

基于分层序列标注的实体关系联合抽取方法

田佳来¹ 吕学强¹ 游新冬^{1,†} 肖刚² 韩君妹²

1. 北京信息科技大学, 网络文化与数字传播北京市重点实验室, 北京 100101; 2. 复杂系统仿真总体重点实验室, 军事科学院系统工程研究院, 北京 100101; † 通信作者, E-mail: youxindong@bistu.edu.cn

摘要 为了提高实体关系联合抽取的效果, 提出一种端到端的联合抽取模型(HSL)。HSL模型采取一种新的标记方案, 将实体和关系的联合抽取转化成序列标注问题, 同时采用分层的序列标注方式来解决三元组重叠问题。实验证明, HSL模型能有效地解决三元组重叠问题, 在军事语料数据集上F1值达到80.84%, 在公开的WebNLG数据集上F1值达到86.4%, 均超过目前主流的三元组抽取模型, 提升了三元组抽取的效果。

关键词 实体关系联合抽取; 三元组重叠; 序列标注; 知识图谱; HSL

Joint Extraction of Entities and Relations Based on Hierarchical Sequence Labeling

TIAN Jialai¹, LÜ Xueqiang¹, YOU Xindong^{1,†}, XIAO Gang², HAN Junmei²

1. Beijing Information Science and Technology University, Beijing Key Laboratory of Internet Culture and Digital Dissemination Research, Beijing 100101; 2. National Key Laboratory for Complex Systems Simulation, Institute of Systems Engineering, Beijing 100101; † Corresponding author, E-mail: youxindong@bistu.edu.cn

Abstract In order to further improve the effect of entity relationship joint extraction, this paper proposes an end-to-end joint extraction model (HSL). HSL model adopts a new labeling scheme to transform the joint extraction of entities and relationships into sequence labeling problems, and uses a layered sequence labeling method to solve the problem of triple overlap. The experiments demonstrates that HSL model can effectively deal with the problem of triple overlap and improve the extraction effect. The F1 value on the military corpus data set reaches 80.84%, and 86.4% on the WebNLG open data set, which exceeds the current mainstream triple extraction model, improving the effect of triple extraction.

Key words entity relationship joint extraction; triple overlap; sequence annotation; knowledge graph; HSL

实体关系三元组抽取是知识图谱构建过程中不可缺少的步骤, 实体关系三元组抽取指从文本中抽取实体对, 并确定实体之间的关系。实体关系三元组结构为(主语, 关系, 宾语), 例如(鲁迅, 属于, 中国)。

早期的三元组抽取一般采用流水线方法: 各个任务依次处理。例如Nadeau等^[1]和Zelenko等^[2]都是首先识别文本中的实体, 然后抽取文本中实体之间的关系。如果原文本句子中的实体相互之间没有关系, 也会强制给任意两实体之间附加一种关

系。但是, 流水线方法忽略两个任务之间的联系^[3], 所以有学者提出实体关系联合抽取模型: 用单一模型抽取文本中的实体关系三元组, 能够增强实体抽取任务与关系抽取任务的联系。Ren等^[4]、Li等^[5]和Miwa等^[6]采用实体关系联合抽取模型抽取文本中的实体关系三元组, 但其模型都需要人工构造大量的特征, 同时也依赖相关的NLP工具包自动抽取。由于NLP工具包抽取结果不稳定, 可能导致错误, 影响抽取的效果。随着神经网络的发展, 其在NLP任务中也取得成效。王国昱^[7]将深度学习方法

国家自然科学基金(61671070)、国家语委重点项目(ZD1135-53)、国防科技重点实验室基金(6142006190301)、北京信息科技大学促进高校内涵发展科研水平提高项目(2019KYNH226)和北京信息科技大学“勤信人才”培育计划(QXTCPB201908)资助

收稿日期: 2020-06-11; 修回日期: 2020-08-14

应用在命名实体识别任务中,取得不错的F1值。Zeng等^[8]和Xu等^[9]使用深度学习的方法来解决关系分类问题,得到的F1值比以往的机器学习模型提升明显。Zheng等^[10]采用基于LSTM的神经网络实体关系联合抽取模型,选取序列标注抽取实体关系三元组,其模型按照就近原则的关系链接方式,然而该模型忽略了实体对完全重叠(EPO)和单一实体重叠(SPO)的情况。Zeng等^[11]首先提出解决三元组重叠问题,并基于Seq2seq思想,提出实体关系联合抽取模型,能够解决三元组重叠问题,但模型依赖解码的结果,导致实体识别不完全。Fu等^[12]采用基于图卷积神经网络改进的方法,进行实体关系三元组抽取,效果比Zeng等^[11]的模型有所提高。

