

刘洋,唐海,朱梦涵,等. 基于视频资源与 WoBERT-AT-BiLSTM-CRF 的命名实体识别方法[J]. 智能计算机与应用,2024, 14(10):63-69. DOI:10.20169/j.issn.2095-2163.241008

# 基于视频资源与 WoBERT-AT-BiLSTM-CRF 的命名实体识别方法

刘洋,唐海,朱梦涵,徐洪胜

(湖北汽车工业学院 电气与信息工程学院,湖北 十堰 442002)

**摘要:** 针对教育领域命名实体识别数据集的缺乏,提出利用视频资源构建相应的学科数据集。传统的语音识别模型存在词错率高、难以处理长时序序列等情况,提出使用端到端的语音识别模型 Whisper。对于实体识别存在误差积累、实体多样性等问题,提出一种以词为单位的 WoBERT-AT-BiLSTM-CRF 命名实体识别方法。数据集通过 WoBERT 预训练模型学习到拥有上下文语义信息的词向量,加入对抗训练生成对抗样本提高模型鲁棒性,再通过 BiLSTM 获得全面的文本表示,最后使用 CRF 利用序列标注之间的相关性来进一步优化命名实体识别结果。实验表明, WoBERT-AT-BiLSTM-CRF 模型识别结果优于其他对比模型,该模型准确率、召回率、 $F_1$  值分别为 94.21%、94.39%、94.30%,说明该方法的可行性,并为教育领域构建命名实体提供了一种新的方案。

**关键词:** 命名实体识别; Whisper; WoBERT; 对抗训练; 双向长短期记忆网络; 条件随机场

**中图分类号:** TP391;G40-057

**文献标志码:** A

**文章编号:** 2095-2163(2024)10-0063-07

## Named entity recognition method based on video resources and WoBERT-AT-BiLSTM-CRF

LIU Yang, TANG Hai, ZHU Menghan, XU Hongsheng

(School of Electrical and Information Engineering, Hubei University of Automotive Technology, Shiyan 442002, Hubei, China)

**Abstract:** Aiming at the lack of named entity recognition data sets in the field of education, this paper proposes using video resources to construct corresponding subject datasets. The traditional speech recognition model has a high word error rate, which is difficult to deal with long sequences. The end-to-end speech recognition model Whisper is proposed. A WoBERT-AT-BiLSTM-CRF named entity recognition method based on word is proposed to solve the problems of error accumulation and entity diversity in entity recognition. The data set learns word vectors with contextual semantic information through WoBERT pre-training model, and adds adversarial training to generate adversarial samples to improve the robustness of the model, then obtains comprehensive text representation through BiLSTM. Finally, CRF is used to further optimize the named entity recognition results by using the correlation between sequence annotations. The experimental results show that WoBERT-AT-BiLSTM-CRF model is superior to other comparison models. The accuracy rate, recall rate and  $F_1$  value of this model are 94.21%, 94.39% and 94.30%, respectively, indicating the feasibility of this method and providing a new scheme for constructing named entities in the field of education.

**Key words:** named entity recognition; Whisper; WoBERT; adversarial training; BiLSTM; CRF

## 0 引言

随着数字化的到来,网络教育和在线学习已成为当今教育领域的热门话题<sup>[1]</sup>。在线教育平台拥有大量的视频学习资源,将如何利用这些学习资源

为学生提供个性化的学习服务则变得至关重要。

命名实体识别是自然语言处理中的一项重要分支任务,在教育领域中是构建学科知识图谱、信息检索、文本分类等任务的基础,其目标是从文本中识别和分类出具有特定意义的实体。因此,对教育领

