# 融合预训练语言模型和标签依赖知识的关系抽取方法

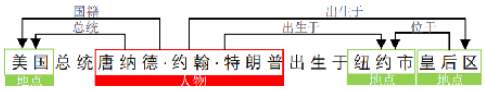

## 湖南数定智能科技有限公司

赵超 谢松昱 郑菲 程琛 彭立宏 曾道建†


## 引言

关系抽取旨在从未经标注的自由文本中抽取实体间的关系。然而，现有的方法大都孤立的预测每一个关系而未考虑关系标签相互之间的丰富语义关联。

关系标签间的联系称之为关系标签依赖，大致可分为两类：弱标签依赖和强标签依赖。弱标签依赖主要与关系标签的层次结构相关，而强标签依赖指关系标签之间具有潜在的逻辑推理蕴涵，以下图为例




该图介绍了美国第 任总统特朗普的人物背景信息。首先，因此与个人属性 出生于 总统 相关的存在弱标签依赖的关系标签应该有更大概率在这句话中出现。其次，实体对特朗普和美国的两种关系总统和国籍存在强标签依赖。

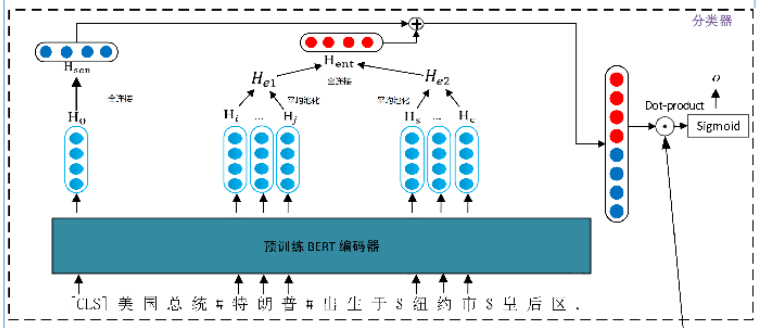
最近，预训练语言模型 对 的各种领域的发展产生了深远的影响。如何将预训练模型 更好的应用于关系抽取任务是目前的研究热点。

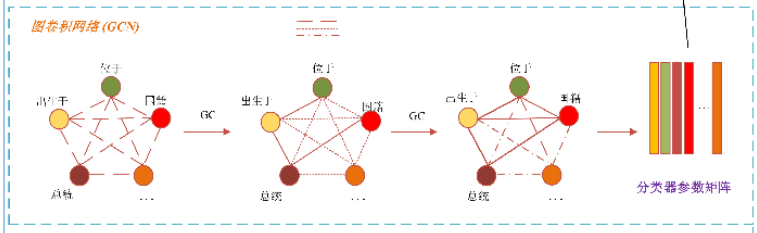
综上，本文提出了一种融合预训练语言模型和标签依赖知识的关系抽取模型。

## 方法

模型的框架如图所示，由三部分组成：

句子和实体编码器，对输入的文本进行编码，得到句子以及实体的上下文语义表征。 图卷积神经网络，对关系标签依赖图进行建模。 关系分类器，完成最终的关系预测任务。




首先，使用 分词器对输入句子进行分词。为了让 对句子进行编码的同时更好的感知目标实体对的信息，在实体的开头和结尾插入一个特殊标记。

利用预训练语言模型获得给定句子的实体表示及句子聚合表示。

$$\boldsymbol{H}_{sen} = f(\boldsymbol{W}_1\boldsymbol{h}_0 + \boldsymbol{b}_1)$$

$$H_{e1} = \frac{\sum_{t=i}^{j} h_t}{j-i+1}$$

$$H_{e2} = \frac{\sum_{t=s}^{n} h_t}{e-s+1}$$

$$\boldsymbol{H}_{ent} = f(\boldsymbol{W}_2[H_{e1}; H_{e2}] + \boldsymbol{b}_2)$$

文通过计算训练语料库中标签的共现情况来构造邻接矩阵建立标签依赖图，并利用图卷积网络对标签依赖建模。

$$\mathbf{V}^{l+1} = f(\mathbf{V}^l, \tilde{\mathbf{A}}) = \delta(\tilde{D}^{-\frac{1}{2}}\tilde{\mathbf{A}}\tilde{D}^{-\frac{1}{2}}\mathbf{V}^l\mathbf{W}^l)$$

综合上述进行最终的关系预测。

$$p_r(r|s,e) = \sigma(\boldsymbol{R}f(\boldsymbol{W}_3[\boldsymbol{H}_{sen}; \boldsymbol{H}_{ent}] + \boldsymbol{b}_3))$$

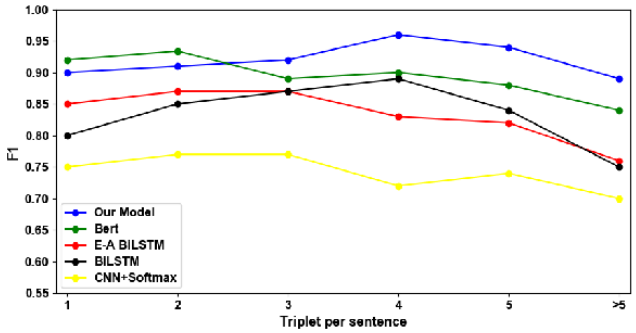
## 实验

为了验证本文提出模型在真实应用场景下的性能，我们采用真实案件受理过程中产生的警情文本作为实验数据。

**表 2.** 警情关系数据集实验结果.

| 模型 | 准确率 | 召回率 | F1 |
|---|---|---|---|
| CNN+Softmax | 79.9% | 69.6% | 74.4% |
| BiLSTM | 85.6% | 80.2% | 82.8% |
| EA-BiLSTM | 86.0% | 84.1% | 85.1% |
| Bert | **92.7%** | 90.4% | 91.5% |
| **Our Model** | 92.4% | **94.7%** | **93.5%** |

此外，为验证一个句子中存在多个三元组时，标签依赖知识可以有效提高关系抽取的性能，我们将测试集按照句子中包含三元组的数量划分为 个子集 ）。




## 结论

本文提出了一种结合预训练语言模型 和图卷积网络的关系抽取模型。相较于传统模型，该模型能更好的捕获上下文语义信息并充分利用和挖掘关系标签之间的相关性。

## 参考文献