (2016) End-to-End Relation Extraction Using LSTMs on Sequences and Tree Structures. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, **1**, 1105-1116. <https://doi.org/10.18653/v1/P16-1105>
- [4] Katiyar, A. and Cardie, C. (2017) Going out on a Limb: Joint Extraction of Entity Mentions and Relations without Dependency Trees. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, **1**, 917-928. <https://doi.org/10.18653/v1/P17-1085>
- [5] Zheng, S., Wang, F., Bao, H., Hao, Y., Zhou, P. and Xu, B. (2017) Joint Extraction of Entities and Relations Based on a Novel Tagging Scheme. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, **1**, 1227-1236. <https://doi.org/10.18653/v1/P17-1113>
- [6] Zhang, Y. and Yang, J. (2018) Chinese NER Using Lattice LSTM. *The 56th Annual Meeting of the Association for Computational Linguistics (ACL)*, **1**, 1554-1564. <https://doi.org/10.18653/v1/P18-1144>
- [7] Gui, T., Ma, R., Zhang, Q., Zhao, L. and Huang, X. (2019) CNN-Based Chinese NER with Lexicon Rethinking. *Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*, 4982-4988. <https://doi.org/10.24963/ijcai.2019/692>
- [8] Li, X., Yan, H., Qiu, X. and Huang, X. (2020) FLAT: Chinese NER Using Flat-Lattice Transformer. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 6836-6842. <https://doi.org/10.18653/v1/2020.acl-main.611>
- [9] Zeng, X., Zeng, D., He, S., Liu, K. and Zhao, J. (2018) Extracting Relational Facts by an End-to-End Neural Model with Copy Mechanism. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, **1**, 506-514. <https://doi.org/10.18653/v1/P18-1047>
- [10] Wei, Z., Su, J., Wang, Y., Tian, Y. and Chang, Y. (2019) A Novel Hierarchical Binary Tagging Framework for Relational Triple Extraction. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 1476-1488. <https://doi.org/10.18653/v1/2020.acl-main.136>
- [11] Dai, Z., Yang, Z., Yang, Y., Carbonell, J. and Salakhutdinov, R. (2019) Transformer-xl: Attentive Language Models beyond a Fixed-Length Context. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 2978-2988. <https://doi.org/10.18653/v1/P19-1285>
- [12] Sahu, S.K., Anand, A., Oruganty, K. and Gattu, M. (2016) Relation Extraction from Clinical Texts Using Domain Invariant Convolutional Neural Network. *Proceedings of the 15th Workshop on Biomedical Natural Language Processing*, Berlin, August 2016, 206-215. <https://doi.org/10.18653/v1/W16-2928>
- [13] Ramamoorthy, S. and Murugan, S. (2018) An Attentive Sequence Model for Adverse Drug Event Extraction from Biomedical Text.
- [14] Giannis, B., Johannes, D., Thomas, D. and Chris, D. (2018) Joint Entity Recognition and Relation Extraction as a Multi-Head Selection Problem. *Expert Systems with Application*, **114**, 34-45. <https://doi.org/10.1016/j.eswa.2018.07.032>