**基金项目:** 湖北省教育科学规划 2022 年度重点课题(2022GA049)。

**作者简介:** 刘洋(1999-),男,硕士研究生,主要研究方向:人工智能,自然语言处理;朱梦涵(1998-),女,硕士研究生,主要研究方向:人工智能;徐洪胜(1968-),男,副教授,主要研究方向:数据挖掘。

**通讯作者:** 唐海(1975-),男,博士,副教授,主要研究方向:人工智能。Email:smile-tang@163.com

**收稿日期:** 2023-06-12

域的命名实体识别进行研究,提出利用视频学习资源构建相关学科数据集,通过对数据集进行命名实体识别任务,在下游任务中可以帮助学生更好地理解视频内容,提高学习效率。

## 1 相关工作

早期的语音识别模型以统计的马尔科夫模型为主,Abdel-Hamid 等学者<sup>[2]</sup>提出深度置信网络(DBN)和马尔科夫模型(HMM)相结合的 NN-HMM 语音识别模型。该模型可以自动学习语音信号的特征,但对噪声和语音质量的变化比较敏感。Chen 等学者<sup>[3]</sup>提出一种自适应语言识别方法,模型学习声学模型和语音模型之间的映射,通过去除语音中的口音和方言来提高语音识别的准确性。Wu 等学者<sup>[4]</sup>提出基于动态稀疏神经网络的语音识别模型,不丧失模型准确率的情况下,提升了效率和计算速度。OpenAI 团队<sup>[5]</sup>在 2022 年提出基于大规模弱监督的语音识别模型 Whisper,模型使用大量的语音数据进行预训练,在不同的语言上有更好的泛化能力。研究同时指出,Whisper 基于 Transformer<sup>[6]</sup>架构。对于中文实体来说,模型可以根据上下文做出最合适的识别。

许多学者将深度学习方法引入到命名实体识别任务中。Lample 等学者<sup>[7]</sup>提出基于 BiLSTM-CRF 的命名实体识别模型,模型能够同时考虑到词性和上下文信息,具有较好的性能表现,常被用作实体识别任务的基准模型。王子牛等学者<sup>[8]</sup>将 BERT 预训练模型加入到命名实体识别任务中,BERT 旨在构建一个双向的语言模型,来更好地捕获语句间的上下文语义。Liu 等学者<sup>[9]</sup>提出了一种优化的 BERT 的预训练方法 RoBERTa,模型使用更大的字典和更好的分词方式,能够准确地切分文本和学习词向量表示,从而提高模型的准确性和泛化能力。Su<sup>[10]</sup>在 RoBERTa 模型的基础上进行预训练,得到一种以词为单位的模型 WoBERT,与以字为单位的模型相比,模型能够更好地理解自然语言中的语义和语法结构,词义的不确定性低,从而降低建模的复杂度。

以上研究为本文在视频资源下命名实体识别的研究提供参考。目前教育领域大多是通过网上爬取文本构建相关的数据集,但数据存在不准确、获取难度高等问题。本文提出通过利用视频资源,使用 Whisper 模型对视频进行文本化,对于构建出的实体存在上下文依赖、长距离特征抽取准确率低<sup>[11]</sup>等问题,本文提出一种基于 WoBERT-AT-BiLSTM-CRF

模型,使用 Wobert 模型获取丰富的语义信息,通过加入对抗训练来有效提高模型的鲁棒性,再使用 BiLSTM 对序列文本的每个词进行建模,最后通过 CRF 来优化模型的输出,从而提高实体识别的精度。

## 2 Whisper 语音识别模型

Whisper 语音识别模型采用端到端的方式,由编码器和解码器构成,模型结构与 Transformer<sup>[6]</sup>模型结构类似。

### 2.1 Whisper 模型结构

Whisper 语音识别模型结构如图 1 所示,可以看到,模型首先将输入数据转换为数据点传递到 2 个 1D 卷积层和 GELU 激活函数得到特征向量,加入位置编码输入 Transformer Encoder 端将每个输入编码成固定长度的向量。Transformer Decoder 端将这个向量作为输入,通过注意力机制生成输出文本。

