

# 基于多粒度认知的命名实体识别方法

李攀锋, 陈樱珏, 钟冷韵, 林 锋

(四川大学计算机学院, 成都 610065)

**摘 要:** 在数据匮乏的领域,命名实体识别效果受限于欠拟合的字词特征表达,引入常规的多任务学习方法可以有所改善,但需要额外的标注成本. 针对这一问题,提出了一种基于多粒度认知的命名实体识别方法,在不产生额外标注成本的前提下,增强字特征信息,提高命名实体识别效果. 该方法从多粒度认知理论出发,以 BiLSTM 和 CRF 为基础模型,将字粒度下的命名实体识别任务与句子全局粒度下的实体数量预测任务相联合,共同优化字嵌入表达. 三个不同类型的数据集上的多组实验表明,引入多粒度认知的方法有效地提升了命名实体识别效果.

**关键词:** 命名实体识别; 多粒度认知; 多任务学习; 自然语言处理

**中图分类号:** TP391 **文献标识码:** A **DOI:** 10.19907/j.0490-6756.2022.022004

## Named entity recognition method based on multi-granularity cognition

LI Pan-Feng, CHEN Ying-Jue, ZHONG Ling-Yun, LIN Feng

(College of Computer Science, Sichuan University, Chengdu 610065, China)

**Abstract:** In the field of data scarcity, the performance of named entity recognition is limited by the expression of under-fitting word features. The named entity recognition effect can be improved by introducing conventional multi-task learning methods, but additional labeling costs are required. Aiming at addressing this problem, we propose a new named entity recognition method based on multi-granularity cognition, which can enhance the character feature information and improve the performance of named entity recognition without incurring additional tagging costs. In order to optimize the expression of word embedding, in this approach, we start from the multi-granularity cognition theory and use BiLSTM and CRF as the basic model, the task of named entity recognition under word granularity is combined with the task of entity number prediction under sentence global granularity. Multiple experiments on three different types of data sets show that the method of introducing multi-granularity cognition method can effectively improve the performance of named entity recognition.

**Keywords:** Named entity recognition; Multi-granularity cognitive; Multi-task learning; NLP

## 1 引 言

命名实体识别是自然语言处理领域的核心任务,其目的是从文本序列数据中识别出表示命名实体的成分,并给出其具体的实体类别. 命名实体识

别是将非结构化文本转化为结构化信息的关键技术,可以为上层的自然语言处理任务提供基础支撑,如知识图谱、搜索引擎、推荐系统等<sup>[1]</sup>.

目前,命名实体识别通常被当作序列标注任务<sup>[2]</sup>,常用的方法有基于词典和规则的模式匹配方

收稿日期: 2021-06-19

基金项目: 国家重点研发计划 (2018YFC0832300, 2018YFC0832303)

作者简介: 李攀锋(1996—),男,四川宜宾人,硕士研究生,研究方向为自然语言处理. E-mail: lipanfeng@stu.scu.edu.cn

通讯作者: 林锋. E-mail: linfeng@scu.edu.cn

法、基于统计机器学习的方法、基于深度学习的方法。基于模式匹配方法如 University of Sheffield NLP 在 2012 年开发的 GATE 框架,有着明确清晰的 NER 规范<sup>[3]</sup>。模式匹配方法准确率高,但需要领域专家设计众多的实体抽取规则,且领域词典需要定期维护。基于统计机器学习的方法如隐马尔科夫模型、最大熵模型、条件随机场等。统计机器学习的方法受限于高质量的大规模标注语料库以及对普适性的特征模板的需要。深度学习的方法如 BiLSTM、BiLSTM+CRF 等<sup>[4-7]</sup>。深度学习的方法引入了基于字向量或词向量的特征表示<sup>[8]</sup>,无需进行额外的特征工程,在中文命名实体识别领域越发占据主导地位。

实际的场景中,某些特定领域由于样本匮乏或者受限于标注成本,可使用的训练数据较少,基于常规深度学习的方法对字词特征向量的学习效果欠佳。为此,出现了很多基于多任务学习的命名实体识别研究,如联合分词、词性标注、语义角色标注等任务<sup>[9]</sup>。这些方法通过与其他相关任务的联合学习,对字词特征向量的捕捉效果有一定的提升。但需要额外进行辅助任务数据标注,在标注成本有限的场景下并不适用。

