计算机工程与应用
Computer Engineering and Applications
ISSN 1002-8331,CN 11-2127/TP

《计算机工程与应用》网络首发论文

题目：基于 BILSTM-CRF 的高校政策语义角色标注研究
作者：徐建国，刘泳慧，刘梦凡
网络首发日期：2020-10-16
引用格式：徐建国,刘泳慧,刘梦凡.基于 BILSTM-CRF 的高校政策语义角色标注研究.计算机工程与应用.
https://kns.cnki.net/kcms/detail/11.2127.TP.20201016.1314.006.html

基于 BILSTM-CRF 的高校政策语义角色标注研究

徐建国，刘泳慧，刘梦凡
山东科技大学 计算机科学与工程学院，山东 青岛 266590

摘　要：采用融合自注意力机制的双向长短期记忆模型(SelfAtt-BILSTM)和条件随机场模型(CRF)，构建一种 SelfAtt-BILSTM-CRF 模型，对政策文本进行语义角色标注，以提取政策主要内容。采用某高校政策文件为实验数据集，利用 BILSTM 模型自动学习序列化语句上下文特征，融合自注意力机制增加重要特征元素的权重，最后通过 CRF 层利用特征进行序列标注，提取语义角色，以实现政策文件的主要内容挖掘。经过对比验证，该模型能够有效的提取政策文本内容，在标注数据集上 F1 值达到 78.99%。实验结果同时表明，自注意力机制能够有效提高神经网络模型的语义角色标注效果。

关键词：双向长短期记忆网络；条件随机场；自注意力机制；语义角色标注；深度学习

文献标志码：A　　**中图分类号：**TP391.1　　doi：10.3778/j.issn.1002-8331.2004-0146

徐建国，刘泳慧，刘梦凡.基于 BILSTM-CRF 的高校政策语义角色标注研究.计算机工程与应用
XU Jianguo, LIU Yonghui, LIU Mengfan. Research on the semantic role labeling of university policy based on BILSTM-CRF. Computer Engineering and Applications

Research on the semantic role labeling of university policy based on BILSTM-CRF

XU Jianguo, LIU Yonghui, LIU Mengfan
College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao, Shandong, 266590, China

Abstract: A SelfAtt-BILSTM-CRF model was constructed by combining the Self Attention Mechanism with Bidirectional Long Short-Term Memory (SelfAtt-BILSTM) and Conditional Random Field (CRF).The semantic role of policy text was annotated to extract the main content of policy. This paper took a college policy file as the experimental data set, used the BILSTM model to automatically learn the context features of serialized statements, integrated the self attention mechanism to increase the weight of important feature elements, and finally used the features to sequence annotation through CRF layer to extract the semantic roles , so as to realize the main content mining of policy files. By contrast, the SelfAtt-BILSTM-CRF model can effectively extract the content of policy text, and the F1 value in the annotation data set reaches 78.99%. The experimental results also show that the Self Attention Mechanism can effectively improve the semantic role annotation effect of the neural network model.

Key words: bidirectional long short-term memory; conditional random field; self attention mechanism; semantic role labeling; deep learning

基金项目：2017 年山东省高校人文社会科学研究计划(思想政治教育专题研究)资助经费项目(No.J17ZZ27)；青岛市哲学社会科学规划项目(No.QDSKL1601121)；山东科技大学 2018 年研究生科技创新项目(No.SDKDYC180339)。

作者简介：徐建国(1964-),男,硕士,副教授、硕士生导师,主要研究领域为为智能数据分析与处理、网络舆情分析；刘泳慧(1996-),女,硕士研究生,主要研究领域为自然语言处理，E-mail:sdkjlyh@163.com；刘梦凡(1996-),女,硕士研究生,主要研究领域为网络舆情分析。

1 引言

自然语言处理分析技术大致分为三个层面：词法分析、句法分析和语义分析。其中语义分析是自然语言处理研究的关键问题，其目标是理解句子表达的真实语义[1]。而语义角色标注其实质上是在句子级别进行浅层的语义分析，按照现代语法知识将词语序列分组，并按照语义角色对它们进行分类，该方法不对整个句子做详细的语义分析，只标注句子中给定谓词(动词、名词等)的语义角色(参数)，使计算机对语句有一个“浅层”的理解[2]。谓词的语义角色有如“施事者”、“受事者”、“时间”、“地点”、“条件”等。语义角色标注可用于问答系统[3]、机器翻译[4]、信息检索[5]、指代消解[6]等自然语言处理领域。广义而言，语义角色标注能够辅助计算机对自然语言进行语义理解，促进自然语言处理技术的发展。

