

基于多维语义映射的关系抽取方法研究

程华龄, 陈艳平, 杨卫哲, 秦永彬, 黄瑞章

引用本文

程华龄, 陈艳平, 杨卫哲, 秦永彬, 黄瑞章. [基于多维语义映射的关系抽取方法研究](#)[J]. 计算机科学, 2022, 49(11): 206-211.

CHENG Hua-ling, CHEN Yan-ping, YANG Wei-zhe, QIN Yong-bin, HUANG Rui-zhang. [Relation Extraction Based on Multidimensional Semantic Mapping](#)[J]. Computer Science, 2022, 49(11): 206-211.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于全局增强图神经网络的序列推荐](#)

Sequence Recommendation Based on Global Enhanced Graph Neural Network

计算机科学, 2022, 49(9): 55-63. <https://doi.org/10.11896/jsjcx.210700085>

[多层注意力机制融合的顺序到序列中国连续手语识别和翻译](#)

Sequence-to-Sequence Chinese Continuous Sign Language Recognition and Translation with Multi-layer Attention Mechanism Fusion

计算机科学, 2022, 49(9): 155-161. <https://doi.org/10.11896/jsjcx.210800026>

[密码学智能化研究进展与分析](#)

Research Progress and Analysis on Intelligent Cryptology

计算机科学, 2022, 49(9): 288-296. <https://doi.org/10.11896/jsjcx.220300053>

[基于物理操作级模型的查询执行时间预测方法](#)

Query Performance Prediction Based on Physical Operation-level Models

计算机科学, 2022, 49(8): 49-55. <https://doi.org/10.11896/jsjcx.210700074>

[基于卷积神经网络的 APP 用户行为分析方法](#)

Analysis Method of APP User Behavior Based on Convolutional Neural Network

计算机科学, 2022, 49(8): 78-85. <https://doi.org/10.11896/jsjcx.210700121>

基于多维语义映射的关系抽取方法研究

程华龄 陈艳平 杨卫哲 秦永彬 黄瑞章

贵州大学计算机科学与技术学院 贵阳 550025

公共大数据国家重点实验室 贵州大学计算机科学与技术学院 贵阳 550025

(hlch.245@foxmail.com)

摘要 关系抽取旨在从句子中识别出实体对之间的关系类型。在关系抽取领域,目前主流的方法都使用了深度学习方法,但大部分方法在输入层没有对词向量进行深层次的讨论。针对这一不足,提出了一种基于多维语义映射的关系抽取方法,该方法的核心思想是将矩阵降维方法应用于神经网络模型输入层。通过将表示文本的词向量进行多维度的降维分解,使分解后的词向量能映射表示同一语句在不同维度上的语义信息。实验结果表明,在 Chinese Literature Text 和 SemEval-2010 Task8 数据集上 $F1$ 值分别达到了 75.3% 和 88.9%,验证了所提方法的有效性。

关键词 关系抽取;神经网络;多维映射;语义信息

中图法分类号 TP391

Relation Extraction Based on Multidimensional Semantic Mapping

CHENG Hua-ling, CHEN Yan-ping, YANG Wei-zhe, QIN Yong-bin and HUANG Rui-zhang

College of Computer Science and Technology, Guizhou University, Guiyang, 550025, China

State Key Laboratory of Public Big Data, College of Computer Science and Technology, Guizhou University, Guiyang 550025, China

Abstract Relation extraction aims to identify relation types between entities from texts. In the field of relation extraction, most of existing methods use deep learning methods, but they do not have in-depth discussion of word vectors in the input layer. To further exploit word vectors, this paper proposes a relation extraction method based on multi-dimensional semantic mapping. The core idea of the method is to reduce dimensionality of text feature matrix before the word vector enters the input layer. Experimental results show that the proposed method not only can reduce dimensionality effectively, but also can represent the semantic information of the same sentence in different dimensions, with its $F1$ of 75.3% and 88.9% on the Chinese Literature Text and SemEval-2010 Task8 datasets, respectively.

