

基于胶囊网络的药物相互作用关系抽取方法

刘宁宁¹, 琚生根^{1*}, 熊熙², 王婧妍¹, 张芮¹

(1. 四川大学 计算机学院, 四川 成都 610065;

2. 成都信息工程大学 网络空间安全学院, 四川 成都 610225)

摘要: 药物相互作用是指药物之间存在的抑制或促进等作用。针对目前药物关系抽取模型在长语句中抽取效果较差以及高层特征信息丢失的问题, 该文提出了一种结合最短依存路径的胶囊网络关系抽取模型, 该方法首先根据原语句解析出两个药物之间的最短依存路径, 然后利用双向长短期记忆网络分别获取原语句和最短依存路径的低层语义表示, 再将二者结合输入到胶囊网络中, 利用胶囊网络的动态路由机制, 动态的决定低层胶囊向高层胶囊传递的信息量, 避免了高层特征信息丢失的问题, 从而提升抽取效果。在 DDIExtraction 2013 药物相互作用关系抽取任务上的实验结果表明, 该文方法的 F1 值优于目前最优方法 1.17%。

关键词: 药物关系抽取; 最短依存路径; 双向长短期记忆网络; 胶囊网络

中图分类号: TP391

文献标识码: A

An Extraction Method for Drug-Drug Relationships Based on Capsule Networks

LIU Ningning¹, JU Shenggen^{1*}, XIONG Xi², WANG Jingyan¹, ZHANG Rui¹

(1. School of Computer Science, Sichuan University, Chengdu, Sichuan 610065, China;

2. School of Cybersecurity, Chengdu University of Information Technology, Chengdu, Sichuan 610225, China)

Abstract: Drug-Drug Interaction refers to the inhibition or promotion between drugs. In order to solve the problem that the current Drug-Drug Interaction relationship extraction model does not perform well in long sentences and a part of high-level information maybe lost, this paper proposes a capsule network extraction model that combines the shortest dependent path. The approach first parses out the shortest dependent path between two drugs according to the original sentence, then obtain the low-level vector representation of the original sentence and the shortest dependent path, through the Bidirectional Long Short Term, and then put the two into the capsule network. Finally, the dynamic routing mechanism of the capsule network is used to dynamically determine the amount of information transmitted by the low-level capsules to the high-level capsules, so that the problem of high-level feature information loss can be avoid, and the extraction effect can be improved. The experimental results on the DDIExtraction2013 show that the F1 value of the approach in this paper is better than the current best approaches of 1.17%.

Key words: Drug Relationship Extraction; Shortest Dependent Path; Bidirectional Long Short Term Memory; Capsule network

0 引言

药物-药物相互作用 (DDI, Drug-Drug Interaction) 是指当患者同时服用两种或多种药物时, 一种药物的药效可能会受到其他药物的影响, 由此可能会产生副作用,

导致治疗费用增加且对患者的生命安全造成威胁, 因此了解药物之间的相互作用具有非常重要的意义与医学应用价值。

目前人们主要是从 DrugBank^[1], PharmGKB^[2] 等医学数据库中获取药物相关知识, 但是随着生物医学文献呈指数级增长, 通过人工手动从中抽取有用的信息来

基金项目: 四川省重点研发项目 (2018GZ0182, 2018GZ0253, 2019YFS0236)

通讯作者: 琚生根 (jsg@scu.edu.cn)

更新数据库的方式变的低效且代价较高，由此造成医学数据库的更新速度较慢，生物医学文献中丰富的医学资源得不到充分利用。DDI自动抽取技术的研究可以实现医学数据库的快速更新，从而可以帮助医学从业者获取较新且较为全面的药物信息。随着DDIExtraction2011^[3]和DDIExtraction2013^[4]这两次药物相互作用关系抽取任务的发布，如何自动的从文本中抽取DDI逐渐成为研究热点。目前已有许多方法成功应用在DDI抽取任务上。