广泛的划分，语义角色标注任务有基于规则、基于统计机器学习、深度学习三种。基于规则的语义角色标注受语料库、知识库、词典等限制，适用于处理小规模数据[7]，在处理小规模数据时具备成本低、可扩展性高的优点[8]，而大规模语料库的构建需要花费大量人力物力，可操作性差。基于统计机器学习的方法可以避免构建规则，主要有最大熵模型(Maximum Entropy，ME)[9]、条件随机场(Conditional Random Field，CRF)[10]、以及支持向量机(Support Vector Machine，SVM)[11][12]等方法。但以上方法特征提取仍需要人工参与，也存在特征稀疏、维数灾难等问题。深度学习方法可以避免人工提取特征而直接对原始数据进行处理[13]，近年来研究学者将其广泛的应用到语义角色标注任务中。张苗苗等[14]引入 Gate 机制对词向量进行调整，采用深度 Bi-LSTM-CRF 模型对 CPB 数据集进行语义角色标注,Gate 机制和 LSTM 模型的门控单元进行互补，有效的提高了标注的准确率。王旭阳等[15]采用语义聚类进行数据预处理以解决稀疏谓词对标注的影响，对词向量进行“模糊化”处理来提升词向量与谓词的相关性，利用 Bi-LSTM 模型和 CRF 模型对 CPB 数据集进行序列标注，F 值最高可达到 81.24%。王明轩等[16]在多层 LSTM 模型中装置新颖的“直梯单元”进行语义角色标注，直梯单元可以减少信息损失，让信息在空间和时间维度上更通畅的传播，更好的发挥深度 LSTM 模型的作用，在 CoNLL-2005、CoNLL-2012 以及领域外数据集上都取得了较好的结果。陈艳平等[17]提出 Bi-LSTM-Attention-CRF 模型对裁判文书进行句法要素的识别，实验证明了 Attention 机制能够有效识别各句法要素的关联性，深度神经网络模型能提高句法要素的识别效率。朱晓霞等[18]采用 BILSTM 和 CRF 模型对微博评论进行语义角色标注以筛选无关语义角色，提取特定语义角色形成情感单元词表，以便后续进行动态主题下的情感演化分析。

本文采用融合自注意力机制的 BILSTM 模型自动学习政策文本上下文特征，CRF 模型实现序列标注，通过对政策文本进行语义角色标注以获取政策的主要内容。BILSTM 模型能够自动学习政策文本的上下文特征，无须人工选取特征，政策文件中长句居多，加入自注意力机制能够更好的捕获特征内部联系，提高了对政策文件中的长句语义角色标注的准确率。采用 CRF 模型进行序列标注也减少了无效标签的数量，如标签 I 应在标签 B 之后，而不是单独存在或者在标签 O 之后。

2 相关知识介绍

2.1 长短期记忆模型

长短期记忆模型(Long Short-Term Memory，LSTM)是循环神经网络(Recurrent Neural Network，RNN)的一种重要变种,其独特的门结构使其保留了 RNN 的处理序列数据的能力,又解决了序列学习过程存在的长期依赖问题，避免梯度消失和梯度爆炸现象。LSTM 由记忆单元、输入门(Input Gate)、遗忘门(Forget Gate)和输出门(Output Gate)四个主要元素组成。其单元结构图如图 1 所示，计算过程如式(1)所示。