Keywords Relation extraction, Neural network, Multidimensional mapping, Semantic information

1 引言

互联网的高速发展带来信息的爆炸式增长,如何高效利用这些信息是信息抽取(Information Extraction, IE)技术的主要任务。关系抽取(Relation Extraction, RE)作为信息抽取领域的主要任务,旨在从非结构化文本中获取已知实体对之间的语义关系,能为事件抽取、机器翻译、自动问答、知识图谱等任务提供有力保障。如句子“The key was in a chest”,其中的实体对 key 和 chest 之间的关系为“Content-Container”。

传统关系抽取方法主要利用手工构建的词法^[1]或依赖人工设计的核函数^[2]。这类方法依靠手工提取,需要耗费大量人力且容易出错,而深度学习方法不需要专业人员来设计

语法和数据特征,而是让模型自动地去学习文本中的语法规则特征,缓解了传统方法错误传递的问题。然而,目前许多深度学习模型在输入层仅简单地做了词嵌入处理,在该方法中,单词输入的词向量维度是固定的,但是单词的含义却会随着上下文的变化而改变。为了获得更丰富的词向量表示,本文提出一种基于多维语义映射的关系抽取方法,该方法的核心思想是将矩阵降维方法应用在神经网络输入层,以获得更全面的语义信息。

本文的主要贡献在于从获取句子的语义信息角度出发,提出了一种基于多维语义映射的关系抽取方法。首先标注实体的相对位置信息来突出句子的结构信息,增强神经网络模型对句子结构特征的学习;然后通过 BERT^[3] 预训练语言

到稿日期:2021-09-14 返修日期:2022-04-18

基金项目:国家自然科学基金通用联合基金(U1836205);国家自然科学基金(62066007, 62066008);贵州省科学技术基金重点(黔科合基础[2020]1Z055)

This work was supported by the State Key Program of the Joint Funds of National Natural Science of China(U1836205), National Natural Science Foundation of China(62066007, 62066008) and Key Project of Guizhou Science and Technology Fund(Qian Ke He Ji Chu[2020]1Z055).

通信作者:秦永彬(ybqin@foxmail.com)

模型输出表示文本的词向量,该词向量经过多维语义映射层后能够获取到句子不同维度的语义信息;最后结合卷积神经网络(Convolutional Neural Networks,CNN)模型对文本特征进行局部捕获,提升关系抽取性能。实验数据显示,通过对多维语义信息的捕获,在 Chinese Literature Text^[4]数据集和 Semeval2010 task8^[5]数据集上 F1 值分别为 75.3% 和 88.9%。

2 相关工作

关系抽取是自然语言处理中一项至关重要的任务,其经典方法主要分为有监督、半监督、弱监督和无监督这 4 种。但这些经典方法都存在着特征提取方法上的误差问题,对关系抽取性能造成了影响。

为了避免传统经典方法依赖人工选择特征,造成特征提取过程中误差累积等问题,研究者将深度学习方法应用到关系抽取任务中。Zeng 等^[6]利用卷积深度神经网络提取词语和句子级别的特征,使用位置嵌入来表示单词和实体之间的距离,将位置嵌入和单词嵌入连接在一起,但是面对距离较远的实体对时性能发挥较差。Liu 等^[7]为了探究单词之间的语义,提出了一种新的关系分类的卷积网络,该网络将词汇特征作为网络的输入,相比人工设计规则更有效。Santos 等^[8]提出的卷积神经网络从词向量和位置向量的角度出发,提出了一个新的排名损失函数。Socher 等^[9]首次提出将循环神经网络(Recurrent Neural Network,RNN)模型应用到关系抽取,忽视了实体对对句子结构的影响。Xu 等^[10]提出了基于长短期记忆人工神经网络(Long Short-Term Memory,LSTM)深度学习方法,该方法将句法类型特征、WordNet 特征和词性特征结合起来进行关系抽取,有效避免了 RNN 梯度消失、梯度下降等问题。Devlin 等提出的 BERT(Bidirectional Encoder Representations from Transformers)模型应用了表达语义能力更强的双向 Transformer 网络结构。然而上述神经网络仅考虑了单词的外部特征,如词汇特征、词性特征,忽视了词向量查找表本身也含有丰富的语言信息。本文将矩阵降维方法应用在神经网络模型输入层,通过将表示文本的词向量进行多维度的降维分解,使分解后的词向量能表示同一语句在不同维度上的语义信息。