针对这个问题,本文从多粒度认知理论出发,以深度学习中经典的 BiLSTM 及机器学习中传统的 CRF 为基础模型,引入了一种基于多粒度认知(MGC)的命名实体识别方法,在不产生额外标注成本的前提下,增强字嵌入特征表示,提高命名实体识别效果。

## 2 BiLSTM 与 CRF 简介

### 2.1 BiLSTM 模型

RNN(Recurrent Neural Networks) 模型由美国认知科学家 Elman 对 Jordan Network 进行了简化而提出<sup>[10]</sup>,被广泛应用于自然语言处理领域,如词性标注、语义角色标注等序列问题<sup>[11]</sup>。RNN 网络在传统神经网络基础上增加了隐层间的信息传递机制,使得序列的历史信息可以保留并为后续时间步提供支持。常规的 RNN 网络结构如图 1 所示。

RNN 网络在设计之初是为了学习长期的序列依赖性,但是实践证明,在处理长序列数据时,RNN 容易产生梯度消失问题。为了解决这个问题,Hochreiter 等提出了长短期记忆网络(LSTM),

用于改进传统的循环神经网络模型<sup>[12]</sup>。LSTM 单元由输入门、遗忘门和输出门 3 个模块组成,如图 2 所示。

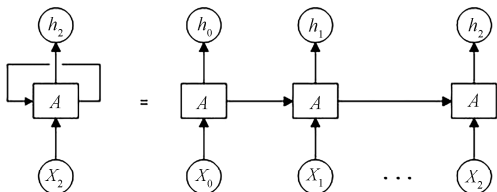


图 1 RNN 结构图

Fig. 1 Structure diagram of RNN

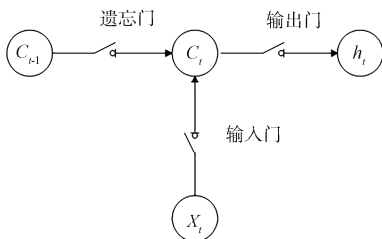


图 2 LSTM 单元结构图

Fig. 2 Structure diagram of LSTM unit

图 2 中,遗忘门决定了上一时刻的单元状态  $C_{t-1}$  有多少保留到当前时刻单元状态  $C_t$ ;输入门决定了当前时刻的输入  $x_t$  有多少保留到了当前时刻细胞状态  $C_t$ ;输出门控制单元状态  $C_t$  有多少传递到当前时刻输出  $h_t$ 。LSTM 单元更新过程如下。

(1) 计算遗忘门值  $f_t$ 。

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

(2) 计算输入门值  $i_t$ 。

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

(3) 计算输出门值  $O_t$ 。

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (3)$$

(4) 确定当前时刻更新信息  $\tilde{C}_t$ 。

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

(5) 更新当前细胞状态  $C_t$ 。

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (5)$$

(6) 计算当前隐层输出  $h_t$ 。

$$h_t = O_t * \tanh(C_t) \quad (6)$$

其中,  $h_{t-1}$  为上一时刻的输出;  $C_{t-1}$  为上一时刻的细胞状态;  $x_t$  为当前时刻的输入;  $\sigma, \tanh$  为激活函数;  $W_f, W_i, W_o, W_c$  为权重矩阵<sup>[13]</sup>。

针对序列处理问题,基于 LSTM 单元的网络模型如图 3 所示。

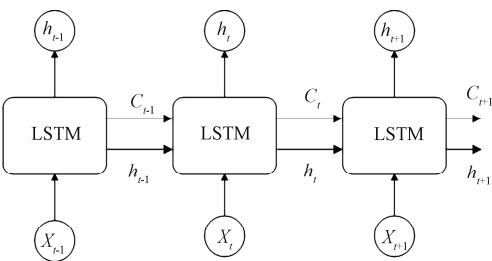


图 3 LSTM 网络结构图  
Fig. 3 Structure diagram of LSTM networks

在自然语言处理领域,单向 LSTM 能够捕捉和记忆前驱语义信息,但无法获取后向语义信息。因此,在目前的自然语言处理相关任务上,通常使用双向的 LSTM 网络<sup>[14]</sup>,即 BiLSTM,网络结构如图 4 所示。