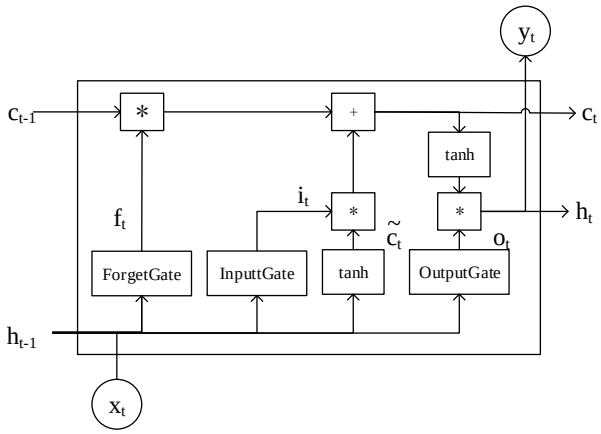

图 1　LSTM 模型单元结构图

Fig.1　LSTM model unit structure

$$\begin{cases} f_t = \sigma\left(W_f\left[h_{t-1}, x_t\right] + b_f\right) \\ i_t = \sigma\left(W_i\left[h_{t-1}, x_t\right] + b_i\right) \\ \tilde{c}_t = \tanh\left(W_c\left[h_{t-1}, x_t\right] + b_c\right) \\ c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \\ o_t = \sigma\left(W_o\left[h_{t-1}, h_t\right] + b_o\right) \\ h_t = o_t * \tanh\left(c_t\right) \end{cases} \tag{1}$$

其中，σ 是激活函数 sigmoid 函数，W 是矩阵乘法操作，b 是函数的偏置项，*表示点乘操作，tanh 为 tanh 函数，f_t、i_t、c_t、o_t、h_t 分别为遗忘门、输入门、记忆细胞、输出门、隐藏状态的输出向量。

2.2　注意力机制

注意力机制(Attention Mechanism)最早是在视觉图像领域中提出，本质来自于人类视觉注意力机制。Attention 函数可以描述为一个源数据 d_b 到目标数据 d_e 的映射，在计算 Attention 时主要分为三步，首先将源数据 d_b 和目标数据 d_e 进行相似度计算得到权重，常用的相似度函数有点积、矩阵拼接和感知机等，如式(2)所示；其次使用一个 softmax 函数对这些权重进行归一化，如式(3)所示；最后将权重和相应的词向量进行加权得到最后的注意力机制处理结果。

$$f\left(d_b, d_s\right) = \begin{cases} d_b{}^T d_e \\ W_a\left[d_b, d_e\right] \\ v_a^T \tanh\left(W_a d_b + U_a d_e\right) \end{cases} \tag{2}$$

$$a_b = soft\ \max\left(f\left(d_b, d_e\right)\right) = \frac{\exp\left(f\left(d_b, d_e\right)\right)}{\sum_{e \in E} \exp\left(f\left(d_b, d_e\right)\right)} \tag{3}$$

2.3　条件随机场

条件随机场最早由 Lafferty 等人于 2001 年提出，其模型思想的主要来源是隐马尔可夫模型，它是一种用来标记和切分序列化数据的统计模型[19]。条件随机场为判别式概率无向图，其联合概率分布可以进行因子分解，对于给定观察序列 $x\left(x_1, \cdots x_i\right)$ 的条件下，相应的标记序列 $y\left(y_1, \cdots y_i\right)$ 的概率[20]，即条件随机场模型为：

$$P\left(y \mid x\right) = \frac{1}{Z\left(x\right)} \exp\left(\sum_i \sum_k \lambda_k t_k\left(y_{i-1}, y_i, x, i\right) + \sum_i \sum_k \mu_k s_k\left(y_i, x, i\right)\right) \tag{4}$$

其中，$Z\left(x\right)$ 为归一化因子，$t_k\left(y_{i-1}, y_i, x, i\right)$ 是整个观察序列和对应标记在 *i-1* 和 *i* 时刻的特征，为转移特征函数，$s_k\left(y_i, x, i\right)$ 为在 i 时刻整个观察序列和标记的特征，为状态特征函数，λ_k 和 μ_k 分别为训练得到的转移特征和状态特征的权重。最后用 Viterbi 算法进行解码，计算条件概率最大的输出序列。

3　模型建立

为提取长篇幅（大于 3 页）政策文件中的主要内容，利用高校政策文件的特点，本文构建一种融合自注意力机制的 BILSTM-CRF 模型，即 SelfAtt-BILSTM-CRF 模型，模型结构如图 2 所示。采用模型对高校政策进行语义角色标注的方式，提取政策文件的主要内容，其理论基础是利用 BILSTM 生成包含上下文特征的词向量，将其与注意力层获得的权重进行加权，然后输入到 CRF 中实现序列标注。本文在 BILSTM-CRF 模型中引入自注意力机制，能有效的获取特定词语在文本中的长距离依赖，能够更准确的捕捉到文本特征的内部相关性，提高语义角色标注的准确率。