早期的DDI抽取方法是基于规则的，需要专业领域人员参与制定一系列规则来完成DDI的抽取，由于语言表达形式的多样性，制定的规则往往难以覆盖所有的药物相互作用关系，因此该方法的召回率较低。随着机器学习技术的不断发展，利用机器学习技术进行DDI抽取的方法逐渐增多，此类方法通常需要利用大量人工定义特征，如词性，句法，语法等特征，且需要利用外部自然语言处理工具生成这些特征，如词性标注器，句法分析器等工具，因此其抽取性能受外部自然语言处理工具的影响较大。由于深度学习技术具有自动学习特征的能力，可以减少人工设计特征所耗费的代价且抽取效果一般比传统的方法好，因此应用深度学习技术进行DDI抽取逐渐成为一种趋势，目前DDI关系抽取领域常用的深度学习模型有卷积神经网络（Convolutional Neural Network, CNN）和循环神经网络（Recurrent Neural Network, RNN）。

胶囊网络是由Sabour^[5]在2017年提出的用于图像处理领域的一种网络结构。针对CNN采用最大池化或平均池化等池化策略会造成信息丢失的问题，胶囊网络的思想是利用动态路由机制动态的决定低层胶囊向高层胶囊传递的信息量，从而可以充分利用高层特征信息。胶囊网络已在图像领域显示出了良好的效果，也有相关研究将其用在自然语言处理领域的任务上，如文本分类^[6]，情感分析^[7,8]，机器翻译^[9]，其

在DDI抽取任务上的应用亟待进一步研究。

本文提出了一种结合最短依存路径的胶囊网络关系抽取模型，针对医学文本中语句较长，句子结构较为复杂的问题，本文通过提取两个药物实体的最短依存路径（Shortest Dependency Path, SDP），并利用双向长短期记忆网络（Bidirectional Long Short-Term Memory, BiLSTM）擅长于捕获长距离序列信息的能力，将原语句与最短依存路径的信息相结合，从而可以获得包含更多句子信息的低层句子表示。针对当前大多药物关系抽取模型未能充分利用高层信息，高层信息存在丢失的问题，本文提出利用胶囊网络的动态传送信息的优势，使高层信息得到充分利用。最后在数据集DDIExtraction2013上验证了本文模型的有效性。

本文章节结构排列方式如下：第一节介绍关于DDI抽取的相关工作；第二节中详细的描述了本文提出的模型；第三节中使用DDIExtraction2013数据集验证本文模型的有效性，并与其它模型进行对比。第四节中总结全文并提出了未来的发展方向。

1 相关工作

目前在药物相互作用关系抽取领域应用的方法主要分为四类，分别是基于规则的方法，基于特征的方法，基于核函数的方法以及基于深度学习的方法。

基于规则的方法是一种比较传统的方法，一般需要根据数据集中语句的结构或由领域专业人员制定一些模式和规则。Tari等人^[10]通过对句子的语法结构进行分析，将药物之间相互作用的知识以规则的形式表示出来，然后利用该规则进行关系的提取。Segura等人^[11]设计了一种结合浅层解析，句法简化与模式匹配的DDI抽取算法，首先将复杂句和复合句分解为简单句，然后利用药剂师制定的一组规则从简单句中抽取相互作用的药物对。

基于特征和基于核函数的方法需要使

用大量事先定义的特征，如词性，语义，药物名等特征来完成对药物关系的抽取。He等人^[12]在使用一般特征的基础上加入了关键词特征，药物数据库中的特征等领域特征。Rastegar等人^[13]将两个药物名之间的词性，词干，以及词的数目作为特征。Kim等人^[14]使用了两个药物之间的最短依存路径作为特征。Chowdhury等人^[15]将特征核，图核和树核相结合，减少了丢失重要特征的风险，取得了比单个核更好的结果。Zhang等人^[16]设计了一种使用句法结构信息的单一内核函数，并通过剪枝方法去除噪声，取得了较好的结果。