在语义映射方面,Cai 等^[11]提出了一种双向递归卷积神经网络模型,该模型将 LSTM 的双通道递归模型和卷积神经网络模型融合在一起,再基于最短依赖路径进行关系抽取,捕获两个相邻词的更多局部语义信息。Chen 等^[12]提出了一种基于多通道的深度神经网络,该网络将实体对划分后的五部分句子结构利用起来。Vu 等^[13]提出了一种基于 CNN 的模型,该模型通过实体对将句子划分为 3 个部分。Liu 等^[14]提出了一种新型的深度神经网络模型,即基于注意力的 BiGRU-CNN 网络,并将其应用于中文问题分类任务。Zhang 等^[15]提出了一种图像增强的多级句子表示网络,该方法从图像角度对句子进行不同维度的语义理解。Alt 等^[16]提出了 TRE(a Transformer based Relation Extraction)模型,该模型与之前的关系提取模型的不同之处在于,TRE 没有使用显式的语言特征来进行关系分类,而是采用预先训练的深度语言

表征和 Transformer 架构相结合,有效地建立起实体之间的长期依赖关系来进行关系抽取。Guo 等^[17]为了避免依存树中的有效信息被硬性粗剪枝方法消除,提出一种将整个依存树作为输入的图卷积神经网络关系抽取模型,在模型中添加注意力机制自动学习剪枝。Lei 等^[18]重点研究了在关系抽取方法中词的形态特征,通过结合词的形态特征和语义特征,充实输入向量的表示。Sun 等^[19]针对如何利用依存树中的有效信息,提出了一种可学习的语法传输注意力图卷积网络。该模型在不破坏原始依赖树的结构的情况下,直接对语法图进行传输操作,通过注意力机制将不同层次的特征融合进行分类。Wang 等^[20]提出了一种基于双向 SDP(Shortest Dependency Paths)的模型。该模型是最短路径依赖的一种新型表示形式,包括原始最短路径和它的反向最短路径,用于捕获关系语义词和方向词。

然而,上述神经网络更多地关注句子的局部语义信息,缺乏对句子全局语义信息的加强。因此本文提出了一种基于多维语义映射的关系抽取方法。通过对词向量进行不同维度的语义捕获,使降维后的词向量能表示同一句子在不同维度上的语义信息;再结合卷积神经网络对句子局部语义信息进行加强,从而获得更全面的语义信息。

3 模型结构

本文从捕获句子的语义信息出发,提出了一种基于多维语义映射的关系抽取模型(Multi-dimension Semantic Mapping Network,MD-SMN)。

本文方法主要包括 3 部分:词嵌入层、多维语义映射层和卷积层。本文关系抽取模型具体结构如图 1 所示,其中词嵌入层的词向量由 d_w 的字向量、 d_p 的位置向量和 d_s 的文本向量拼接而成。随后进入多维语义映射层,该层的主要目的是将词向量进行不同维度的映射,使得映射后的矩阵包含同一个句子在不同维度的语义信息。同时映射后的矩阵维度相比词嵌入层的词向量矩阵维度有所降低,再进入卷积层。不同大小的卷积核在提取句子语义时有更大的优势。

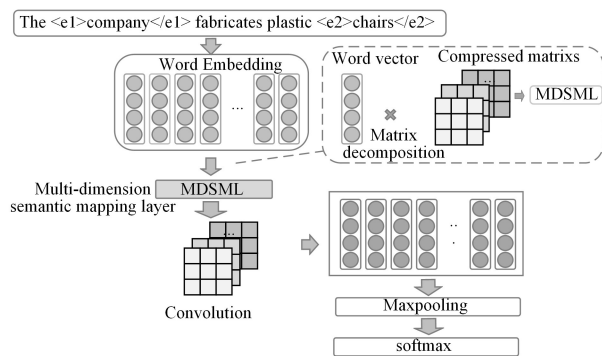


图 1 模型框架

Fig. 1 Model framework

为了提取多尺度的特征,本文使用了 3 个不同的卷积核来提取语义信息。最后进入池化层,对不同尺度的特征进行筛选,取多组特征的最大值并传递到分类层。

3.1 实例输入层

本文实验使用的数据集为 Chinese Literature Text 和 Se-

meval2010 Task8 数据集。它们的数据以分层结构存储在文件中,在处理数据集时,本文采用树的形式表示其中的文章,树的节点代表实体和关系,用树形解析的方法将数据输入存储为以下形式,采用符号 $\langle e \rangle$ 将实体对标记出来,以更好地体现句子的结构特征,具体如下:

CLT:河道清澈见底,鱼在水中随意游动, $\langle e2 \rangle$ 河边 $\langle /e2 \rangle$ $\langle e1 \rangle$ 绿树 $\langle /e1 \rangle$ 倒影在河中。

Semeval2010 task8: The $\langle e1 \rangle$ company $\langle /e1 \rangle$ fabricates plastic $\langle e2 \rangle$ chairs $\langle /e2 \rangle$

3.2 字嵌入层

机器学习的方法往往不能直接处理文本数据,因此本文需要找到合适的方法将文本数据转换为数字数据,从而引入词嵌入的概念。给定一个由 n 个单词组成的句子 $x, x = \{w_1, w_2, \dots, w_n\}$, 每个单词 w_n 转换为实值向量 r^{w_n} , 且 $r^{w_n} \in R^{d_w}$ 。通过查找字符嵌入矩阵对字符嵌入过程进行编码,得 $W_{\text{character}} \in R^{V \times d_w}, V$ 是训练词典的大小, d_w 是字符向量的维度。

本文在词嵌入层使用了 BERT 预训练模型, BERT 的预训练模型是在 wiki 百科语料库的基础上训练得到的语言模型。每个字在文本中出现的位置不同就具有不同的位置向量,再结合句子信息赋予每个字字向量和文本向量,结合这 3 个向量对输入的线性序列进行预处理得到字嵌入的输出,也是多维度语义信息转换层的输入,可表示为:

$$E_x = \{r^{w_1}, r^{w_2}, \dots, r^{w_n}\} \tag{1}$$

3.3 多维语义映射层

在机器学习中,为了避免计算和分析大量的数据,许多学者选择了降维的方法,通过降维可以降低预测变量的维度和数量,确保变量之间相互独立以及提供框架来解释实验结果。目前常见的降维方法有特征选择和特征变换两种。本文的关系抽取模型应用了特征变换方法。

特征变换会产生一组新的特征,而这些新特征压缩的效率更高。在关系抽取模型中,这个优点可以帮助神经网络更好地理解句子的语义信息。在多维度语义映射层,本文应用了矩阵分解方法,该方法本质上是一种基于原始特征空间的低秩近似。矩阵分解的优点在于降维后的特征更加突出,在神经网络模型中可以为后续提供一种可加模型。

E_x 是一个由字嵌入层输出的特征矩阵。设 k 是特征压缩后的维数,则通过矩阵分解的方法可以得到输出特征矩阵 A , 表示为:

$$A = \min_{W, H} \|E_x - WH\|_2 \tag{2}$$

其中,因为 k 的秩小于 E_x 的秩,所以 WH 是 E_x 的一个良好低秩近似。

通过该方法,神经网络的模型复杂度得以降低,效率也得到提高。为了让特征矩阵 A 包含更多的语义信息,本文随机选取了 5 个不同的矩阵 $W_1 H_1, W_2 H_2, \dots, W_5 H_5$ 对矩阵 A 进行矩阵分解,得到多维度语义空间转换层最后的输出 A_x 。 A_x 作为多维度语义映射层的输出,不仅降低了词向量的维度,而且捕获了 5 个不同维的句子语义信息,简化了卷积层的输入并加强了模型对语义信息的捕获。

3.4 卷积层

CNN 是一种多层的监督学习神经网络,网络结构由输入

层、隐藏层和输出层 3 个部分构成。隐藏层中的卷积层和池采样层是实现卷积神经网络特征提取功能的核心模块。通常将输入层到隐藏层的映射称为一个特征映射,本文利用这一特性对文本特征进行提取。