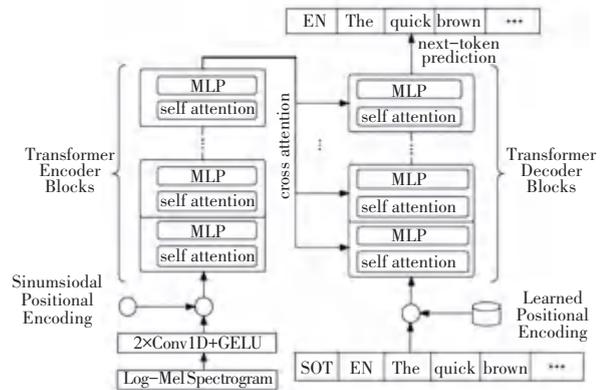


图 1 Whisper 模型结构

Fig. 1 Whisper model structure

### 2.2 Whisper 数据预处理

Whisper 语音识别模型将音频重采样到 16 000 Hz 作为输入,再转换为 80 通道的 Log Scale-Mel Spectrogram, Spectrogram 指在时间轴上的每个固定时间点上切割时序窗口,在每个时序窗口上进行傅里叶变换来计算每个频率上的能量,使一个一维信号变成二维信号,既有时间维度、也有频谱维度。Log-Scale 指计算出的能量进行对数运算得到分贝, Mel 指调节在不同频率之间的分辨率。Whisper 使用 80 个通道抽取 80 个维度的特征。每个时间窗口固定 25 ms,1 s 中有 16 000 个信号,先进行傅里叶变换,再进行分辨率调节,每次窗口则向后滑动 10 ms,使每 30 s 语言信号转换成 3 000 个数据点,每个数据点上维度为 80,将输出结果归一化到 1 和-1 之间。

### 2.3 Whisper 模型多任务设计

预训练方面, OpenAI 公司收集了 68 万小时的语音数据, 使用多任务和多语言的训练方法对 Whisper 模型进行预训练, 模型根据特殊的标记来执行不同的任务。模型中包括的 4 种识别任务分别为: 英文语音转英文文本; 任何语音转英文文本; 任意语言转任意语言文本; 识别出只有背景音乐。模型以 zero-shot<sup>[12]</sup> 的方式运作, 可以帮助模型更好地适应不同场景。

## 3 WoBERT-AT-LSTM-CRF 模型

WoBERT-AT-LSTM-CRF 模型结构如图 2 所示。

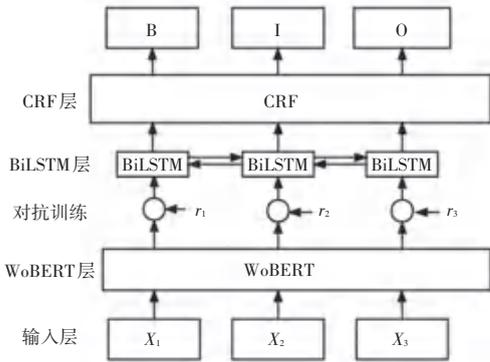


图 2 WoBERT-AT-BiLSTM-CRF 模型结构

Fig. 2 WoBERT-AT-BiLSTM-CRF model structure

由图 2 可知, 模型主要由 3 部分构成: 加入对抗训练的 WoBERT 层、BiLSTM 层、CRF 层。其中, WoBERT 层加入对抗训练生成扰动值  $r$ , 将扰动值  $r$  与词向量计算得到新特征向量, 可以防止各种扰动的输入, 从而提高模型在任务中的准确性。输入的词向量经过加入对抗训练的 WoBERT 层提取文本的语义特征。WoBERT 层使用预训练的编码器对文本进行编码, 产生词向量序列, 该序列表示文本的语义信息。词向量序列被输入到 BiLSTM 层中, BiLSTM 层用于捕获序列中的上下文信息, 生成一个新的特征序列。最后经过 CRF 层, 对序列进行标注得到每个词对应的实体标签。

### 3.1 WoBERT 模型

本文采用以词为单位的预训练模型 WoBERT, 相对于以字为单位的模型, 如 BERT<sup>[8]</sup>、MACBERT<sup>[13]</sup> 等, 以词为单位的模型可以更好地捕获语言的语义信息。词是语言中最基本的语义单位, 而字则无法完整地表达单词的含义, WoBERT 模型可以更好地处理词之间的关系, 从而提高模型的