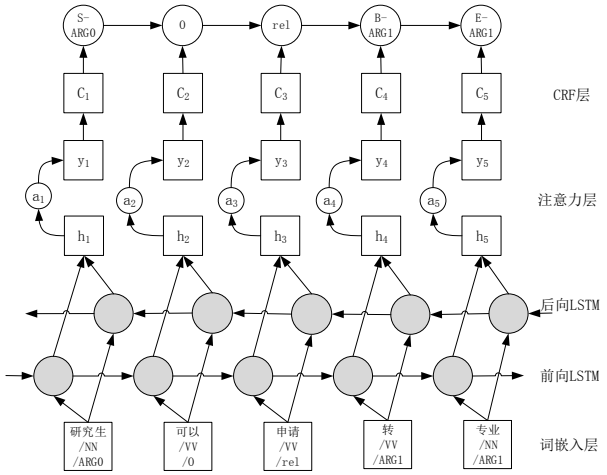

图 2　SelfAtt-BILSTM-CRF 模型结构

Fig.2　SelfAtt-BILSTM-CRF model structure

3.1　特征向量表示

本文除了使用词特征之外，还使用了词性和语义角色特征，将输入的政策文本序列转化为特征向量表示输入到词嵌入层。词性是词语的基本语法属性，通常也称为词类。语句中的词性特征也蕴含着信息，结合词语的位置，词性特征能表示词语的用途和功能[21]。例如，文本序列“特制定本规定”中，“规定”本身有名词和动词词性，其位于句末标注为名词，表示谓词“制定”的受事方，可以看出词性对语义角色标注具有重要参考价值。语义角色标注是以谓词为中心，分析句子中各成分与谓词的关系，语义角色有动作的施事方、受事方、目的原因等，其能表示浅层语义信息。设一个长度为 n 的句子 S，则第 i 个词 w_i 的词向量为 e_i，词性向量为 p_i，语义角色向量为 r_i。将其拼接得到句子的特征向量为 $[e_1,\cdots e_n]\oplus[p_1,\cdots p_n]\oplus[r_1,\cdots r_n]\in R^{n\times(d+l+t)}$，其中 d，l，t 分别为词向量，词性向量和语义角色向量的维度。拼接得到的特征向量输入词嵌入层(Word Embedding)，将高维的特征向量训练为语义信息更丰富的低维向量作为BILSTM模型的输入。

3.2　融合自注意力机制的 BILSTM-CRF 模型

标准LSTM模型理论上只能按照文本序列的输入处理上文的信息，而下文的信息对于序列文本的处理也有重要意义。BILSTM 由两层 LSTM 模型组成，词嵌入层得到的低维词向量从左到右按序作为正向LSTM的输入，每一次的输出都会结合后一个词向量开始下一轮的预测，以此类推，得到正向输出 $h_1,h_2,\cdots,h_t$；反向 LSTM 的过程与正向 LSTM 正好相反，得到反向输出 $\overleftarrow{h_t},\overleftarrow{h_{t-1}},\cdots,\overleftarrow{h_1}$，将两层LSTM的输出进行串联得到包含上下文的特征 h_t，如式(5)所示。这种结构使得 BILSTM 可以充分学习输入序列数据中文本的上下文信息。

$$h_t=\left[\overrightarrow{h_t},\overleftarrow{h_t}\right] \tag{5}$$

自注意力机制(Self Attention Mechanism)是注意力机制的一种特殊形式，其源数据和目标数据相同，计算同一个样本数据中每个元素的重要程度，获得样本数据内部的联系[22]。本文采用的相似度函数为点积，由式(5)中包含上下文的特征 h_t 计算权重 a_t，如式(6)所示，将特征 h_t 与对应的权重 a_t 相乘，最终得到注意力层的输出特征向量 y_t，如式(7)所示。

$$a_t=\frac{\exp\left(h_t^{T}h_t\right)}{\sum_{t=1}^{n}\exp\left(h_t^{T}h_t\right)} \tag{6}$$