在本文的关系抽取模型中,卷积神经网络由输入层、卷积层、池化层、全连接层、输出层组成,每一部分都互相协调,共同抽取文本特征。为了提取句子不同尺度的特征,文中使用了 3 个不同的卷积核并行提取语义信息,分别是 $3 \times 3, 4 \times 4, 5 \times 5$, 通过这样的方法获得文本的抽象化特征,对该序列进行的卷积操作可形式化表示为:

$$c_i = f_c(W_c^T \cdot X^T + b) \tag{3}$$

其中, $W_c^T \in R^{K \times H}$ 是卷积神经网络中的滤波器, b 为偏置值, f_c 为非线性函数(如 ReLU 函数)。 c_i 的维度为 H , 当对 $\{r^{w_1}, r^{w_2}, \dots, r^{w_n}\}$ 进行卷积运算时,可使用式(4)表示:

$$c_i = \text{Conv}(r) \tag{4}$$

卷积神经网络可以对句子的局部特征进行有效捕捉,该网络将 K -gram 矩阵 $[r_i, r_{i+1}, \dots, r_L]$ 映射到一个高阶特征表示,通过高阶映射可以学习到 K -gram 的句法语义信息。为了得到对神经网络模型更有效的特征,将每一个卷积结果 c 进行最大池化操作,表示为:

$$p = \text{Pooling}(r) = \text{Max}(c_1, c_2, \dots, c_n) \tag{5}$$

经过池化层后,再利用全连接层对神经网络进行全局调剂,通过 Softmax 层对各关系类别概率进行输出,表示为:

$$y = \text{Softmax}(\text{Conv}(p)) \tag{6}$$

根据上述公式获得标记实例文本 A_x , 通过传统的卷积神经网络进行关系抽取的整体过程可表示为:

$$y = \text{Softmax}(\text{Conv}(\text{Pooling}(\text{Conv}(A_x)))) \tag{7}$$

4 实验与分析

4.1 实验数据及评价指标

本文采用中文数据集 Chinese Literature Text (CLT) 和英文数据集 Semeval-2010 Task8。作者从数百篇中文文章中建立了一个命名实体识别(Named Entity Recognition, NER)和关系抽取数据集。常见的数据集通常建立在句子级别,而该数据集是在篇章级别建立的。该数据集设置了 9 个不同的类,以便更好地了解实体之间的联系。经过处理,训练集有 16484 条,测试集有 1675 条,验证集有 1347 条。

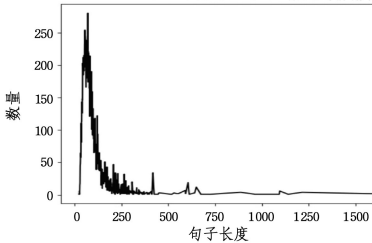
英文数据集 Semeval-2010 Task8 包含 8000 条训练句子和 2717 条测试句子,该数据集有 10 个关系类别,前 9 个关系是有向的,最后一个类型“Other”是无向的。

此外,本文采用评价指标 PRF 值对实验结果进行分析与评价,其中 P 为准确率(Precision), R 为召回率(Recall), F 即 F1 值。通过宏平均求 P 和 R 的调和平均值,用于反映模型综合性能,公式表示为 $F1 = 2 \times P \times R / (P + R)$ 。

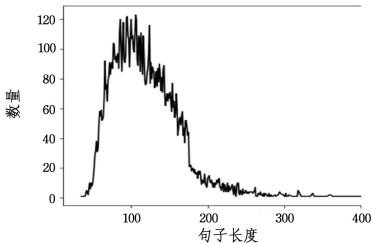
4.2 实验设置

考虑到句子语义信息的获取,本文参考了两个数据集文本长度的分布情况,如图 2 所示。一方面,在 CLT 数据集中,实例长度分布最集中的部分在 100~200 长度, Semeval-2010 Task8 数据集中实例长度分布最集中的部分在 80~125 长度;另一方面,CLT 数据集中实例长度超过 125 和 Semeval-