近年来基于神经网络的方法在药物关系抽取领域取得了很好的进展。Liu等人^[17]首次提出利用CNN模型进行药物关系抽取，只需要将单词嵌入和位置嵌入信息输入到卷积网络中，实验结果表明，该方法可以在不提取任何词性，句法等特征的基础上取得较好的结果，由此证明深度学习方法适用于解决药物关系抽取问题，并且可以取得比基于特征和基于核函数的方法更好的结果。为了更好的获取句子语义的表示，Quan^[18]采用多通道卷积神经网络的

方法，使用了在不同医学文献上训练的词向量作为CNN的输入，可以更好的表示单词的语义信息。Kavaluru等人^[19]将词级别和字符级别的RNN模型相融合，以更好的表示句子语义信息。考虑到句子中不同部分的重要程度不同，也有相关论文通过加入注意力机制提高模型的抽取效果^[20,21]。Asada等人^[20]考虑到在一句话中不同单词的重要性不同，提出加入单词注意力机制，给不同的单词分配不同的注意力权重，给重要的单词赋予较大的权重，从而提升模型的预测效果。Zhou等人^[21]考虑到单词的位置信息对于判断药物之间的关系有很重要的作用，提出将bilstm层产生的隐层状态与位置嵌入结合生成位置感知注意力，从而可以更好的利用位置信息。考虑到外部现有的丰富的生物医学资源知识有助于提升药物关系抽取的准确性，Asada等人^[22]将图卷积神经网络的中间层与卷积神经网络的中间层相拼接，以更好的利用药物的分子结构信息。Xu等人^[23]通过将外部现有的生物医学资源与词汇信息和实体位置信息结合在一起，来完成从生物文献中抽取DDI。

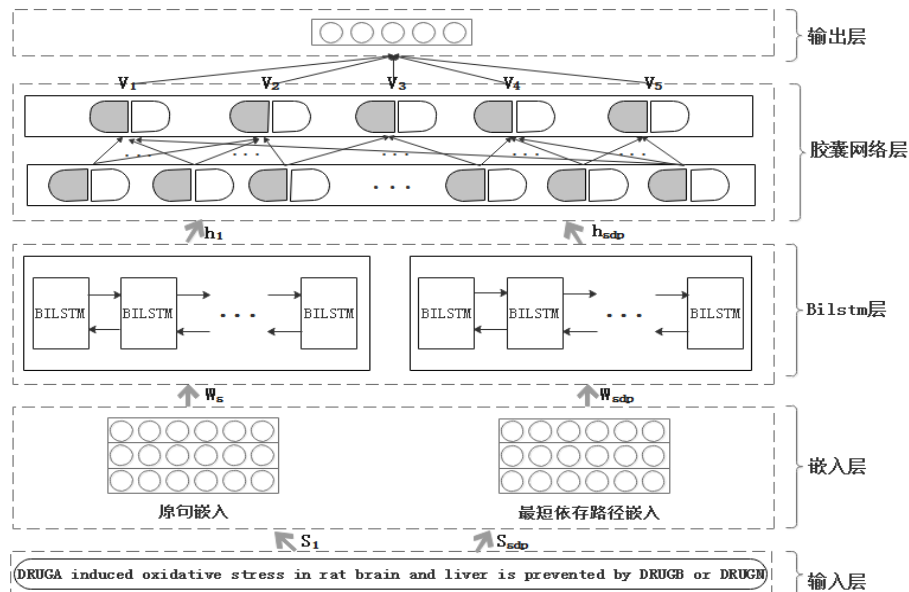


图1 模型整体框架图

2 本文方法

本文提出了一种结合最短依存路径的利用胶囊网络进行药物关系抽取模型，模型整体框架图如图 1 所示。该模型首先根

据输入语句 s_1 解析出语句的最短依存路径 s_{sdp} ，然后分别将原语句和最短依存路径信息输入到嵌入层，利用嵌入层将语句转换为向量表示，再将转换为向量表示的原句嵌入和最短依存路径嵌入信息分别输入到 BiLSTM 层，从而捕获 s_1 和 s_{sdp} 的低层语义特征 h_1 和 h_{sdp} ，将 h_1 和 h_{sdp} 进行拼接后送入到胶囊网络，获取高层特征，最后输出两个药物实体的类别。