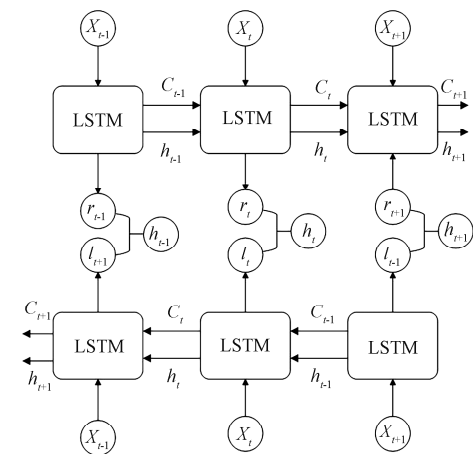


图 4 BiLSTM 网络结构图  
Fig. 4 Structure diagram of BiLSTM networks

## 2.2 CRF 模型

条件随机场模型(CRF)由 Lafferty 等于 2001 年提出,是一种判别式概率无向图模型,是一个标准的序列标注算法<sup>[15]</sup>。最常用的是线性链条件随机场,它接收一个观测序列  $x$ ,输出目标序列  $y$ ,如图 5 所示。

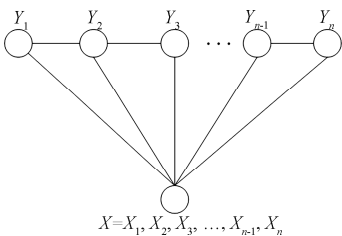


图 5 线性链条件随机场示意图  
Fig. 5 Diagram of linear-CRF

设  $P(Y|X)$  是线性链条件随机场,  $X$  为观测序列变量,  $Y$  为目标序列变量,用  $P(y|x)$  表示给定  $X$

的观测值  $x$  的情况下,随机变量  $Y$  取值为  $y$  的条件概率。

对于观测序列  $x$  的任意位置  $x_i(i=1,2,3,\dots,n)$ ,都有  $K_1$  个局部特征函数和  $K_2$  个节点特征函数,共有  $K=K_1+K_2$  个特征函数,特征函数取值  $\{0,1\}$ ,即满足特征和不满足特征。其中,局部特征函数是关于  $y_{i-1}, y_i, x, i$  的函数,用  $t_k$  表示;节点特征函数是关于  $y_i, x, i$  的函数,用  $s_r$  表示。其中,  $y_{i-1}$  表示当前节点的前驱节点的目标值;  $y_i$  表示当前节点的目标值<sup>[16]</sup>。用  $h_k$  统一表示特征函数。

$$h_k(y_{i-1}, y_i, x, i) = \begin{cases} t_k(y_{i-1}, y_i, x, i), & k \in A \\ s_r(y_i, x, i), & k \in B \end{cases} \quad (7)$$

其中,  $A = \{1, 2, \dots, K_1\}$ ;  $B = \{K_1 + r\}$ ;  $r = 1, 2, \dots, K_2$ 。

则对于整个序列,定义全局特征函数如下。

$$f_k(y, x) = \sum_{i=1}^n h_k(y_{i-1}, y_i, x, i) \quad (8)$$

同时定义特征函数权重系数  $w_k$  如下式。

$$w_k = \begin{cases} \lambda_k, & k = 1, 2, \dots, K_1 \\ \mu_r, & k = K_1 + r, r = 1, 2, \dots, K_2 \end{cases} \quad (9)$$

则有

$$P(y | x) = \frac{1}{Z(x)} \exp \sum_{k=1}^K w_k f_k(y, x) \quad (10)$$

其中,  $Z(x)$  为规范化因子,如下式。

$$Z(x) = \sum_y \exp \sum_{k=1}^K w_k f_k(y, x) \quad (11)$$

将特征函数  $f_k$  与权重系数  $w_k$  张量化表示为

$$w = (w_1, w_2, \dots, w_K)^T \quad (12)$$

$$F(y, x) = (f_1(y, x), \dots, f_K(y, x))^T \quad (13)$$

则  $P(Y|X)$  的内积形式为

$$P(y | x) = \frac{\exp(w \cdot F(y, x))}{\sum_y \exp(w \cdot F(y, x))} \quad (14)$$