$$y_t=a_th_t \tag{7}$$

采用融合自注意力机制的BILSTM模型进行特征提取，对于政策文件中无法拆分的长句，BILSTM 模型可以充分利用上下文信息，获得长距离依赖，自注意力机制可以计算增加重要元素的权重，精简元素较多的语义角色。最后采用 CRF 以注意力层输出的特征向量作为输入，以标注序列为监督信号，完成序列标注。一个句子的预测特征即为一个标注序列，可以看作 CRF 的一个特征函数，特征函数的输出值越接近 1 就表示该标注序列越符合特征，越接近 0 则表示越不符合。融合自注意力机制的BILSTM模型可以有效的获取上下文特征，但是在标注过程中无法使用特征依赖信

息，会出现大量非法标注的情况。例如复合词正确的标签顺序为“BIE”，会出现“IIE”等非法标注结果。CRF 模型计算条件概率最大的输出序列，能够通过特征函数得出合理的标注结果。

对政策文件内容进行语义角色标注时，例如，政策文件中句子“为进一步加强我校本科教学工作，提高本科教学质量，把 xxxx 作为一项基本制度，现结合我校实际，特制定本规定。”此文本序列中谓词为“制定”，其目的状语较长，CNN 倾向于获取局部静态信息，而 RNN 无法解决长期依赖问题，其都不能得到完整的目的状语。采用融合自注意力机制的 BILSTM 模型不仅能够捕捉到动态的时序信息，还能结合上下文解决长期依赖的问题，最终通过 CRF 模型完成序列标注，进而可以得到较为完整的标注正确的语义角色。

3.3 SelfAtt-BILSTM-CRF 模型在高校政策语义角色标注中的应用

高校政策是高校发布的对高校人、财、物等进行组织和管理的一系列文件，获取高校政策的主要内容对后续进行政策作用分析研究具有重要意义。对于 3 页以内的政策文件，可以快速的获取政策主要内容，而大于 3 页的政策文件中信息较多，无法快速清晰的获取政策中的主要内容。因此，利用模型对长篇幅（大于 3 页）政策文件进行语义角色标注提取，有利于把握长篇幅政策文件中的主要内容，具有较大的现实意义。

本文将 SelfAtt-BILSTM-CRF 模型应用于高校政策文本的语义角色标注，具体流程图如图 3 所示。首先，爬取某高校的政策文件，统一文件格式并对文件内容进行初步清洗；然后，将政策文本进行切分长句、分词、去除停用词、词性标注等形成以句子为单位的文本序列，通过 SelfAtt-BILSTM-CRF 模型计算出每个句子对应的标签序列；最后，根据标签序列结合文本找出对应的语义角色并进行提取。

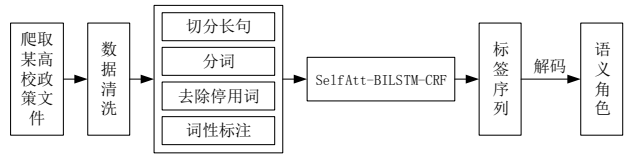

图 3　政策文件语义角色标注流程图

Fig.3　Policy documents semantic role labeling flowchart

4　实验结果分析

为验证本文提出的模型对高校政策文件内容的挖掘效果，通过互联网爬取某高校办公系统中的部分政策文件进行实验与对比。

4.1　数据收集与预处理

本次实验爬取某高校 2016 年 1 月 1 日至 2019 年 12 月 31 日在办公系统中发布的政策文件共 140 篇，除去内容少于 3 页的文件，余下 53 篇长篇幅的政策文件进行实验验证。爬取得到的政策文件为 PDF 和 Word 格式，首先将文件转换为文本文件，并对文件内容进行初步清洗，包括去除文件红头标题、页码以及带有表格的附件等。

政策文件中长句居多，其中由一些短的并列句组成的长句，用模型进行角色标注时，容易漏标一些语义角色，因此将文件内容进行分句时，以“。；：？！”作为分句的标记符号将由并列句组成的长句切分为短句。同时政策文件有固定的格式，章节条目分明，其中“第一章”、“第一条”等字样会对模型产生一定干扰，因此在去除停用词时将章节条目词去掉。