还有许多学者专注于三元组的抽取研究。李明耀等^[13]对中文实体关系三元组抽取进行研究,根据依存句法分析和中文语法制定抽取规则,F1值达到76.78%。黄培馨等^[14]采用一种融合对抗学习的方法,利用带有偏置的激活函数来增强信息的多通道传输特性,取得不错的效果。赵哲煊^[15]对生物学实体关系三元组进行抽取,首先通过多标签卷积神经网络对实体进行抽取,最后用领域词典查询的方法抽出实体关系三元组。张永真等^[16]针对专利文本三元组抽取,通过机器学习模型,分析词性特征、位置特征和上下文特征的重要性,剔除弱的特征,提升了专利文本三元组抽取的效果。王昊^[17]构建知识库来协助实体关系抽取任务,当目标实体在知识库三元组中出现的次数大于某个阈值时,将其关系定义为关系高频词,同时采用Word2vec语言模型训练嵌入词向量,用于增强模型语义信息,并通过定义关系高频词和增加先验特征来提高模型效果。

尽管目前主流的三元组抽取模型可以在一定程度上解决三元组重叠问题,但是由于模型结构的原因导致编码能力弱,抽取效果差,F1值低于50%。为了提高具有三元组重叠中实例的三元组抽取效果,本文提出一种端到端的联合抽取模型(HSL),HSL模型采取一种新的标记方案,将实体与关系的联合提取转化成序列标注问题,同时采用分层的序列标注方式来解决三元组重叠问题。实验中采用人工标记的军事语料和WebNLG公开数据集。结果证明,无论在特定领域的语料上还是在公开语料上,HSL模型的准确率和召回率都比目前主流的三元组抽取模型有所提升,能够更有效地抽取三元组。

1 实体关系三元组抽取模型-HSL

三元组抽取目标为抽取句子中的 (s, p, o) ,其中, s 为主语, o 为宾语, p 为 s 与 o 的关系。本文的三元组抽取模型设计思路来源于百度三元组抽取比赛中的一个Baseline^[18],参考Seq2Seq模型^[19]的思路,先抽取主语,然后根据主语的先验信息抽取关系及宾语。Seq2Seq模型的解码公式如下:

$$P(y_1, y_2, \dots, y_n | x) = P(y_1 | x)P(y_2 | x, y_1) \dots P(y_n | x, y_1, y_{n-1}), \quad (1)$$

其中, x 为先验句子。给定一个 x ,在所有的 y 上面建模,使生成 y_1, y_2, \dots, y_n 的概率最大似然。首先输入 x 得到第一个 y_1 词语,再将 x 和 y_1 作为先验特征输入模型中,解码出 y_2 ,依此类推,解码出 y_3, \dots, y_n 。由此,可以得出三元组抽取公式:

$$P(s, p, o | x) = P(s | x)P(p, o | x, s), \quad (2)$$

其中, x 是输入的句子。首先抽取主语 s ,然后将主语 s 和句子 x 解码出关系 p 和宾语 o 。

从式(2)来看,一个句子中只能抽取出一组三元组,所有我们将三元组抽取问题转化成序列标注问题。序列标注时,对一个字符序列中的每一个字符打上相应的标签(图1),根据标签抽取命名实体。将主语 s 的识别过程转化成序列标注问题,一个句子可以抽取 n 个主语 s ,再分别将各个主语 s 作为先验特征进行关系 p 和客体 o 的抽取。另外,一个句子只能识别出一个客体和一种关系,所以根据主语 s 抽取的思想,同样通过对句子进行序列标注来抽取客体和关系。但是,一个句子只能生成一条序列标签,无法确定主语与宾语的关系。因此,在序列标注时,将标签设计成带有关系类别的标签,以便确定主语和宾语的关系。这种方法不能解决三元组抽取的实体对重叠问题(EPO),即两个三元组主语与宾语完全重叠而仅仅关系不同的情况。

为了解决实体对重叠(EPO)问题,本文采用分层的序列标注方法,在抽取宾语 o 和关系 p 时,每层标注序列产生的宾语 o 都与主语 s 对应一种提前设定好的关系,最终生成的标注序列数量与关系的数量相同。