## 3 基于多粒度认知的 MGC\_NER 模型

### 3.1 多粒度认知

认知科学是揭示人类智能和行为的一门学科,重点研究神经系统和脑机制中,信息的表达、处理和转化过程<sup>[17]</sup>。随着人工智能的发展,认知计算逐渐出现。认知计算是对新一代智能系统特点的概括,基于认知计算的智能系统具备人类的某些认知能力,能够出色完成对数据的发现、理解、推理和决策等特定认知任务,帮助决策者从不同类型的数

据中揭示潜在意义,以实现不同程度的感知、记忆、学习和其他认知活动。

人的认知往往是多层次、多视角的,依托人类处理问题时所表现出全局观和近似求解能力,多粒度认知成为了解决人工智能问题的重要手段<sup>[18]</sup>。在语义理解任务中,基于篇章粒度与句子粒度实现了语义融合;在分词任务中,基于长短词层次粒度增强了分词效果。

在命名实体识别任务中,本质是对字粒度的文字序列进行标签标注,考虑到人类在对文本认知时,综合考虑文本整体,对文本有一定的全局认知。基于此,本文提出了一种结合字粒度与句子全局粒度的命名实体识别方法。在字粒度层面,模型解决命名实体识别任务,在句子全局粒度层面,模型解决命名实体数量预测任务。两个粒度认知任务联合优化,以提高命名实体识别效果。

### 3.2 MGC\_NER 模型设计

本文提出的基于多粒度认知的命名实体识别方法,以 BiLSTM+CRF 为基本的命名实体识别模型,实现字粒度的序列标记任务,同时采用 BiLSTM 结合注意力机制,实现句子全局粒度的命名实体数量预测任务,网络结构如图 6 所示。

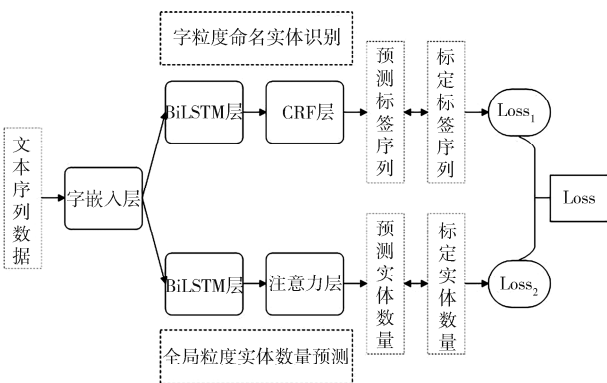


图 6 MGC\_NER 网络结构图

Fig. 6 Structure diagram of MGC\_NER networks

下面对图 6 中模型的各个部分进行说明。

(1) 字嵌入层。模型的输入为中文句子,通过预处理构建字表,将句子转化为高维的 one-hot 矩阵,传入字嵌入层。字嵌入层将 one-hot 矩阵映射为压缩的低维的字嵌入矩阵。

(2) 命名实体识别 BiLSTM 层。接收字嵌入后的文本矩阵,通过捕捉前后向语义信息,初步获取命名实体识别任务的标签预测概率。

(3) CRF 层。以命名实体识别 BiLSTM 层的输出为发射矩阵,通过状态转移矩阵进一步修正命

名实体识别任务的标签预测概率。

(4) 实体数量预测 BiLSTM 层。接收字嵌入后的文本矩阵,通过捕捉前后向语义信息,进一步获取句子的语义特征。

(5) 注意力层。接收实体数量预测 BiLSTM 层的输出特征矩阵,通过注意力机制获取句子全局的重要特征信息,最终预测出该句子中包含的实体数量。

总体上,字粒度的命名实体识别任务与句子全局粒度的实体数量预测任务共享字嵌入层,通过两个粒度认知任务的损失函数共同优化。

### 3.3 MGC\_NER 模型优化目标

本文提出的 MGC\_NER 模型有两部分优化目标。

(1) 命名实体识别的优化 Loss 采取 log 似然如下式。

$$Loss_1 = -\log \frac{P_{\text{realpath}}}{P_1 + P_2 + \dots + P_n} \quad (15)$$