精度。

WoBERT 模型在 ROBERTa 模型<sup>[14]</sup> 基础上进行预训练, 使用 MLM (Masked Language Model) 的预训练任务。初始化阶段, 将每个词用 ROBERTa 自带的 Tokenizer 切分为字, 再用字 Embedding 的平均作为词 Embedding 的初始化。该模型结构如图 3 所示。

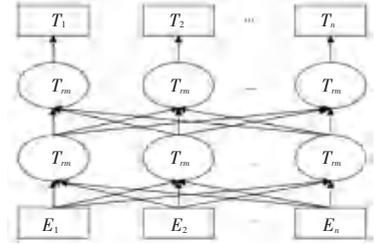


图 3 WoBERT 模型结构

Fig. 3 WoBERT model structure

首先, 序列文本通过 Embedding 转换为词向量  $E = (E_1, E_2, \dots, E_n)$ , 每个  $E_n$  包含 Token Embeddings、Position Embeddings、Segment Embeddings。其中, Token Embeddings 将文本中的每个单词转化为对应的向量表示。Segment Embeddings 用于处理多句子输入的场景。Position Embeddings 是一种序列化的数据结构, 使得模型可以更好地理解文本中的位置信息, 最后输出得到有语义特征的向量  $T = (T_1, T_2, \dots, T_n)$ 。

多头注意力机制是特征抽取的核心, 多头注意力机制的主要思想是计算词之间的关联程度调整权重系数来获取词新的权重。其结构如图 4 所示。

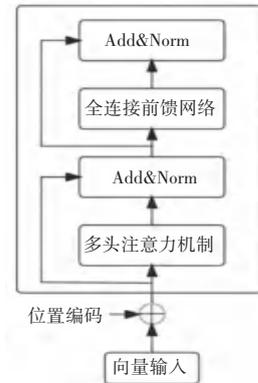


图 4 Transformer 中编码器结构

Fig. 4 Encoder structure in Transformer

在多头注意力机制中, 每个词向量经过线性变换得到矩阵  $Q$ 、矩阵  $K$ 、矩阵  $V$ , 计算公式如下:

$$Attention(Q, K, V) = \text{Softmax} \left( \frac{QK^T}{\sqrt{d_k}} \right) V \quad (1)$$

计算点积得到多头注意力机制的分数值, 再将

这些分数值进行线性变换和拼接得到特征信息作为多头注意力机制的输出。相关公式分别表示为:

$$head_i = Attention(Qw_i^q, Kw_i^k, Vw_i^v) \quad (2)$$

$$Multi-head(Q, K, V) = concat(head_1, \dots, head_h) W^O \quad (3)$$

WoBERT 的 Mask 机制在 BERT<sup>[8]</sup> 的基础上做了改进。采用了全词遮盖的方式,避免模型将一个词语中的某些字母单独遮盖,从而更好地保留了整个词的语义信息。

### 3.2 对抗训练

在 WoBERT 模型中使用 FGSM<sup>[15]</sup> 算法生成有扰动的对抗样本,FGSM 算法计算样本梯度加上扰动,使模型的预测结果发生错误,让模型更好地学习到输入文本中的干扰和噪声,来提高模型的鲁棒性和泛化性能。WoBERT 中对抗训练流程如图 5 所示。

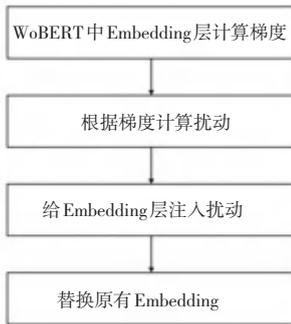


图5 对抗训练流程

Fig. 5 Adversarial training process

在对抗训练中,将 WoBERT 模型看作前向传播函数  $f(x)$ ,  $x$  表示输入样本,计算  $x$  的梯度生成对抗样本,将其加入到输出样本中得到扰动后的样本  $x'$ 。研究推得的公式为:

$$x' = x + \epsilon \text{sign}(\nabla_x J(f(x), y)) \quad (4)$$

其中,  $\text{sign}(\nabla_x J(f(x), y))$  是输入样本  $x$  梯度的符号,  $\nabla$  表示梯度方向;  $J(f(x), y)$  表示模型的损失函数;  $y$  表示标签;  $\epsilon$  表示扰动大小。