2.1 输入层

首先根据原语句中的药物实体，生成相互作用的药物对，然后进行药物盲化处理，即将药物实体分别用“DRUGA”，“DRUGB”和“DRUGN”代替，其中“DRUGA”和“DRUGB”代表要判断相互作用关系的药物实体，“DRUGN”代表无关药物实体。本文使用的数据集中已标明了药物实体的位置，故不需要进行命名实体识别等操作识别药物实体。

假设原语句 s 为：“Cypermethrin induced oxidative stress in rat brain and liver is prevented by vitamin E or allopurinol.”，其中“Cypermethrin”，“vitamin E”以及“allopurinol”表示药物实体，该句中一共有三个药物实体，在经过两两组合之后，可得三个药物对，即产生三个输入语句，如表 1 所示。

表 1 输入语句处理

药物对	类别	语句
(Cypermethrin, vitamin E)	effect	DRUGA induced oxidative stress in rat brain and liver is prevented by DRUGB or DRUGN.
(Cypermethrin, allopurinol)	effect	DRUGA induced oxidative stress in rat brain and liver is prevented by DRUGN or DRUGB.
(vitamin E, allopurinol)	false	DRUGN induced oxidative stress in rat brain and liver is prevented by DRUGA or DRUGB.

2.2 嵌入层

嵌入层是将输入文本语句转换为向量

的形式。在嵌入层中包括两个模块，分别是原句嵌入和最短依存路径嵌入。其中原句嵌入模块的输入信息包含单词信息，位置信息以及药物类型信息。最短依存路径嵌入模块的输入是利用斯坦福解析器获取的两个药物实体之间的最短依存路径。

假设输入原句 $S = \{w_1, w_2, \dots, w_n\}$ ，其中 w_i 代表句子中的第 i 个词，通过设置两个位置向量 p_1 和 p_2 ，来表示句子中的每个单词到两个药物实体的距离，此外用 t_1 和 t_2 表示两个药物的类型。在原句嵌入中，本文采用以词嵌入，位置嵌入和类型嵌入相拼接的方法表示句子，其中词嵌入是将单词以向量的形式表达出来，本文采用的是在大规模语料库上训练的 glove 词向量。则单个嵌入向量表示为如公式 (1) 所示。

$$x_i = d_w \oplus d_{p_1} \oplus d_{p_2} \oplus d_{t_1} \oplus d_{t_2} \quad (1)$$

其中 $x_i \in R^{d+2p+2t}$ ， d 代表单词嵌入的维度， p 代表位置嵌入的维度， t 代表类型嵌入的维度。则原句嵌入向量可以表示为 $W_s = \{x_1, x_2, \dots, x_n\}$ ，最短依存路径嵌入向量可以表示为 W_{sdp} 。

2.3 BiLSTM 层

长短期记忆网络 (Long Short Term Memory, LSTM) 是一种特殊结构的 RNN 模型，有效解决了循环神经网络模型面临的梯度消失问题。LSTM 由输入门，遗忘门，输出门以及细胞状态组成，采用公式 (2) 至公式 (7) 进行更新。

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (3)$$

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

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

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t \odot \tanh(C_t) \quad (7)$$

其中 i_t 表示输入门， f_t 表示遗忘门， o_t 表示输出门， h_{t-1} 和 C_{t-1} 表示前一个时间步 $t-1$ 的隐藏状态和细胞状态， h_t 和 C_t 表示当前时间步 t 的隐藏状态和细胞状态。