其中,  $P_{\text{realpath}}$  为在给定观测序列  $x$  的条件下,真实的目标序列  $y_{\text{realpath}}$  的概率;  $P_1 \sim P_n$  为在给定观测序列  $x$  的条件下,预测目标序列  $y_n$  的概率。

(2) 实体数量预测的优化 Loss 采取交叉熵损失如下式。

$$Loss_2 = -\sum_{i=1}^n y^{(i)} \log \hat{y}^{(i)} \quad (16)$$

其中,  $y$  为实体数量标定类别张量;  $\hat{y}$  为实体数量预测概率张量。

为实现两个粒度联合优化,通过系数  $d$  (超参数,通过验证集调整) 将  $Loss_1$  与  $Loss_2$  联合,得到最终模型优化目标。

$$Loss = Loss_1 + d * Loss_2 \quad (17)$$

## 4 实验与结果

### 4.1 实验数据

本文选取新浪财经简历数据集和人民日报数据集两个公共数据集,自行构建了涉案财物领域法律数据集,并对数据集规模进行了限制。其中,新浪财经简历数据集与涉案财物领域法律数据集为特定领域数据集,人民日报数据集为跨领域数据集。通过多领域数据集以验证本文方法的有效性。

本文的命名实体识别任务采取 BIO 三位序列标注法,各数据集规模及标签类别划分如表 1 所示。

表 1 数据集信息表		
Tab. 1 Table of data sets information		
数据集	规模/字	标签类别
新浪财经	271 413	28
人民日报	247 121	9
涉案财物	285 009	13

本文按照 5 : 1 : 1 的比例随机划分了训练集、验证集和测试集. 在各个数据集的训练集上实体数量分布情况如图 7~图 9 所示.

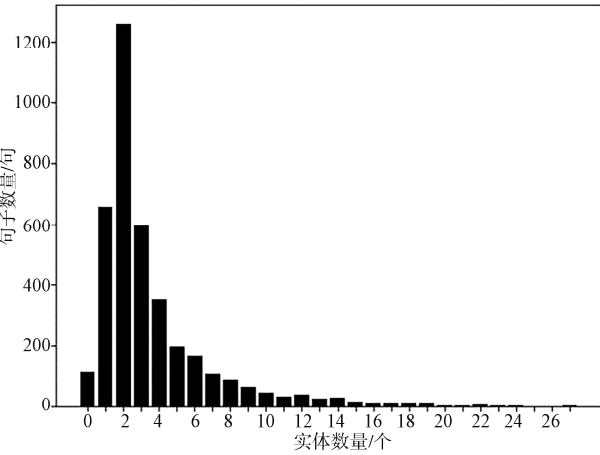


图 7 新浪财经数据集实体数量分布图  
Fig. 7 Distribution map of the number of entities in the Sina Finance data set

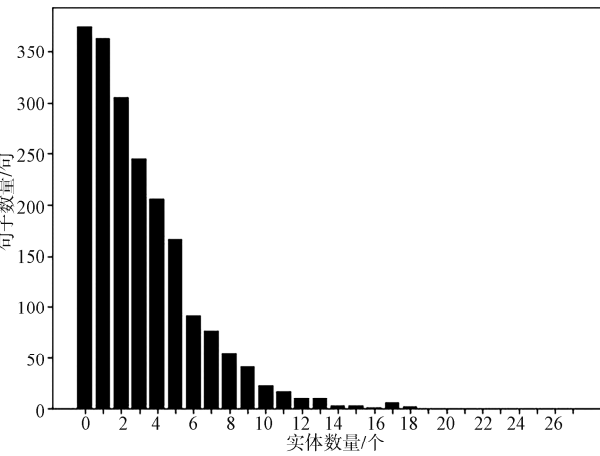


图 8 人民日报数据集实体数量分布图  
Fig. 8 Distribution map of the number of entities in the People's Daily data set

由图 7~图 9 可知,在新浪财经简历数据集、人民日报数据集和涉案财物处置领域法律数据集上,各种实体数量类型下的句子分布差异较大,在实体数量预测任务中将产生数据偏斜问题. 针对这个问题,本文通过区间划分,将实体数量区间预测作为实体数量预测的目标,避免了类别不平衡问题.