### 3.3 BiLSTM 模型

LSTM 是一种特殊的循环神经网络模型, LSTM 模型能够更好地捕捉序列中的长期依赖性,但是由于单向的 LSTM 模型无法处理上下文信息特征, Huang 等学者<sup>[16]</sup> 提出 BiLSTM。BiLSTM 在 LSTM 的基础上增加了一个反向的 LSTM 层,能够更好地捕捉序列中的信息,提高模型的性能。

BiLSTM 通过门控单元来控制信息的流动,包括输入门、遗忘门和输出门。这些门控制信息的流入

和流出,从而控制记忆单元的状态。在每个时间步, LSTM 单元根据输入和前一个时间步的隐状态计算当前时间步的隐状态,并输出一个表示该时间步的值。这种双向建模可以更好地利用序列数据中的上下文信息,提高模型性能。BiLSTM 模型结构如图 6 所示。

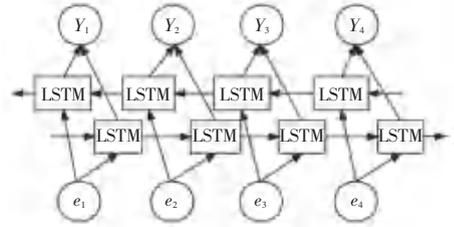


图6 BiLSTM 模型结构

Fig. 6 BiLSTM model structure

### 3.4 CRF 模型

在命名实体识别中, BiLSTM 能够获取序列的上下文信息,但是对于标签之间的相互作用并不能进行有效处理。因此,在 BiLSTM 模型后加入 CRF 模型<sup>[17]</sup>, 约束标签序列来提升标注结果的精确性。

CRF 模型参数包括  $(k+2) \times (k+2)$  的转移矩阵  $A$ , 这里  $k$  表示标签的数量,加 2 表示在句子中添加一个开始和结束位置,  $A_{i,j}$  表示标签  $i$  转移到标签  $j$  的概率。研究中,当一个长度和矩阵长度的标签序列  $y = (y_1, y_2, \dots, y_n)$  相等,则模型计算标签  $x$  的句子等于  $y$  的得分,可由如下公式求出:

$$\text{score}(x, y) = \sum_{i=1}^n P_{i, y_i} + \sum_{i=1}^{n+1} A_{y_{i-1}, y_i} \quad (5)$$

模型预测使用维特比算法解得最优路径,即预测的标签序列,计算方法见下式:

$$y^* = \text{argmax}(\text{score}(x, \tilde{y})) \quad (6)$$

## 4 实验与分析

### 4.1 实验数据集

本文利用视频资源构建相应的学科数据集。以计算机专业为例,选取《计算机网络》课程作为数据集的构造来验证实验的有效性。爬取《计算机网络》完整课程共计 38.3 h, 相关学科实体种类繁多,通过目录分析和相关专业知识整合,创建五大实体类型、11 类实体标签。为保证标注数据的质量,采用数据一致性进行评估。

对数据进行预处理等操作后,按句号进行数据集的分割,再对数据进行 5 种实体类型的 BIO 人工标注策略,为保证标注数据的质量,使用 YEDDA 标

注工具<sup>[18]</sup>, YEDDA 工具通过机器学习模型进行辅助标注, 提高标注效率和准确性。其中, “B” 表示实体开始, “I” 表示实体内部, “O” 表示非实体。实验数据共计 3 840 条句子, 12 347 个实体, 以句号按 8:1:1 划分为训练集、测试集、验证集。表 1 为计算机网络学科命名实体类型样例。

表 1 实体类型定义与样例

实体类别	实体英文	实体示例	实体数量
层	Layer	物理层	3 035
协议	Protocol	TCP/IP	2 298
硬件	Hardware	路由器	2 586
网络	Network	局域网	2 366
子定义	Definition	电子邮件	2 062