1.1 标签设计

本文中三元组抽取顺序是先抽取主语,再根据主语的先验特征来预测客体和关系。HSL模型先标

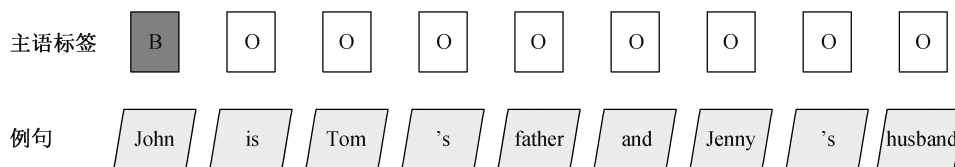


图 1 主语标签设计

Fig. 1 Subject label design

注句子序列,生成主语的标注序列,再生成宾语的标注序列。

图 1 是抽取主语时,句子经过序列标注后生成对应标签的一个例子。根据标签,可以容易地提取出主语。采用 BIO 的标注方案, B 代表当前字符是主语的首个字符, I 表示当前字符是主语的中间或结尾部分, O 表示当前字符与主语无关。图 1 的句子中, John 对应的字符是 B, 表示其是主语的开头字符, 向后搜寻, 如果下一个字符对应的标签为 I, 说明当前字符也是主体的一部分, 直到下一个字符为 O 标签, 则主语提取完毕。所以, 从图 1 的句子中最终提取出来的主语为 John。

如图 2 所示, 经上一步骤抽取出主语 John 后, 结合主语的先验特征 John, 对句子进行分层序列标注, 生成带有与主语对应关系类别的宾语标签序列, 这种方法称为分层序列标注。HSL 模型基于有监督学习, 所以事先预定的关系类别是固定的, 有几种关系类别就会生成几条宾语标签序列, 从代表某一关系类别的宾语标签序列中提取的宾语就表示该宾语与先验主语之间的关系为该类别。同样采用 BIO 的标注方案, 在图 2 的句子中提取的宾语为 Jenny 和 Tom, Jenny 所在的宾语标签序列对应的关系类别为 wife, Tom 所在的宾语标签序列对应的关系类别为 son。所以, 最终抽取出两个三元组, 一个是

(John, wife, Jenny), 另一个是(John, son, Tom)。如果 Tom 被判断为主语, 同样会重复以上抽出三元组的操作。

1.2 三元组模型结构

在早期序列标注任务中, 通常采用条件随机场和马尔可夫模型。近年来, 序列标注任务得到飞速发展, 随着神经网络的出现, 端到端模型逐渐应用于序列标注任务中。本文的端到端联合抽取模型(HSL)将三元组抽取分成两个序列标注任务, 模型结构如图 3 所示。首先, HSL 采用语言模型和位置编码, 将文本转化成具有语义和距离信息的词向量; 然后, 将词向量经过 12 层的 GLU Dilated CNN 编码, 得到句子编码向量, 再通过 Self Attention 机制, 进一步提取特征, 解码出主语; 最后, 将主语作为先验特征输入 BILSTM 模型中, 与句子编码向量相加和, 通过 Self Attention 机制, 进一步提取特征, 解码出关系和宾语。

1.3 编码器

采用 12 层 GLU Dilated CNN 编码。卷积神经网络最早应用在图像领域中, 能充分地提取图片中的特征。在自然语言处理领域中, Kim^[20]最早提出利用文本卷积进行文本分类任务, 发现卷积能够充分地提取文本特征与挖掘词语之间的关联。Dauphin 等^[21]提出一种新的非线性单元 GLU (gated linear