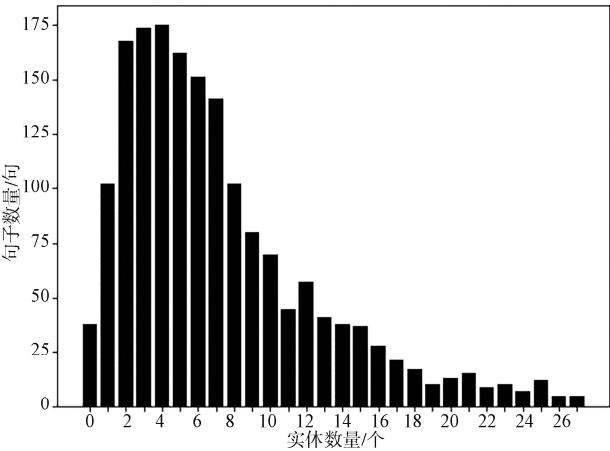


图 9 涉案财物数据集实体数量分布图  
Fig. 9 Distribution map of the number of entities in the Property Involved data set

4.2 实验设计

本文在 BiLSTM+CRF 命名实体识别模型上引入基于多粒度认知的联合优化方法(MGC\_NER 模型). 为进一步说明本文提出的多粒度认知机制的有效性,在单纯的 BiLSTM 模型上引入基于多粒度认知的联合优化方法构建了对比实验. 并在随机初始化字向量和载入预训练字向量两种情况下进行实验,总体设计如图 10 所示.

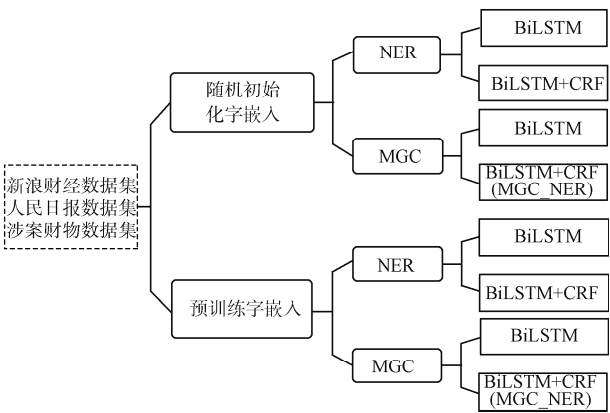


图 10 实验总体设计图  
Fig. 10 The overall design of the experiment

其中,NER 组模型为常规的命名实体识别模型,引入了 MGC 机制的 BiLSTM+CRF 模型(MGC\_NER)为本文提出的最终模型. 特别的,引入了 MGC 机制的 BiLSTM 模型用于证明 MGC 机制的有效性. 各组实验均在相同情况下进行,字嵌入维度 300 维, BiLSTM 隐层维度 256 维,注意力隐层维度 100 维,实体数量预测区间数为 5,粒度联合系数  $d$  为 0.3.

4.3 实验结果

本文依照图 10 的实验结构进行了实验,评价指标采取在实体识别严格比对下的准确率(Precision)、召回率(Recall)和  $F_1$  值. 计算公式如下.

$$\text{Precision}=\frac{P_{\text{correct}}}{P_{\text{all}}}$$

(18)

$$\text{Recall}=\frac{P_{\text{correct}}}{C_{\text{all}}}$$

(19)

$$F_1=\frac{2 * \text{Precision} * \text{Recall}}{\text{Precision}+\text{Recall}}$$

(20)

其中, $P_{\text{correct}}$ 表示预测正确的实体数量; $P_{\text{all}}$ 表示预测结果中实体数量的总和; $C_{\text{all}}$ 表示标定数据中实体数量总和. 实验结果如表 2~表 4 所示.

其中,粗体数据为引入了本文提出的多粒度机制方法的实验数据. 对比各组实验结果可知,本文提出的方法在保持基础模型的准确率不降低或基本不降低的情况下,能够显著提升召回率. 总体上,在各个数据集的各项实验上,本文所提方法的  $F_1$  值都高于对比方法的  $F_1$  值,验证了本文方法的有效性.