BiLSTM是LSTM的一种变形，它采用两个LSTM分别获取语句的前向信息和后向信息，然后将二者得到的向量拼接起来作为BiLSTM的输出，假设前向过程的输出的隐藏状态用 \vec{h}_1 表示，后向过程的输出隐藏状态用 \overleftarrow{h}_n 表示，利用公式(8)可得到BiLSTM的输出。

$$h = [\vec{h}_n, \overleftarrow{h}_1] \quad (8)$$

使用BiLSTM网络可以较好的捕获句子的全局序列信息，适用于长句子的处理，故本文采用BiLSTM结构分别获取原句的向量表示 h_1 和最短依存路径信息的向量表示 h_{sdp} ，将二者拼接起来得到最终的向量表示 h_{all} ，将其作为胶囊网络层的输入。

2.4 胶囊网络层

胶囊是由一组神经元组成的，每个胶囊负责确定对象的单个部分，所有胶囊共同决定对象的整体架构。胶囊网络的结构如图2所示。主要分为低层胶囊层，动态路由层和高层胶囊层。胶囊网络利用动态路由机制，将低层胶囊的信息动态的传送到高层胶囊中从而可以克服CNN中池化层信息丢失的问题。

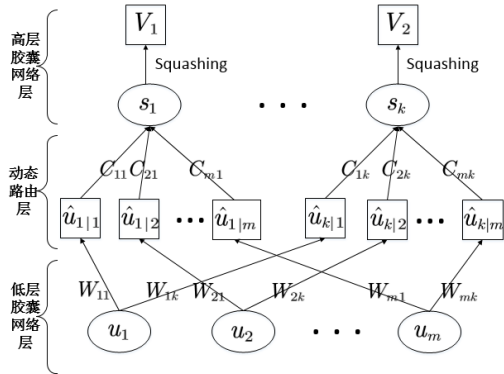


图2 胶囊网络结构

假设 u_i 是第 i 个低层胶囊的输出，则它预测的高层胶囊 j 的输出可以由公式(9)计算得到。

$$\hat{u}_{j|i} = W_{ij} u_j \quad (9)$$

其中 W_{ij} 代表低层胶囊 i 和高层胶囊 j 之间的权重矩阵，该矩阵是在训练过程中学习得到。

在高层胶囊网络中 s_j 是高层胶囊网络的输入， v_j 是高层胶囊网络的输出，其中 s_j 是根据低层胶囊预测向量 $\hat{u}_{j|i}$ 和低层胶囊相对应的耦合系数 c_{ij} 得到的，如公式(10)所示。

$$s_j = \sum_i c_{ij} \cdot \hat{u}_{j|i} \quad (10)$$

在胶囊网络中，向量的长度代表概率值，故用如下非线性压缩函数，将高层胶囊输出向量的长度转换为0到1之间。如公式(11)所示。

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \frac{s_j}{\|s_j\|} \quad (11)$$

低层胶囊网络和高层胶囊网络之间信息传递量是由耦合系数 c_{ij} 决定的， c_{ij} 的计算过程如公式(12)所示，其中 b_{ij} 表示胶囊 i 和胶囊 j 之间耦合的对数概率，初始化为0， b_{ij} 会在路由迭代中，动态的更新。

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_{i=1}^K \exp(b_{ik})}, \quad b_{ij} \leftarrow b_{ij} + \hat{u}_{j|i} \cdot v_j \quad (12)$$

由上式可以看出，当低层胶囊将信息传送到正确的高层胶囊时，耦合系数 c_{ij} 会变大，当传送到错误的高层胶囊时，耦合系数 c_{ij} 会变小，当 $c_{ij}=0$ 时表示低层胶囊 i 和高层胶囊 j 之间没有信息传递， c_{ij} 由动态路由算法迭代更新。

动态路由算法的伪代码如下所示：

算法1: Routing($\hat{u}_{j|i}, r, l$)