## 4.2 实验环境和参数配置

实验使用 Windows11 操作系统, i7-12700H CPU、2060 GPU、Python3.7 版本的编程语言。使用 Tensorflow1.15.0 深度学习框架; 2.3.1 Keras 深度学习库, 具体超参数取值见表 2。

表 2 超参数取值表

Table 2 Hyperparameter value table

参数	参数值	参数	参数值
Max-length	128	学习率	1e-5
WoBERT	12	批次大小	16
BiLSTM	128	优化器	Adam
Dropout	0.1	扰动强度	0.5
Epoch	20	隐层维度	768
Batch_size	32	优化器	Adam

## 4.3 评价指标

结合命名实体识别数据集的构建, 采用总词错率和英文词错率作为语音识别评价指标。其数学定义公式分别如下:

$$\text{总词错率} = \frac{\text{错词个数}}{\text{文本总词数}} \times 100\% \quad (7)$$

$$\text{英文词错率} = \frac{\text{英文错词个数}}{\text{文本总体英文词数}} \times 100\% \quad (8)$$

实验选取常用的命名实体识别评价指标: 准确率  $P$ 、召回率  $R$  和  $F_1$  值, 计算公式具体如下:

$$P = \frac{T_p}{T_p + F_p} \quad (9)$$

$$R = \frac{T_p}{T_p + F_n} \quad (10)$$

$$F_1 = \frac{2PR}{P + R} \quad (11)$$

其中,  $T_p$  表示模型正确识别个数;  $F_p$  表示模型识别非相关实体个数;  $F_n$  表示模型未识别到实体个数。

对标注数据采用内部一致性进行评估, 对于同一个实体是否给出相同的标注结果。为了评估内部一致性使用 Cohen's kappa 系数进行衡量, 计算公式如下:

$$kappa = (P_o - P_e) / (1 - P_e) \quad (12)$$

其中,  $P_o$  表示观察者间实际一致性,  $P_e$  表示观察者间机会一致性。

Cohen's kappa 系数是一种评估分类标注一致性的指标。该指标考虑了偶然性的影响。指标的区间为  $[-1, 1]$ , 其中  $-1$  表示完全相反的意见一致,  $0$  表示随机标注结果,  $1$  表示完全一致。

## 4.4 实验结果分析

为保证标注数据的质量, 将标记数据分为不同组进行重复标记, 通过不同分组标注数据 Cohen's kappa 系数进行来对 Cohen's kappa 系数大于 0.9 时为合格, 否则进行数据矫正, 直到所有分组系数高于 0.9。

为验证语音识别模型的有效性, 比较 Whisper 与目前主流的模型在语音识别上的表现, 实验选取 1 h 的课程进行对比实验, 实验结果见表 3。

表 3 语音识别模型词错率

识别方式	总词错率	英文词错率
Whisper-Tiny	9.64	9.29
Whisper-Base	7.74	9.15
Whisper-Small	4.31	13.24
Whisper-Medium	2.78	10.67
Whisper-Large	1.66	9.87
阿里云	1.59	25.32
科大讯飞	1.93	28.59
腾讯云	2.47	18.21

由表 3 可知, 阿里云的语音识别模型词错率效果最优, 比 Whisper-Large 低了 0.07%, 但阿里云在英文词错率较高, 比 Whisper-Large 高了 15.45%, 科大讯飞和腾讯的语音模型虽然在总体上有好的识别结果, 但是在识别英文时错误指数明显增加。结合针对不同学科所构建的学科数据集中的中英混杂情况, 综合比较 Whisper-Large 模型实验结果最优, 实验使用 Whisper-Large 模型进行数据集的构建。

设计 5 种不同的模型在相同的数据集下进行对

比实验,实验结果见表4。

从表4可知,BiLSTM-CRF模型效果较差。加入了预训练模型后模型整体的分类效果有明显的提升。BERT-LSTM-CRF模型比BERT模型有更高的准确率,效果提升了5.2%。基于词的WoBERT模型能够捕捉到单词之间的语义关系,从而更好地表达语言的含义,模型有更高的准确率,比BERT-BiLSTM-CRF模型准确率提升了2.41%。实验表明,本文提出的模型在命名实体识别任务中有着很好分类效果。