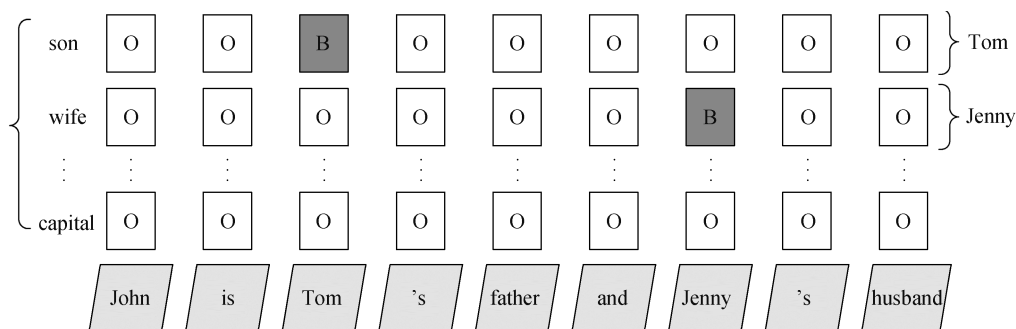


图 2 宾语和关系标签设计

Fig. 2 Label design of object and relationship

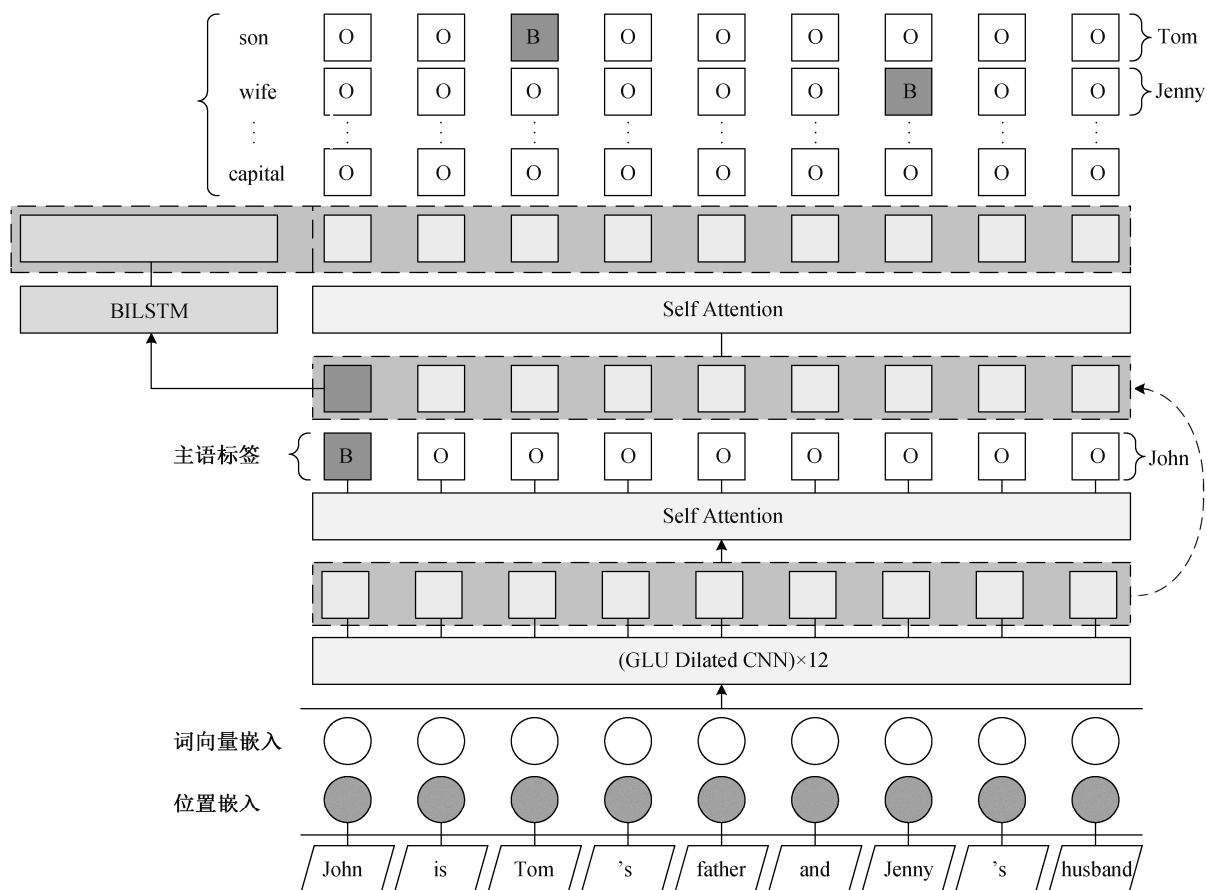


图 3 实体关系三元组抽取模型结构

Fig. 3 Structure diagram of entity relationship triples extraction model

units), 将激活函数转化成另一种表达方式, 可以防止梯度消失现象, Gehring 等^[22]在 Facebook 文章中也引用 GLU 方法。Yu 等^[23]提出 Dilated 卷积方法, 过程如图 4 所示, 当膨胀率为 1 时, 卷积为标准卷积; 当膨胀率为 2 时, 卷积操作会跳过中间词语, 将输入向量 w_1 和 w_3 关联起来, 能够增加远距离词语间的相互关联性。将词嵌入后得到的向量 W_{all} 通过带有 GLU 方法的 Dilated 卷积, 得到向量与通过阀门数值控制的 W_{all} 加和, 最终可以得到编码后的向量 H 。