1. for every capsule j in layer $l+1$ and capsule i in layer l : $b_{j|i} \leftarrow 0$.
2. for r interactions do
3. for every capsule i in layer l :
 $c_i \leftarrow \text{soft max}(b_i)$ corresponds to Eq. (12)
4. for every capsule j in layer $l+1$:
 $s_j \leftarrow \sum_i c_{ij} \hat{u}_{j|i}$
5. for every capsule j in layer $l+1$:
 $v_j = \text{squash}(s_j)$ corresponds to Eq. (11)
6. for every capsule i in layer l and capsule j in layer $l+1$:
 $b_{j|i} = b_{j|i} + \hat{u}_{j|i} \cdot v_j$
7. End for
8. return v_j

2.5 输出层

由于在胶囊网络中，高层胶囊的向量长度代表该类别的概率，故在该层中，选择具有最大输出向量长度的高层胶囊类别作为最终模型预测的类别。

模型的损失函数如公式（13）所示。高层胶囊网络中的胶囊数量与数据集中药物关系的类别数目相同，每个胶囊代表一个关系类别，若输入语句的关系类别为 k ，则 T_k 值为1，否则 T_k 值为0，其中 m^+ ， m^- 和 λ 是需要事先指定的超参数。

$$L_k = T_k \max(0, m^+ - \|v_k\|)^2 + \lambda(1 - T_k) \max(0, \|v_k\| - m^-)^2 \quad (13)$$

3 实验

3.1 数据集介绍及评价指标

本文采用为DDIExtraction2013数据集进行实验，该数据集中药物之间相互作用关系分为如下5种类型：

- 1) Advice: 文本中描述了关于两种药物同时使用的建议。
- 2) Effect: 文本中明确指出两种药物相互作用的结果。
- 3) Mechanism: 文本中明确讨论了药物动力学机制。
- 4) Int: 文本中说明两种药物存在一定关系，但未定义具体的关系类型。
- 5) Negative: 两个药物之间不存在相互作用关系。

数据集的详细信息如表2所示。

表2 DDIExtraction2013 数据描述

	训练集	测试集	总数
Advice	826	221	1047
Effect	1687	360	2047
Mechanism	1319	302	1621
Int	188	96	284
Negative	23772	4737	28509

本文采用精准率，召回率以及F1值这三个指标对实验结果进行评价。

3.2 参数设置

本文模型中使用的参数的取值如表3所示。

表3 参数取值

参数	取值
Epoch	20
Batchsize	64
词嵌入维度	300
位置嵌入维度	10
药物类型嵌入	10
学习率	0.001
m^+	0.9
m^-	0.1
λ	0.25
路由迭代次数	3

3.3 实验结果

本文模型在DDIExtraction2013上的实验结果如图3和图4所示，其中图3是模型的损失值曲线图，图4是模型的F1值曲线图。由图3可看出该模型在开始训练时损失降低的速度很快，然后趋于平稳。由图4可看出，模型训练的前期F1值提升较快，后续增加速度较慢，然后逐渐趋于平稳。