表4 实验结果对比

Table 4 Comparison of experimental results %

模型	P	R	F <sub>1</sub>
BiLSTM-CRF	86.60	79.78	83.05
BERT	90.72	91.21	90.65
BERT-BiLSTM-CRF	91.80	92.15	91.97
RoBERTa-BiLSTM-CRF	92.40	93.27	92.83
WoBERT-AT-BiLSTM-CRF	94.21	94.39	94.30

为进一步验证WoBERT-AT-BiLSTM-CRF模型的表现,对每个实体进行评估。各实体的指标对比见表5。从表5看出,本文提出模型在多数实体的识别指标优于其他模型,有较少的实体评价指标略低,可能个别实体数量偏少降低了实体的分类能力,但总体指标优于其他模型。

表5 各实体的指标对比

Table 5 Comparison of indicators by entity %

模型	实体类型	P	R	F <sub>1</sub>
BERT-	Layer	100.00	80.25	89.04
BiLSTM-	Definition	96.82	85.39	90.75
CRF	Protocol	90.91	71.73	80.19
	Hardware	95.62	71.20	81.62
	Network	89.80	72.13	80.00
RoBERT-	Layer	100.00	81.84	89.80
BiLSTM-	Definition	95.73	88.20	81.81
CRF	Protocol	92.31	70.89	80.19
	Hardware	97.76	71.20	82.39
	Network	93.75	73.77	82.75
WoBERT-	Layer	98.51	81.48	89.19
AT-	Definition	95.15	88.20	91.55
BiLSTM-	Protocol	95.03	82.57	82.30
CRF	Hardware	100.00	91.74	93.54
	Network	94.23	90.33	86.73

WoBERT层使用对抗训练减少过拟合风险,增

强模型的抗攻击能力。为验证对抗训练的有效性,优化后的WoBERT-AT-BiLSTM-CRF模型与WoBERT-BiLSTM-CRF模型进行对比实验。F<sub>1</sub>值随着训练次数的变化如图7所示。

从图7看出,加入对抗训练的WoBERT-AT-BiLSTM-CRF模型的整体识别效果得到提升,模型在第16个Epoch上取得了最大的F<sub>1</sub>值94.30%。加入对抗训练优化前后模型识别结果见表6。

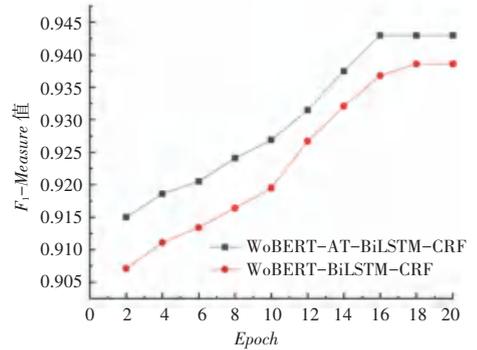


图7 F<sub>1</sub>值变化图

Fig. 7 Diagram of F<sub>1</sub> value variation

表6 优化前后模型识别结果

Table 6 Model identification results before and after optimization

模型	P / %	R / %	F <sub>1</sub> / %
WoBERT- BiLSTM-CRF	93.56	93.91	93.86
WoBERT-AT-BiLSTM-CRF	94.21	94.39	94.30

从表6看出,加入对抗训练的WoBERT-AT-BiLSTM-CRF模型识别结果得到提升,准确率、召回率、F<sub>1</sub>值分别为94.21%、94.39%、94.30%。与没有加入对抗训练的WoBERT-BiLSTM-CRF相比,准确率、召回率、F<sub>1</sub>值分别提高了0.65%、0.48%、0.44%,说明对抗训练对模型的识别效果有了一定提升。