GLU Dilated CNN 的结构如图 5 所示, 计算方法如式(3)和(4)所示。GLU 方法能够提高模型信息

多通道传输能力, 膨胀卷积方法可以提高模型特征抽取能力。

$$\sigma = \sigma(\text{Dilated Conv1}(W_{all})), \quad (3)$$

$$H = \text{Dilated Conv2}(W_{all}) \otimes \sigma + (1 - \sigma) \otimes W_{all} \quad (4)$$

1.4 解码器

编码完成后, 首先解码主语, 再解码关系及宾语。在两次解码过程中, 都采用 Self Attention 机制^[24]来进一步提取特征。

首先进行主体解码。图 6 展示经过 12 层 GLU Dilated CNN 编码得到的 H 向量通过 Self Attention 机制的具体操作, H 向量经过 3 个不同的全连接层,

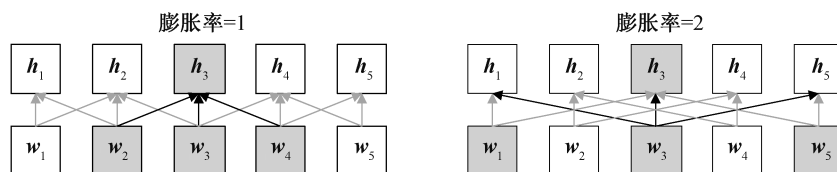


图 4 膨胀卷积^[23]

Fig. 4 Dilated convolution^[23]

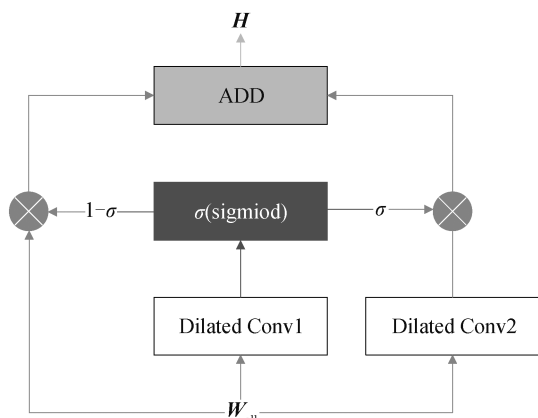


图 5 GLU Dilated CNN 结构
Fig. 5 Structure of GLU Dilated CNN

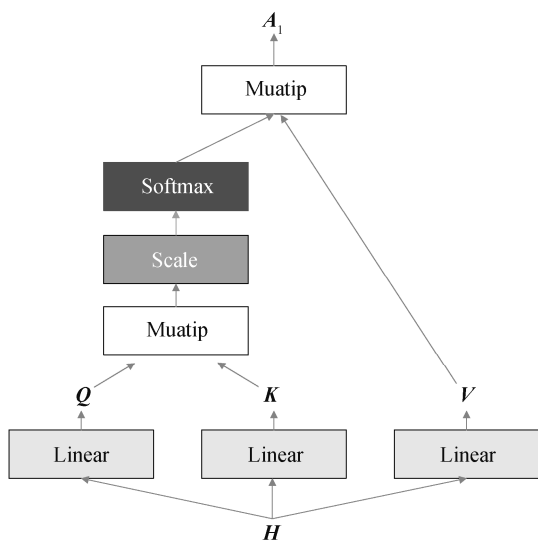


图 6 Self Attention 机制
Fig. 6 Self Attention mechanism

得到 Q , K 和 V 。向量 QK^T 表示词语与其他词语的相关程度，对 QK^T 进行标准化，并输入 Softmax 激活函数，得到词语之间的相关程度向量。将相关程度向量与 V 做点乘，得到向量 A_1 ，Self Attention 机制表达如下：

$$A_1 = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V, \quad (5)$$

其中， d_k 为经验参数，能够使训练时梯度更加稳定。

Linear 层为全连接层，能通过点积的方式，得到输出维度为标签数量的向量，从而获得每个字对应标签的概率。由于主语只需要一条标签序列，所以对向量 A_1 只需做一次 Linear 层的操作。由于标签维度为 3，所以 Linear 层输出维度为 3，Linear 层

的激活函数采用 Softmax， A_1 通过 Linear 层得到每个词语对应的标签概率，每个词语取其对应标签中概率最大的标签为最终标签，生成主语标签序列 (图 3)。主语解码过程如下：

$$\text{Subject tagger} = \text{MaxpTag}(\text{softmax}(\text{linear}(A_1))), \quad (6)$$