最终得到有效数据 4341 条，实验采用 k 折交叉验证方式训练，将数据集等比例划分为 4 份，每次 3 份作为训练集，1 份作为测试集，采用各评判指标的平均数作为最终结果。通过在 python 中安装 NLTK 库对文本进行分词和词性标注处理。

4.2　数据集标注

在正式训练模型之前需要对文本进行语义角色标注，要提取政策文件中比较全面的内容则需要标注多种类型的语义角色，本文参考了 Chinese Proposition Bank(CPB)语料的标注方法，对数据集进行标注的语义角色包括 5 类核心语义角色

ARG0-ARG4，12 种附加语义角色，其中部分释义如表 1 所示，为了更好的识别语义角色的边界、对复合词进行标注，我们采用 IOBES 标注策略，部分结果示例如表 2 所示，B 表示复合词词首，I 表示复合词词中，E 表示复合词词尾，S 表示简单词，O 表示语义角色以外的其他词。

表 1　部分语义角色释义表

Table 1　Part of the semantic role definition table

语义角色类型	说明	语义角色类型	说明
ARG0	动作的施事方	ARGM-TPC	Topic(主题)
ARG1	动作的受事方	ARGM-CND	Condition (条件)
ARGM-DIR	Direction (方向)	ARGM-FRQ	Frequency (频率)
ARGM-MNR	Manner (方式)	ARGM-TMP	Temporal (时间)
ARGM-LOC	Locative (地点)	ARGM-PRP	Purpose or reason (目的或原因)

表 2　部分 IOBES 标注结果示例

Table 2　Partial IOBES annotation result examples

词	角色标注	词	角色标注
评审	B-ARG0	办公室	S-ARG0
领导	I-ARG0	设在	Rel
小组	E-ARG0	研究生	B-ARG1
下设	rel	院部	E-ARG1
办公室	S-ARG1	。	O

4.3　模型性能分析

本文由 Tensorflow 进行模型搭建，实验参数对整个实验有较大影响，决定了实验结果的准确度，通过固定一个参数同时更改其他参数的方法，反复对比试验，最终得到最佳的实验参数。为防止训练数据发生过拟合，dropout 值设为 0.5。数据集的 batch_size 为 32，词嵌入维度为 300，词性嵌入维度为 20，语义角色嵌入维度为 30，训练学习率为 0.001，优化器采用 Adam，模型训练轮数为 50，保存在测试集上 F1 最高的模型参数。实验结果采用准确率，召回率，F1 值作为评判指标，本文采用 CRF、BILSTM、BILSTM-CRF、SelfAtt-BILSTM-CRF 四种模型对政策文本进行语义角色标注，对比结果如表 3 所示。

表 3　各模型语义角色标注结果对比（%）

Table 3　The model results of semantic role labeling（%）

模型	准确率	召回率	F1 值
CRF	68.69	65.82	67.22
BILSTM	72.27	70.37	71.31
BILSTM-CRF	78.76	75.69	77.20
SelfAtt-BILSTM-CRF	81.55	76.59	78.99

由表 3 可以看出，使用单一模型的实验结果总体低于组合模型的实验结果，其中 CRF 模型依赖于人工设计的特征模板，模板的优劣对模型实验结果有较大影响，而 BILSTM 模型无需人工设计特征，其独特的门结构可以存储历史信息，有效地捕获上下文特征，但 BILSTM 模型进行标注时无法利用上下文依赖信息，会标注大量无效标签。第三组实验中，将两个单一的模型组合成 BILSTM-CRF 模型，利用 BILSTM 模型自动学习上下文特征，而 CRF 是概率结构化模型，在 BILSTM 模型后接入 CRF 模型可以有效的弥补 BILSTM 模型标注偏执，不能全局归一的问题。第四组实验中，在 BILSTM-CRF 模型中融合了自注意力机制，自注意力机制通过增加重要特征元素的权重，能够更好的捕获文本序列内部的联系，获得长距离依赖关系。因此，在四种模型中，数据集的 F1 值呈现依次递增的情况。对比 CRF 模型，SelfAtt-BILSTM-CRF 模型的 F1 值提高了 11.77%。实验证明，本文提出的模型能够提高对政策文本进行语义角色标注的准确率。