## 5 结束语

针对教育领域中命名实体识别没有公开数据集,本文提出利用教育领域的视频资源构建相关数据集,为学科数据集的构建提供一种新的方案。

对于中文命名实体识别中存在识别误差积累、上下文依赖等问题,本文提出一种以词为单位的WoBERT-AT-BiLSTM-CRF命名实体识别方法,在WoBERT模型中加入对抗训练来提升模型的抗扰动性,加入BiLSTM和CRF进一步提高命名实体识别结果。在实验中各个指标均高于其他几个模型,证明本模型在实体识别任务的有效性,接下来为学生

情感分析、学科知识图谱的构建等任务打下基础。

## 参考文献

- [1] 石映辉, 蒲秋钰, 曲阜, 等. 在线课程资源学习满意度的现状、影响因素及其提升策略[J]. 现代教育技术, 2023, 33(3): 91-99.
- [2] ABDEL-HAMID O, MOHAMED A, JIANG H, et al. Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition [C]//2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Kyoto, Japan; IEEE, 2012: 4277-4280.
- [3] CHEN Yichen, YANG Zhaojun, YEH C F, et al. Aipnet: Generative adversarial pre-training of accent-invariant networks for end-to-end speech recognition [C]//ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Barcelona, Spain; IEEE, 2020: 6979-6983.
- [4] WU Zhaofeng, ZHAO Ding, LIANG Qiao, et al. Dynamic sparsity neural networks for automatic speech recognition [C]//ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Toronto, Canada; IEEE, 2021: 6014-6018.
- [5] RAFORD A, KIM J W, XU Tao, et al. Robust speech recognition via large-scale weak super-vision[J]. arXiv preprint arXiv:2212.04356, 2022.
- [6] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Advances in Neural Information Processing Systems. Long Beach, USA; NIPS Foundation, 2017, 30: 5998-6008.
- [7] LAMPLE G, BALLESTEROS M, SUBRAMANIAN S, et al. Neural architectures for named entity recognition [J]. arXiv preprint arXiv:1603.01360, 2016.
- [8] 王子牛, 姜猛, 高建瓴, 等. 基于 BERT 的中文命名实体识别方法[J]. 计算机科学, 2019, 46(S2): 138-142.
- [9] LIU Yinhan, OTT M, GOYAL N, et al. Roberta: A robustly optimized Bert pretraining approach [J]. arXiv preprint arXiv: 1907.11692, 2019.
- [10] SU J L. Speeding up without losing points: Chinese WoBERT based on word granularity [EB/OL]. [2020-09-18]. <https://kexue.fm/archives/7758>.
- [11] 冀振燕, 孔德焱, 刘伟, 等. 基于深度学习的命名实体识别研究 [J]. 计算机集成制造系统, 2022, 28(6): 1603-1615.
- [12] HUANG Jiacheng, LI Zuxin, ZHOU Zhe. A simple framework to generalized zero-shot learning for fault diagnosis of industrial processes [J]. IEEE/CAA Journal of Automatica Sinica, 2023, 10(6): 1504-1506.
- [13] CUI Yiming, CHE Wanxiang, LIU Ting, et al. Revisiting pre-trained models for Chinese natural language processing [J]. arXiv preprint arXiv:2004.13922, 2020.
- [14] LIU Shuang, YANG Hui, LI Jiayi, et al. Chinese named entity recognition method in history and culture field based on BERT [J]. International Journal of Computational Intelligence Systems, 2021, 14: 1-10.
- [15] GOODFELLOW I J, SHLENS J, SZEGEDY C. Explaining and harnessing adversarial examples [J]. arXiv preprint arXiv:1412.6572, 2014.
- [16] HUANG Zhiheng, XU Wei, YU Kai. Bidirectional LSTM-CRF models for sequence tagging [J]. arXiv preprint arXiv:1508.01991, 2015.
- [17] 游乐圻, 裴忠民, 罗章凯. 融合自注意力的 ALBERT 中文命名实体识别方法 [J]. 计算机工程与设计, 2023, 44(2): 605-611.
- [18] YANG Jie, ZHANG Yue, LI Linwei, et al. YEDDA: A lightweight collaborative text span annotation tool [J]. arXiv preprint arXiv:1711.03759, 2017.