其中，MaxpTag() 为对应标签概率最大的函数。

图 7 为宾语和关系的解码结构。在对宾语和关系解码时，需要将式 (6) 得到的主语作为先验特征，加入宾语及关系解码结构中，找到主语在文中的开始和结尾索引位置，从 GLU Dilated CNN 编码后的 H 向量中取出对应的向量 H_{subject} ，使用 BILSTM 模型进一步提取其特征。将 H 向量与 Self Attention 机制得到的结果向量和 BILSTM 模型的结果向量相加，得到最终向量。由于关系数量与宾语标签序列数量相同，假设关系有 n 种，解码时将最终向量解码成 n 种关系类别的宾语标签序列，每个宾语标签序列提取的宾语与主语的关系就是本条宾语标签序列预定义的关系，即需要 n 个 Linear 层和 Softmax 层，最终生成图 2 所示的宾语标签序列。

2 实验

2.1 实验数据

采用两种数据源来验证 HSL 模型的有效性：1) 从环球军事网采集数据并自行标注构建的军事语料数据集；2) WebNLG 数据集。数据集的具体信息如表 1 所示。

环球军事网包含大量武器装备信息，内容偏军事文本描述。自行标注数据 2925 条，其中 2625 条

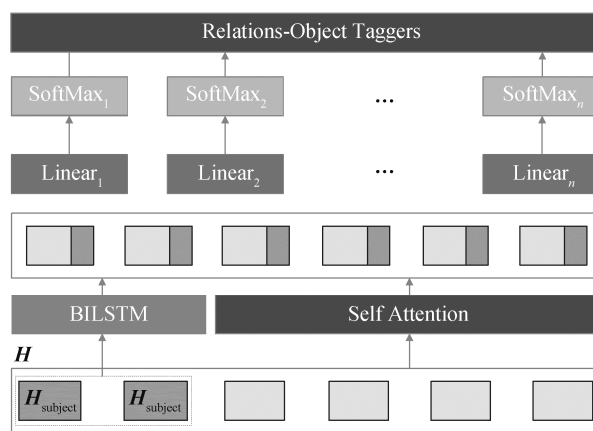


图 7 宾语和关系的解码结构
Fig. 7 Structure figure of object and relationship decoding

表 1 军事语料数据集和 WebNLG 数据集的重叠情况
Table 1 Overlap of military corpus data set and WebNLG data set

数据集	重叠类别	训练集	测试集
军事语料数据集	正常	985	92
	单个实体重叠	1640	207
	所有类别	2625	299
WebNLG 数据集	正常	1596	246
	完全重叠	227	26
	单个实体重叠	3406	457
	所有类别	5019	729

用于训练, 299 条用于测试。

WebNLG 数据集是评估三元组抽取模型效果最常用的通用领域数据集, 其中的每个句子都会包含多个实体关系三元组。公平起见, 本文采用其发布的 WebNLG 部分数据集进行模型评估, 其中 5019 条句子进行训练, 500 条进行验证, 729 条进行测试。根据句子的三元组重叠程度, 将句子分为正常(Normal)、主体客体完全重叠(EPO)和单个实体重叠(SEO)3 个类别。从表 1 可以看出, WebNLG 数据集中大多数句子都属于 SEO 类别, 如果模型在该语料上的 F1 值较好, 则说明模型具备解决单个实体重叠问题的能力。

2.2 评价指标

为了验证本文模型的有效性, 采用与 Gardent 等^[25]相同的准确率(P)和召回率(R)相结合的 F1 值来评判。当预测产生的三元组与真实三元组的名称和类别完全一致时, 称为正确识别的三元组。

$$P = \frac{\text{正确识别三元组总数量}}{\text{识别出的三元组总数量}} \times 100\%, \quad (10)$$

$$R = \frac{\text{正确识别三元组总数量}}{\text{测试数据中三元组总数量}} \times 100\%, \quad (11)$$

$$F1 = \frac{2 \times P \times R}{P + R} \times 100\%。 \quad (12)$$

2.3 实验及结果分析

HSL 模型运行在戴尔服务器的 Ubuntu 16.04 操作系统中, 服务器运行内存为 64 G。GPU 为 8 块 Tesla V100 显卡, 每块显存为 16 G, 编码语言为 python3.6, 采用 Keras 深度学习框架。实验结果表明, 不同的参数对实体关系三元组的识别结果有一定程度的影响。经过调参, 最终确定的最优参数如