5 结束语

本文将深度神经网络模型应用到高校政策文件的语义角色标注中，构建 SelfAtt-BILSTM-CRF 模型提取高校政策文件的主要内容。将词、词性和语义角色特征向量化进行拼接输入到 BILSTM 模型，加入自注意力机制赋予不同特征向量相应的权重，最后利用 CRF 模型完成序列标注。对比 CRF、BILSTM、BILSTM-CRF 和 SelfAtt-BILSTM-CRF 四种模型在高校政策文件数据集上的实验结果，验证了 SelfAtt-BILSTM-CRF 模型的可操作性和有效性。但由于缺乏规范、完善的高校政策文件语料库，人工对高校政策文件进行标注的语义角色存在不可避免的误差，现阶段使用的数据集规模也较小，无法实现模型的高准确率。对文件内容进行语义角色标注时选取标注的语义角色数量较多，也影响了模型的准确率，在尽可能多的提取政策文件内容的前提下，扩展标注数据集和权衡语义角色数量以提高模型的准确率是下一步的研究方向。

参考文献：

[1] 中国中文信息学会.中文信息处理发展报告(2016)[R].北京:中国中文信息学会,2016.

Chinese Information Society of China. Chinese information processing development report (2016)[R]. Beijing: Chinese Information Society of China, 2016.

[2] 祝娜,王效岳,白如江.语义角色标注及其在科技情报分析中的应用研究[J].情报理论与实践,2015,38(1):98-103.

ZHU N,WANG X Y,BAI R J. Semantic role labeling and applied research in science and technology Intelligence Analysis[J]. Information Studies: Theory & Application, 2015,38(1):98-103.

[3] Narayanan S, Harabagiu S. Question answer based on semantic structures[C]//Proceedings of the 20th International Conference on Computational Linguistics. Geneva: Association for Computational Linguistics, August 1, 2004: 1-10.

[4] Wu Dekai,Fung P.Can semantic role labeling improve SMT[C]//Proceedings of the 13th Annual Conference of the European Association for Machine Translation. Barcelona: European Association for Machine Translation, 2009: 218-225.

[5] Surdeanu M, Harabgiu S, Willams J, et al. Using predicate-argument structures for information extraction[C]//Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics. Sapporo: Association for Computational Linguistics,2003:8-15.

[6] Kong Fang,Zhou Guodong,Zhu Qiaoming.Employing the centering theory in pronoun resolution from the semantic perspective[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Suntec: Association for Computational Linguistics, 2009: 987-996.

[7] 吴林静,劳传媛,范佳林等.一种规则与统计相结合的应用题句子语义角色识别方法[J].计算机应用研究，2018,35(8): 2299-2303.

Wu Linjing,Lao Chuanyuan,Fan Guilin, et al. Hybrid method based on rules and statistics for semantic role annotation in math word problems[J]. Application Research of Computers,2018,35(8):2299-2303.

[8] 黄炜,黄建桥,李岳峰.基于BiLSTM-CRF的涉恐信息实体识别模型研究[J/OL].情报杂志. 2019, 38(12): 149-156 [2019-09-12]http://kns.cnki.net/kcms/detail/61.1167.g3.20190910.1730.010.html.

Huang Wei, Huang Jianqiao, Li Yuefeng. Research on entity identification model of terrorism-related information based on BiLSTM-CRF[J/OL]. Journal of Intelligence.2019,38(12):149-156[2019-09-12]http://kns.cnki.net/kcms/detail/61.1167.g3.20190910.1730.010.html.

[9] 于江德，王希杰，余正涛. 基于最大熵模型的语义角色标注[J].微电子学与计算机,2010,27(8):173-176+180.

YU Jiangde, WANG Xijie, YU Zhengtao. Semantic role labeling based on maximum entropy model[J]. Microelectronics & Computer, 2010,27(8):173-176+180.

[10] 王荣洋,鞠久朋,李寿山等.基于CRFs的评价对象抽取特征研究[J].中文信息学报,2012,26(2):56-61.

WANG Rongyang,JU Jiupeng,LI Shoushan,et al.Feature engineering CRFs based opinion target extraction[J]. Journal of Chinese Information Processing, 2012, 26(2): 56-61.