2010 Task8 数据集中实例长度超过 100 的样本实例较少。因此,在训练时为了防止特征稀疏造成性能下降的问题,将 CLT 数据集和 Semeval-2010 Task8 数据集的最大句子长度分别设为 125 和 100。



(a) Chinese literature text 数据分布折线图



(b) Semeval-2010 Task8 数据分布折线图

图 2 数据集长度分布图

Fig. 2 Diagram of dataset length distribution

模型参数设置如表 1 所列。本文经过 15~20 次迭代后获取了最优参数。

表 1 模型参数

Table 1 Hyperparameters

Hyperparameter	Description	Parameter value
l_s	Max sentence length	125/100
d_w	Size of word level vector	768
d_m	Size of multi-dimension semantic mapping layer	450
Learning rate	Convolutional layer	3×10^{-5}
Batch size	Size of Mini-Batch	32

4.3 实验结果与分析

4.3.1 局部与整体模型对比

本文的实验以图 1 中没有添加多维语义映射层的模型作为基线系统(Baseline)。本文的方法在没有添加多维语义映射层下,分类性能在 Chinese Literature Text 数据集和 Semeval-2010 Task8 数据集上分别能达到 71.6%和 85.4%。

在第一组实验中,本文在基线模型的基础上添加了多维语义映射层(MD-SML)。两个数据集的实验结果如表 2 所列。

表 2 数据集添加多维语义映射层的实验结果

Table 2 Experimental results of adding multidimensional semantic mapping layer to dataset

(单位:%)

Method	F1 (CLT dataset)	F1 (Semeval2010 Task8)
Baseline+	71.6	85.4
MD-SML	75.3	88.9

从表 2 的实验结果可以看出,在基线模型上添加多维语义映射层后,模型的分类效果得到有效提升。原因在于单独

的卷积循环网络关注的重点为局部特征语义,而多维语义映射层的核心思想是通过矩阵降维方法中的特征变换产生一组新的特征,而这些新特征压缩的效率更高。

在关系抽取神经网络模型中,该优点可以帮助神经网络更好地理解句子的语义信息,由此证明了添加 MD-SML 不仅降低了词向量的维度,而且捕获了 5 个不同维的句子语义信息,从而提高模型运算的效率,并简化卷积层的输入。

4.3.2 与现有方法对比

本文选取了当前主流的关系抽取模型与本文提出的基于多维语义映射的方法(MD-SMN)进行对比,如表 3 所列。

CNN+Softmax 模型。Zeng 等^[6]利用卷积深度神经网络提取词汇和句子级别的特征进行关系抽取,在 CLT 数据集上 F1 值达到了 47.6%。

CR-CNN 模型。Santos 等^[8]使用卷积神经网络来解决关系分类任务,他们从词向量和位置向量的角度出发,在 CLT 数据集上 F1 值达到了 54.1%。

SDP-LSTM 模型。Xu 等^[10]提出了一种新颖的神经网络来对句子中两个实体的关系进行分类,在 CLT 数据集上 F1 值达到了 55.3%。

BRCNN 模型。Cai 等^[11]专注于建模两个实体之间的最短依赖路径(SDP),利用卷积或递归神经网络进一步探索如何利用充分依赖关系,通过合并 SDP 中的信息,在 CLT 数据集上 F1 值达到了 55.6%。

表 3 MD-SMN 在 CLT 数据集上与其他模型的对比

Table 3 Comparison of MD-SMN with other models on CLT dataset

(单位:%)

Method	Features used	F1
CNN+softmax	Word embedding, position embedding, WordNet	47.6
CR-CNN	Word embedding, position embedding	54.1
SDP-LSTM	Word embedding, POS embedding, WordNet, grammar relation embedding	55.3
BRCNN	Word embedding, POS embedding, NER, WordNet	55.6
MD-SMN	Word embedding	75.3

表 4 列出了进行对比的模型(表 3 中已出现的模型不再赘述)。

CNN+Softmax 模型。Zeng 等^[6]提出的模型,在 Semeval-2010 Task8 数据集上 F1 值达到了 82.7%。

CR-CNN 模型。Santos 等^[8]提出的模型,在 Semeval-2010 Task8 数据集上 F1 值达到了 84.1%。

SDP-LSTM 模型。Xu 等^[10]提出的模型,在 Semeval-2010 Task8 数据集上 F1 值达到了 83.7%。

Bi-SDP-Att 模型。Wang 等^[20]提出了一种具有双最短路径依赖注意机制的模型,该模型通过形成一个平行的注意力权重来提升分类性能,在 Semeval-2010 Task8 数据集上 F1 值达到了 85.1%。