表 2 所示。

将 HSL 模型与目前在三元组抽取任务中效果最好的 4 个模型(NovelTagging 模型^[10]、CopyR 模型^[11]、GraphRel 模型^[12]和 Baseline 模型^[18](BL 模型))进行比较。除 BL 模型外, 其余模型在相同 WebNLG 数据集上的得分都是从原始文章复制而来, 并在军事语料数据集上对各个模型进行测试。在 WebNLG 数据集上的实验结果如表 3 所示。

在 WebNLG 数据集上, HSL 的 F1 值比目前最优的 BL 模型高 42.4%, 准确率和召回率均超过其他模型, 证明了 HSL 模型的有效性。同时, HSL 模型的召回率和准确率相差不大, 说明模型比较稳定。另外, WebNLG 数据集属于通用数据集, 实体关系种类大于 100 种, 表明 HSL 在通用数据集上效果好且稳定。在 WebNLG 数据集中, 单个实体重叠(SEO)的句子占比非常大, 由于 NovelTagging 方法假设每个实体标签只能对应一种关系, 忽略了三元组抽

表 2 模型参数设置

Table 2 Model parameter settings

参数名称	参数值
Word2vec 词向量维度	200
位置编码词向量维度	200
自编码词向量维度	200
GLU Dilated CNN 层数	12
GLU Dilated CNN 膨胀率	1, 2, 5, 1, 2, 5, 1, 2, 5, 1, 1, 1
模型训练迭代次数	500
Self Attention 隐藏层维度	768
BILSTM 词向量维度	200
50 轮次前学习率	0.01
50 轮次后学习率	0.0001

表 3 WebNLG 数据集和军事语料数据集的实验结果(%)
Table 3 Experimental results of each model in the WebNLG data set and military corpus data set (%)

数据集	模型	P	R	F1
WebNLG 数据集	NovelTagging	52.5	19.3	28.3
	CopyROneDecoder	32.2	28.9	30.5
	CopyRMutilDecoder	37.7	36.4	37.1
	GraphRel1p	42.3	39.2	40.7
	GraphRel2p	44.7	41.1	42.9
	BL	75.6	75.3	75.4
	HSL	85.4	84.4	86.4
军事语料数据集	BL	96.05	60.50	74.24
	HSL	85.98	76.28	80.84

取的三元组重叠问题, 所以其召回率仅为 19.3%。CopyR 模型和 GraphRel 模型考虑了重叠问题, 所以召回率有所上升。由于 BL 模型是先识别主语, 后识别关系宾语, 能够解决三元组重叠问题, 所以召回率和准确率达到 74.2% 和 75.1%。HSL 在 BL 的基础上增加更多先验信息(字词向量和位置向量), 增加 GLU Dilated CNN 编码器和 Self Attention 机制, 比 BL 模型的 F1 值提升 10%, 充分说明 HSL 的有效性。

军事语料数据集是武器装备领域的中文数据集, 在军事语料数据集上不同模型的实验结果如表 3 所示。可以看出, BL 模型的准确率最高, 但由于其编码能力弱和无先验信息, 导致识别出的三元组偏少, 所以召回率只有 60.50%。HSL 模型的召回率为 76.28%, 说明其具有较强编码能力, 能够识别更多的三元组。由于其他模型需要先分词, 再进行分类任务, 所以 F1 值肯定不理想, 本文不做比较。

为进一步证明 HSL 具有良好的解决三元组重叠问题能力, 分别在不同重叠程度的 WebNLG 数据集和军事语料数据集上统计 HSL 的准确率、召回率和 F1 值(表 4)。可以看出, HSL 在不同重叠程度数据集上的 F1 值都高于 75%, 说明其具有解决三元组重叠问题的能力。

3 结语

本文提出一种基于分层序列标注的实体关系三元组抽取模型(HSL), 通过加入位置编码向量来增加模型对位置的注意力, 采用 GLU Dilated CNN 和 Self Attention 机制来增加模型特征抽取能力, 通过分层序列标注的方式来解决三元组重叠问题。实

表 4 HSL 在不同重叠程度的军事语料数据集和 WebNLG 数据集上的实验结果(%)

Table 4 HSL experimental results on military corpus data set and WebNLG data set with different degrees of overlap (%)

数据集	重叠程度	P	R	F1
军事语料数据集	Normal	86.02	83.33	84.65
	SPO	85.98	75.29	80.28
WebNLG 数据集	Normal	71.93	82.00	76.64
	SPO	86.75	87.16	86.95
	EPO	84.32	81.81	83.05