[11] 王步康,王红玲,周国栋.基于树核函数的中文语义角色标注[J].计算机工程,2011,37(22):128-130.
WANG Bukang, WANG Hongling, ZHOU Guodong. Semantic role labeling in chinese language based on tree kernel function[J]. Computer Engineering, 2011, 37(22): 128-130.

[12] 毛小丽, 何中市, 邢欣来, 等. 基于语义角色的实体关系抽取[J].计算机工程,2011,37(17):143-145.
MAO Xiaoli,HE Zhongshi,XING Xinlai,et al. Entity relation extraction based on semantic role[J]. Computer Engineering, 2011,37(17):143-145.

[13] BENGIO Y, SCHWENK H, SENCALJ S, et al. Neural probabilistic language models[J]. Journal of Machine Learning Research,2001,3(6):1137-1155.

[14] 张苗苗,刘明童,张玉洁等.融合Gate过滤机制与深度Bi-LSTM-CRF的汉语语义角色标注[J].情报工程, 2018, 4(2): 45-53.
ZHANG Miaomiao,LIU Mingtong,ZHANG Yujie,et al. The integration of gated filtering mechanism and deep Bi-LSTM-CRF for chinese semantic role labeling[J]. Technology Intelligence Engineering, 2018,4(2):45-53.

[15] 王旭阳,朱鹏飞.基于模糊机制和语义密度聚类的汉语自动语义角色标注研究[J].计算机应用与软件, 2019, 36(9): 76-82+92.
Wang Xuyang,Zhu Pengfei. Chinese automatic semantic role labeling based on fuzzy mechanism and semantic density clustering[J]. Computer Applications and Software, 2019, 36(9):76-82+92.

[16] 王明轩,刘群.基于深度神经网络的语义角色标注[J].中文信息报,2018,32(2):50-57.
WANG Mingxuan,LIU Qun.A simple and effective deep model for semantic role labeling[J]. Journal of Chinese Information Processing, 2018,32(2):50-57.

[17] 陈艳平,冯丽,秦永彬,等.一种基于深度神经网络的句法要素识别方法[J/OL].山东大学学报(工学版), 2019, 50(2): 1-6[2020-03-12].http://kns.cnki.net/kcms/detail/37.1391.T.20200311.0858.002.html.
CHEN Yanping,FENG Li,QIN Yongbin,et al. A syntactic element recognition method based on deep neural network[J/OL]. Journal Of Shandong University (Engineering Science), 2019, 50(2): 1-6 [2020-03-12]. http://kns.cnki.net/kcms/detail/37.1391.T.20200311.0858.002.html.

[18] 朱晓霞,宋嘉欣,孟建芳.基于主题——情感挖掘模型的微博评论情感分类研究[J].情报理论与实践, 2019, 42(5): 159-164.
ZHU Xiaoxia,SONG Jiaxin,MENG Jianfang. Research on the classification of emotion in microblog comments based on the theme-emotion mining model[J]. Information Studies: Theory & Application, 2019, 42(5): 159-164.

[19] Lafferty J D, McCallum A, Pereira F C N. Conditional random fields: probabilistic models for segmenting and labeling sequence data[C]//Proceedings of the Eighteenth International Conference on Machine Learning, 2001: 282-289.

[20] 邬伦,刘磊,李浩然等.基于条件随机场的中文地名识别方法[J].武汉大学学报(信息科学版),2017,42(02):150-156.
WU Lun,LIU Lei, LI Haoran,et al.A chinese toponym recognition method based on conditional random field[J].Geomatics and Information Science of Wuhan University, 2017,42(02):150-156.

[21] 李卫疆,李涛,漆芳.基于多特征自注意力BLSTM的中文实体关系抽取[J].中文信息学报,2019,23(10):47-56+72.
LI Weijiang, LI Tao, QI Fang. Chinese entity relation extraction based on multi-features Self-Attention Bi-LSTM[J]. Journal of Chinese Information Processing, 2019,23(10):47-56+72.

[22] 杨善良,孙启.基于注意力机制的循环神经网络评价对象抽取模型[J].计算机应用与软件,2019,36(03):202-209.
Yang Shanliang, Sun Qi. Evaluation object extraction model of recurrent neural network based on attention mechanism[J]. Computer Applications and Software, 2019, 36(03): 202-209.