表 2 新浪财经数据集实验结果

Tab. 2 Experimental results of the Sina Finance data set

字嵌入形式	模型	准确率	召回率	$F_1$
随机初始化	BiLSTM	0. 906 4	0. 897 5	0. 901 9
	<b>MGC_BiLSTM</b>	<b>0. 914 5</b>	<b>0. 912 8</b>	<b>0. 913 7</b>
	BiLSTM+CRF	0. 929 1	0. 925 7	0. 927 4
	<b>MGC_BiLSTM+CRF</b>	<b>0. 935 4</b>	<b>0. 933 1</b>	<b>0. 934 2</b>
预训练初始化	BiLSTM	0. 913 1	0. 909 2	0. 911 1
	<b>MGC_BiLSTM</b>	<b>0. 920 2</b>	<b>0. 920 2</b>	<b>0. 920 2</b>
	BiLSTM+CRF	0. 927 4	0. 917 7	0. 922 6
	<b>MGC_BiLSTM+CRF</b>	<b>0. 929 8</b>	<b>0. 926 3</b>	<b>0. 928 0</b>

表 3 人民日报数据集实验结果

Tab. 3 Experimental results of the People's Daily data set

字嵌入形式	模型	准确率	召回率	$F_1$
随机初始化	BiLSTM	0. 821 7	0. 505 5	0. 625 9
	<b>MGC_BiLSTM</b>	<b>0. 808 5</b>	<b>0. 691 5</b>	<b>0. 745 5</b>
	BiLSTM+CRF	0. 849 3	0. 645 4	0. 733 5
	<b>MGC_BiLSTM+CRF</b>	<b>0. 840 6</b>	<b>0. 704 2</b>	<b>0. 766 4</b>
预训练初始化	BiLSTM	0. 769 0	0. 593 0	0. 669 6
	<b>MGC_BiLSTM</b>	<b>0. 774 6</b>	<b>0. 748 8</b>	<b>0. 761 6</b>
	BiLSTM+CRF	0. 806 8	0. 677 2	0. 736 3
	<b>MGC_BiLSTM+CRF</b>	<b>0. 770 3</b>	<b>0. 767 8</b>	<b>0. 769 1</b>

表 4 涉案财物数据集实验结果

Tab. 4 Experimental results of the Property Involved data set

字嵌入形式	模型	准确率	召回率	$F_1$
随机初始化	BiLSTM	0. 905 5	0. 845 7	0. 874 6
	<b>MGC_BiLSTM</b>	<b>0. 909 2</b>	<b>0. 877 2</b>	<b>0. 892 9</b>
	BiLSTM+CRF	0. 904 5	0. 877 2	0. 890 6
	<b>MGC_BiLSTM+CRF</b>	<b>0. 905 8</b>	<b>0. 902 3</b>	<b>0. 904 1</b>
预训练初始化	BiLSTM	0. 895 0	0. 848 2	0. 870 9
	<b>MGC_BiLSTM</b>	<b>0. 901 1</b>	<b>0. 872 1</b>	<b>0. 886 4</b>
	BiLSTM+CRF	0. 893 0	0. 872 7	0. 882 8
	<b>MGC_BiLSTM+CRF</b>	<b>0. 912 2</b>	<b>0. 889 7</b>	<b>0. 900 8</b>

此外,实验数据还表明了如下 3 种情况.

(1) 通过对比随机初始化字向量与预训练字向量的实验结果,说明在命名实体识别领域,预训练字向量不一定能够起到积极的效果,甚至可能起到负面作用. 本文认为,预训练的字向量更侧重于语义信息,在命名实体识别任务上,信息量冗杂的预训练向量干扰了任务的进行.

(2) 通过对比 BiLSTM 和 BiLSTM+CRF 实验结果,说明 CRF 对于命名实体识别任务有一定的辅助作用. 本文认为,由于 CRF 模型在训练过程中学习到了相邻字标签之间的约束关系,从而降低了命名实体识别的错误率.

(3) 在人民日报数据集上,本文的方法对比基础模型准确率有所下降,本文认为,与人民日报数据集自身特点有关. 从图 8 可知,该数据集上实体数量为零的句子较多,为了维持区间均衡,将实体数量为零的句子与其他句子划分为一个区间,导致了实体数量预测任务中对数量为零的句子造成了干扰,从而间接影响了该数据集上命名实体识别的准确率.