LST-AGCN 模型。Sun 等^[19]针对如何利用依赖树中的有效信息,提出了一种可学习的语法传输注意力图卷积网络,在 Semeval-2010 Task8 数据集上 F1 值达到了 86.0%。

BRCNN 模型。Xu 等^[10]提出的模型,在 Semeval-2010 Task8 数据集上 F1 值达到了 86.3%。

Multi-channel 模型。Chen 等^[12]提出了一种多通道深度神经网络的方法。该方法使用两个命名实体将一个句子分成 5 个通道,这些通道在递归传播过程中不会相互作用,从而使神经网络能够学习不同的表示形式,在 Semeval-2010 Task8 数据集上 F1 值达到了 84.5%。

Entity Attention Bi-LSTM 模型。Lee 等^[21]为了充分利用与实体有关的信息,将实体注意机制与潜在实体类型(Latent Entity typing,LET)方法相结合,将实体及其潜在类型作为特征,在 Semeval-2010 Task8 数据集上 F1 值达到了 85.2%。

Bi-LSTM-IE 模型。Lei 等^[18]在关系抽取方法中重点研究了词的形态特征。通过结合词的形态特征和语义特征,充实输入向量的表示,证明了该方法对多重关系抽取是有效的,但是该模型在处理重叠关系上没有明显优势。在 Semeval-2010 Task8 数据集上 F1 值达到了 85.3%。

Alt 等^[16]提出了 TRE 模型,TRE 没有使用显式的语言特征来进行关系分类,而是采用预先训练的深度语言表征和 Transformer 架构相结合,有效建立起实体之间的长期依赖关系。在 Semeval-2010 Task8 数据集上 F1 值达到 87.1%。

表 4 MD-SMN 在 Semeval-2010 Task8 数据集上与其他模型的对比
Table 4 Comparison of MD-SMN with other models on Semeval-2010

Task8 dataset		
(单位:%)		
Method	Features used	F1
CNN+softmax	Word embedding, position embedding, WordNet	82.7
SDP-LSTM	Word embedding, POS, WordNet, grammar relation embedding	83.7
CR-CNN	Word embedding, position embedding	84.1
Multi-channel	Word embedding with multi-channel	84.5
Bi-SDP-Att	Word embedding, position embedding,	85.1
Entity Attention Bi-LSTM	Latent Entity Typing	85.2
Bi-LSTM-IE	Word embedding	85.3
LST-AGCN	Word embedding	86.0
BRCNN	Word embedding, POS embedding, NER, WordNet	86.3
TRE	Wording embedding, position embedding-	87.1
MD-SMN	Word embedding	88.9

由表 2 和表 3 可知,本文提出的面向关系抽取的多维语义映射方法优于其他方法,在 Chinese Literature Text 数据集上 F1 值达到了 75.3%,在英文数据集 Semeval-2010 Task8 上 F1 值达到了 88.9%。实验证明,利用本文方法对句子的语义信息进行多维度角度捕获是有效的,能够提升关系抽取性能的效果。

综上所述,以及通过表 3、表 4 的实验结果证明,本文方法对句子的语义信息进行多维度捕获是有效的,对关系抽取任务是有帮助的。

结束语 本文针对大部分现有的深度学习模型在输入层只是简单地将文本语句变换为词向量,而没有对词向量进行深层次的探讨这一问题,提出了一种基于多维语义映射的

关系抽取方法。该方法的核心思想是将矩阵降维方法应用在神经网络模型输入层。通过将表示文本的词向量进行多维度的降维分解,使分解后的词向量能表示同一语句在不同维度上的语义信息,从而提升抽取效果。在 Chinese Literature Text 和 SemEval-2010 Task8 数据集上的实验结果表明,基于多维语义映射的方法在关系抽取任务上表现优秀,说明本文方法对于许多基于神经网络的语言处理模型而言都是有帮助的。

后续工作会在此基础上继续探究添加外部特征(如结构特征、词汇特征等)对关系抽取任务的影响。

参 考 文 献

[1] KAMBHATLA N. Combining lexical, syntactic, and semantic features with maximum entropy models for information extraction[C]// Proceedings of the ACL Interactive Poster and Demonstration Sessions. 2004:178-181.