验证明, 无论在何种三元组重叠情况下, HSL 模型的 F1 值均高于目前主流的三元组抽取模型; 在 WebNLG 数据集和军事语料数据集上的 F1 值均高于主流三元组抽取模型。

参考文献

- [1] Nadeau D, Sekine S. A survey of named entity recognition and classification. *Linguisticae Investigationes*, 2007, 30(1): 3–26
- [2] Zelenko D, Aone C, Richardella A. Kernel methods for relation extraction. *Journal of Machine Learning Research*, 2003, 3(6): 1083–1106
- [3] Chan Y S, Roth D. Exploiting syntacticosemantic structures for relation extraction // Meeting of the Association for Computational Linguistics: Human Language Technologies. Portland, 2011: 551–560
- [4] Ren Xiang, Wu Zeqiu, He Wenqi, et al. Cotype: joint extraction of typed entities and relations with knowledge bases // 26th International Conference. International World Wide Web Conferences Steering Committee. Perth, 2017: 1015–1024
- [5] Li Qi, Ji Heng. Incremental joint extraction of entity mentions and relations // Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Baltimore, 2014: 402–412
- [6] Miwa M, Sasaki Y. Modeling joint entity and relation extraction with table representation // Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, 2014: 1858–1869
- [7] 王国昱. 基于深度学习的中文命名实体识别研究[D]. 北京: 北京工业大学, 2015
- [8] Zeng Daojian, Liu Kang, Lai Siwei, et al. Relation classification via convolutional deep neural network // 25th International Conference on Computational Linguistics: Technical Papers. Dublin, 2014: 2335–2344
- [9] Xu Kun, Feng Yansong, Huang Songfang, et al. Semantic relation classification via convolutional neural networks with simple negative sampling. *Computer Science*, 2015, 71(7): 941–9
- [10] Zheng Suncong, Wang Feng, Bao Hongyun, et al. Joint extraction of entities and relations based on a novel tagging scheme [EB/OL]. (2017–07–07)[2020–05–20]. <https://arxiv.org/abs/1706.05075>
- [11] Zeng Xiangrong, Zeng Daojian, He Shizhu, et al.

- 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 (Volume 1: Long Papers). Melbourne, 2018: 506–514
- [12] Fu T J, Li P H, Ma W Y. GraphRel: modeling text as relational graphs for joint entity and relation extraction // Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, 2019: 1409–1418
- [13] 李明耀, 杨静. 基于依存分析的开放式中文实体关系抽取方法. 计算机工程, 2016, 42(6): 201–207
- [14] 黄培馨, 赵翔, 方阳, 等. 融合对抗训练的端到端知识三元组联合抽取. 计算机研究与发展, 2019, 56(12): 2536–2548
- [15] 赵哲焕. 生物医学实体关系抽取研究[D]. 大连: 大连理工大学, 2017
- [16] 张永真, 吕学强, 申闫春, 等. 基于SAO结构的中文专利实体关系抽取. 计算机工程与设计, 2019, 40(3): 706–712
- [17] 王昊. 面向网络的中文实体关系抽取的研究[D]. 北京: 中国科学院大学, 2015
- [18] 苏剑林. 百度三元组抽取比赛 Baseline [EB/OL]. (2019–06–03)[2020–05–20]. <https://github.com/bojone/kg-2019-baseline>
- [19] Sutskever I, Vinyals O, Le Q V. Sequence to sequence learning with neural networks [EB/OL]. (2014–12–14)[2020–05–20]. <https://arxiv.org/abs/1409.3215>
- [20] Kim Y. Convolutional neural networks for sentence classification [EB/OL]. (2014–09–03) [2020–05–20]. <https://arxiv.org/abs/1408.5882>
- [21] Dauphin Y N, Fan A, Auli M, et al. Language modeling with gated convolutional networks // Proceedings of the 34th International Conference on Machine Learning-Volume 70. Sydney, 2017: 933–941
- [22] Gehring J, Auli M, Grangier D, et al. Convolutional sequence to sequence learning // Proceedings of the 34th International Conference on Machine Learning-Volume 70. Sydney, 2017: 1243–1252
- [23] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions [EB/OL]. (2016–04–30) [2020–05–20]. <https://arxiv.org/abs/1511.07122>
- [24] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need // Advances in Neural Information Processing Systems. Red Hook, 2017: 5998–6008
- [25] Gardent C, Shimorina A, Narayan S, et al. Creating training corpora for NLG micro-planning // Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. Vancouver, 2017: hal-01623744