5 结 论

本文提出了一种基于多粒度认知的命名实体识别方法,从多粒度认知理论出发,基于 BiLSTM 和 CRF 构建了命名实体识别与实体数量预测联合模型,通过共同优化字嵌入表达,提高句子语义信息捕捉能力. 本文在 3 个数据集上进行了多组实验,实验表明,引入多粒度认知机制有助于改善命名实体识别效果,证明了本文方法的有效性. 后续将对实体数量预测中实体数量区间的划分进行深入研究,细化区间划分,有望进一步提升模型效果.

参考文献:

[1] 刘浏,王东波. 命名实体识别研究综述[J]. 情报学报, 2018, 37: 329.

[2] 张晗,郭渊博,李涛. 结合 GAN 与 BiLSTM-Attention-CRF 的领域命名实体识别[J]. 计算机研究与发展, 2019, 56: 1851.

[3] 焦凯楠,李欣,朱容辰. 中文领域命名实体识别综述[J]. 计算机工程与应用, 2021, 57: 1.

[4] 郑余祥,左祥麟,左万利,等. 基于时间关系的 Bi-LSTM+GCN 因果关系抽取[J]. 吉林大学学报:理学版, 2021, 59: 6.

[5] 李静,程芄森,许丽丹,等. 基于局部对抗训练的命名实体识别方法研究[J]. 四川大学学报:自然科学版, 2021, 58: 023003.

[6] 张若彬,刘嘉勇,何祥. 基于 BLSTM-CRF 模型的安全漏洞领域命名实体识别[J]. 四川大学学报:自然科学版, 2019, 56: 469.

[7] 彭嘉毅,方勇,黄诚,等. 基于深度主动学习的信息安全领域命名实体识别研究[J]. 四川大学学报:自然科学版, 2019, 56: 457.

[8] 尹学振,赵慧,赵俊保,等. 多神经网络协作的军事领域命名实体识别[J]. 清华大学学报:自然科学版, 2020, 60: 648.

[9] 罗凌,杨志豪,宋雅文,等. 基于笔画 ELMo 和多任务学习的中文电子病历命名实体识别研究[J]. 计算机学报, 2020, 43: 1943.

[10] 杨丽,吴雨茜,王俊丽,等. 循环神经网络研究综述[J]. 计算机应用, 2018, 38: 1.

[11] 司念文,王衡军,李伟,等. 基于注意力长短时记忆网络的中文词性标注模型[J]. 计算机科学, 2018, 45: 66.

[12] Lindemann B, Maschler B, Sahlab N, *et al.* A survey on anomaly detection for technical systems using LSTM networks [J]. Comput Ind, 2021, 131: 103498.

[13] 权波,杨博辰,胡可奇,等. 基于 LSTM 的船舶航迹预测模型[J]. 计算机科学, 2018, 45: 126.

[14] 翟社平,段宏宇,李兆兆. 基于 BiLSTM-CRF 的知识图谱实体抽取方法[J]. 计算机应用与软件, 2019, 36: 269.

[15] 宋青松,张超,陈禹,等. 组合全卷积神经网络和条件随机场的道路分割[J]. 清华大学学报:自然科学版, 2018, 58: 725.

[16] 王宇,杨艺,王宝山,等. 深度神经网络条件随机场高分辨率遥感图像建筑物分割[J]. 遥感学报, 2019, 23: 1194.

[17] 陈伟宏,安吉尧,李仁发,等. 深度学习认知计算综述[J]. 自动化学报, 2017, 43: 1886.

[18] 王国胤,李帅,杨洁. 知识与数据双向驱动的多粒度认知计算[J]. 西北大学学报:自然科学版, 2018, 48: 488.

引用本文格式:  
中 文: 李攀峰,陈樱珏,钟泠韵,等. 基于多粒度认知的命名实体识别方法[J]. 四川大学学报:自然科学版, 2022, 59: 022004.  
英 文: Li P F, Chen Y J, Zhong L Y, *et al.* Named entity recognition method based on multi-granularity cognition [J]. J Sichuan Univ: Nat Sci Ed, 2022, 59: 022004.