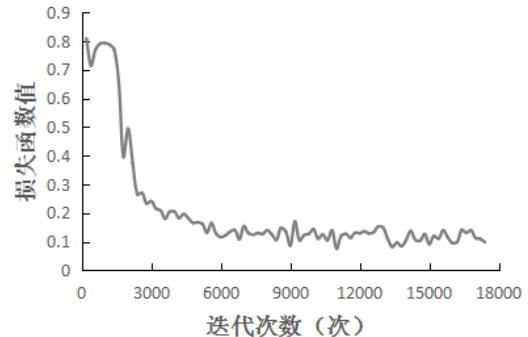


图3 损失值曲线图

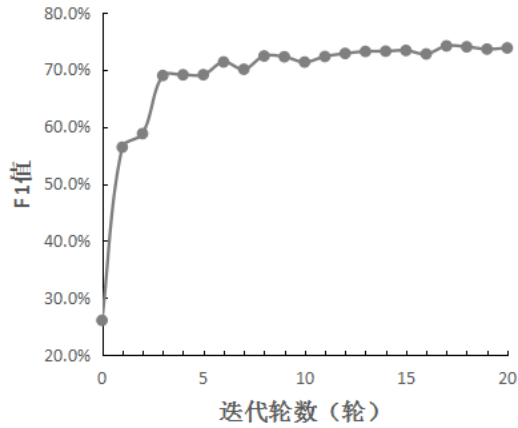


图4 F1 值曲线图

为了分析模型不同模块对抽取效果的影响程度, 本文设计了消融实验, 如表4所示, 分别验证了BiLSTM, Capsule和SDP对实验结果的影响。

表4 消融实验

模型	P (%)	R (%)	F1 (%)
本文模型	78.78	70.05	74.16
-BiLSTM	75.47	65.52	70.15
-Capsule	73.88	68.74	71.22
-SDP	76.87	68.33	72.35

由消融实验结果可以看出, 去掉BiLSTM层后, 模型的效果降低4.01%, 说明在不使用BiLSTM获取低层语句特征信息的情况下, 仅单独使用胶囊网络进行关系抽取的效果不是很理想。去掉最短依存路径信息后, 模型的F1值降低1.81%, 说明结合最短依存路径信息可以更好的丰富句子的语义信息, 比单纯使用原语句信息的效果要好。去掉胶囊网络层后, 模型的F1值降低2.94%, 表明胶囊网络动态利用高层信息的能力可以显著的提升模型的效果。消融实验的结果表明了本文提出的模型可以充分结合最短依存路径, BiLSTM以及胶囊网络这三者的优势, 从而较好的提升模型的抽取效果。

将本文模型的实验结果与现有的结果进行对比。对比结果如表5所示。可以看出本文模型的F1值比其中效果最好的模型高1.17%。

表5 准确率对比

文献	P (%)	R (%)	F1 (%)
Liu et al. (2016) ^[17]	75.72	64.66	69.75
Quan et al. (2016) ^[18]	75.99	65.25	70.21
Wang et al. (2017) ^[24]	72.53	71.49	72.00
Yi et al. (2017) ^[25]	73.67	70.79	72.20
Xu et al. (2018) ^[23]	71.52	70.79	71.15
Asada et al. (2018) ^[22]	73.31	71.81	72.55
Zhou et al. (2018) ^[21]	75.8	70.38	72.99
本文	78.78	70.05	74.16

4 结论

本文充分利用BiLSTM捕获长距离信息的能力和胶囊网络动态结合高层信息的优势, 将二者结合用于药物相互作用关系抽取, 并使用最短依存路径信息丰富语句信息, 实验结果表明, 本文模型在从复杂的医学文本中抽取药物相互作用的效果较好。在未来的工作中可以尝试加入注意力机制, 即给原句信息和最短依存信息分配合适的权重, 从而提升模型的效果。针对数据集中负例较多的现象, 在未来的工作中可以考虑使用数据增强等操作, 以平衡不同类型数据之间的数量差异。

参考文献:

- [1] DrugBank 4.0: Shedding new light on drug metabolism[J]. *Nucleic Acids Research*, 2014, 42(D1):D1091-D1097.
- [2] Thorn C F, Klein T E, Altman R B. PharmGKB: The Pharmacogenomics Knowledge Base[J]. *Methods in Molecular Biology*.
- [3] Segura Bedmar I, Martinez P, Sánchez Cisneros D. The 1st DDIE Extraction-2011 challenge task: Extraction of Drug-Drug Interactions from biomedical texts[J]. 2011.
- [4] Segura Bedmar I, Martínez P, Herrero Zazo M. Semeval-2013 task 9: Extraction of drug-drug interactions from biomedical texts (ddiextraction 2013)[C]. *Association for Computational Linguistics*, 2013.
- [5] Sabour S, Frosst N, Hinton G E. Dynamic Routing Between Capsules[J]. 2017.
- [6] Zhao W, Ye J, Yang M, et al. Investigating