[2] BUNESCU R, MOONEY R. A shortest path dependency kernel for relation extraction[C]// Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing. 2005:724-731.

[3] DEVLIN J, CHANG M W, LEE K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[J]. arXiv:1810.04805, 2018.

[4] XU J, WEN J, SUN X, et al. A discourse-level named entity recognition and relation extraction dataset for chinese literature text [J]. arXiv:1711.07010, 2017.

[5] HENDRICKX I, KIM S N, KOZAREVA Z, et al. Semeval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals[J]. arXiv:1911.10422, 2019.

[6] ZENG D, LIU K, LAI S, et al. Relation classification via convolutional deep neural network[C]// Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers. 2014:2335-2344.

[7] LIU C Y, SUN W B, CHAO W H, et al. Convolution neural network for relation extraction[C]// International Conference on Advanced Data Mining and Applications. Berlin:Springer, 2013: 231-242.

[8] SANTOS C N D, XIANG B, ZHOU B. Classifying relations by ranking with convolutional neural networks[J]. arXiv: 1504.06580, 2015.

[9] SOCHER R, HUVAL B, MANNING C D, et al. Semantic compositionality through recursive matrix-vector spaces[C]// Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. 2012:1201-1211.

[10] XU Y, MOU L, LI G, et al. Classifying relations via long short term memory networks along shortest dependency paths[C]// Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. 2015:1785-1794.

[11] CAI R, ZHANG X, WANG H. Bidirectional recurrent convolutional neural network for relation classification[C]// Proceedings of the 54th Annual Meeting of the Association for Compu-

tational Linguistics, 2016;756-765.

[12] CHEN Y P, WANG K, YANG W Z, et al. A Multi-Channel Deep Neural Network for Relation Extraction [J]. IEEE Access, 2020, 8: 13195-13203.

[13] VU N T, ADEL H, GUPTA P, et al. Combining recurrent and convolutional neural networks for relation classification[J]. arXiv, 1605. 07333, 2016.

[14] LIU J, YANG Y, LV S, et al. Attention-based BiGRU-CNN for Chinese question classification[J]. Journal of Ambient Intelligence and Humanized Computing, 2019, 1: 1-12.

[15] ZHANG K, LV G, WU L, et al. Image-enhanced multi-level sentence representation net for natural language inference[C] // 2018 IEEE International Conference on Data Mining (ICDM). IEEE, 2018;747-756.

[16] ALT C, HÜBNER M, HENNIG L. Improving relation extraction by pre-trained language representations[J]. arXiv: 1906. 03088, 2019.

[17] GUO Z, ZHANG Y, LU W. Attention guided graph convolutional networks for relation extraction[J]. arXiv: 1906. 07510, 2019.

[18] LEI M, HUANG H, FENG C, et al. An input information enhanced model for relation extraction[J]. Neural Computing and Applications, 2019, 31(12): 9113-9126.

[19] SUN K, ZHANG R, MAO Y, et al. Relation extraction with

convolutional network over learnable syntax-transport graph [C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2020; 8928-8935.

[20] WANG H, QIN K, LU G, et al. Direction-sensitive relation extraction using Bi-SDP attention model [J]. Knowledge-Based Systems, 2020, 198: 105928.

[21] LEE J, SEO S, CHOI Y S. Semantic relation classification via bi-directional lstm networks with entity-aware attention using latent entity typing[J]. Symmetry, 2019, 11(6): 785.



CHENG Hua-ling, born in 1996, post-graduate. Her main research interests include natural language processing and relation extraction.



QIN Yong-bin, born in 1980, Ph.D, professor. His main research interests include big data governance and application, and multi-source data fusion.

(责任编辑:何杨)