- Capsule Networks with Dynamic Routing for Text Classification[J]. 2018.
- [7] Wang Y, Sun A, Han J, et al. Sentiment analysis by capsules[C]//Proceedings of the 2018 World Wide Web Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2018: 1165-1174.
- [8] Wang Y, Sun A, Huang M, et al. Aspect-level Sentiment Analysis using AS-Capsules[J]. 2019.
- [9] Zheng Z, Huang S, Tu Z, et al. Dynamic Past and Future for Neural Machine Translation[J]. 2019.
- [10] Tari L, Anwar S, Liang S, et al. Discovering drug-drug interactions: a text-mining and reasoning approach based on properties of drug metabolism[J]. *Bioinformatics*, 2010, 26(18):i547-i553.
- [11] Segura-Bedmar I, Paloma Martínez, César de Pablo-Sánchez. A linguistic rule-based approach to extract drug-drug interactions from pharmacological documents[J]. *BMC Bioinformatics*, 2011, 12 Suppl 2(Suppl 2):S1.
- [12] He L, Yang Z, Zhao Z, et al. Extracting drug-drug interaction from the biomedical literature using a stacked generalization-based approach[J]. *PloS one*, 2013, 8(6): e65814.
- [13] Rastegar-Mojarad M, Boyce R D, Prasad R. UWM-TRIADS: classifying drug-drug interactions with two-stage SVM and post-processing[C]//Second Joint Conference on Lexical and Computational Semantics (* SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013). 2013, 2: 667-674.
- [14] Kim S, Liu H, Yeganova L, et al. Extracting drug-drug interactions from literature using a rich feature-based linear kernel approach[J]. *Journal of Biomedical Informatics*, 2015, 55:23-30.
- [15] Chowdhury M F M, Lavelli A. Exploiting the scope of negations and heterogeneous features for relation extraction: A case study for drug-drug interaction extraction[C]//Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2013: 765-771.
- [16] Yijia Zhang, Hongfei Lin, Zhihao Yang, et al. A Single Kernel-Based Approach to Extract Drug-Drug Interactions from Biomedical Literature[J]. *PLoS ONE*, 2012, 7(11):e48901-.
- [17] Shengyu L, Buzhou T, Qingcai C, et al. Drug-Drug Interaction Extraction via Convolutional Neural Networks[J]. *Computational and Mathematical Methods in Medicine*, 2016, 2016:1-8.
- [18] Chanqin Q, Lei H, Xiao S, et al. Multichannel Convolutional Neural Network for Biological Relation Extraction[J]. *BioMed Research International*, 2016, 2016:1-10.
- [19] Kavuluru R, Rios A, Tran T. Extracting Drug-Drug Interactions with Word and Character-Level Recurrent Neural Networks[C]// IEEE International Conference on Healthcare Informatics. IEEE, 2017.
- [20] Asada M, Miwa M, Sasaki Y. Extracting drug-drug interactions with attention CNNs[C]//BioNLP 2017. 2017: 9-18.
- [21] Zhou D, Miao L, He Y. Position-aware deep multi-task learning for drug-drug interaction extraction[J]. *Artificial Intelligence in Medicine*, 2018:S0933365717306310.
- [22] Asada M, Miwa M, Sasaki Y. Enhancing Drug-Drug Interaction Extraction from Texts by Molecular Structure Information[J]. *arXiv preprint arXiv:1805.05593*, 2018.
- [23] Xu B, Shi X, Zhao Z, et al. Leveraging biomedical resources in bi-lstm for drug-drug interaction extraction[J]. *IEEE Access*, 2018, 6: 33432-33439.
- [24] Wang W, Yang X, Yang C, et al. Dependency-based long short term memory network for drug-drug interaction extraction[J]. *BMC bioinformatics*, 2017, 18(16): 578.
- [25] Yi Z, Li S, Yu J, et al. Drug-drug interaction extraction via recurrent neural network with multiple attention layers[C]//International Conference on Advanced Data Mining and Applications. Springer, Cham, 2017: 554-566.