

问答模式下结合属性语义的实体属性抽取研究

常露予, 张晓滨

(西安工程大学 计算机科学学院, 陕西 西安 710048)

摘要: 实体属性抽取任务中常面临属性标签过多时模型存在爆炸风险的问题, 且目前大多数属性抽取模型对文本均分配一致的注意力因子, 未将上下文的变化考虑在内。为解决上述问题, 提出一种基于问答模式的结合属性语义的实体属性抽取方法。该方法的要点在于, 将文本看作上下文, 把属性视为查询, 从上下文中提取到的答案等同于期望的属性值。文中对文本和属性的语义表示进行建模, 并提出一个动态注意力机制用于捕捉二者间的语义交互、实现信息融合, 同时自适应地控制属性信息融入文本向量的程度。为了验证该方法的有效性, 将模型与目前广泛应用的 BiLSTM 模型、BiLSTM-CRF 模型、OpenTag 模型和 Open Tagging 模型在包含大量属性标签的数据集 AE-110K、AE-650K 上进行对比实验, 结果表明, 模型在结合属性语义信息且采用动态 Attention 的条件下, 其预测准确度、召回率和 F1 值更高。

关键词: 问答模式; 实体属性抽取; 动态注意力; 语义交互; 信息融合

中图分类号: TP398.1

文献标识码: A

文章编号: 1673-629X(2024)04-0174-06

doi:10.20165/j.cnki.ISSN1673-629X.2024.0026

Research on Entity Attribute Extraction Combined with Attribute Semantics in Question Answering Mode

CHANG Lu-yu, ZHANG Xiao-bin

(School of Computer Science, Xi'an Polytechnic University, Xi'an 710048, China)

Abstract: Entity attribute extraction tasks often face the problem of model explosion risk when there are too many attribute labels, and at present, most attribute extraction models assign consistent attention factors to texts, and do not take context changes into account. To solve the above problems, an entity attribute extraction method based on question-answering mode combined with attribute semantics is proposed. The key point of this method is that the text is regarded as the context, the attribute is regarded as the query, and the answer extracted from the context is equivalent to the expected attribute value. The semantic representation of text and attributes is modeled, and a dynamic attention is proposed to capture the semantic interaction, realize information fusion, and adaptively control the degree to which attribute information is integrated into the text vector. In order to verify the effectiveness of the proposed method, the model was compared with the currently widely used models such as BiLSTM, BiLSTM-CRF, OpenTag and Open Tagging on datasets AE-110K and AE-650K containing a large number of attribute tags. It is showed that under the condition of combining attribute semantic information and adopting dynamic Attention, the model has higher prediction accuracy, recall rate and F1 value.

Key words: question answering mode; entity attribute extraction; dynamic attention; semantic interaction; information fusion

0 引言

近些年来, 迅速增长的互联网信息产生了海量数据源, 促使以实体为核心的信息抽取技术^[1-2]的提出。由于实体具有各自的属性, 需要找到一种方法有效地提取实体的属性信息, 属性抽取正是在这种背景下产生^[3]。

早期, 属性抽取主要按照人工指定的规则抽取实体属性。Hu 等人^[4]采用关联规则实现属性抽取, 使用

词典 WordNet^[5]获取评论观点, 对识别罕见的属性表示具有较好的效果。Zhuang 等人^[6]通过依存关系构建关系模板, 根据评价对象、评价意见成对出现的语法规则, 从影评文本中提取单元对。Qiu 等人^[7]通过信息共享获取评价对象、评价意见间的信息交互, 实现交叉迭代的双重抽取。Qian 等人^[8]针对不同样本的依存关系提出规则挑选算法, 利用规则制定相应的抽取模式。此类方法依赖大量手工规则, 无法提取规则之

收稿日期: 2023-06-07

修回日期: 2023-10-09

基金项目: 陕西省自然科学基金基础研究计划项目(2023-JC-YB-568)

作者简介: 常露予(1999-), 女, 通讯作者, 硕士研究生, 研究方向为知识图谱; 张晓滨(1970-), 男, 副教授, 硕导, 硕士, 研究方向为数据挖掘和个性化服务技术与应用。

外的属性,不适用属性类别丰富的开放域抽取任务。

自从 Jakob 等人^[9]引入条件随机场(Conditional Random Field, CRF)后,属性抽取逐渐被解释成序列化标注任务。具体地, Jakob 等人^[9]利用词汇、句法结构等多种特征构建特征表示,将目标建模成信息抽取任务,解决领域可移植性问题。徐冰等人^[10]在 CRF 的基础上引入浅层句法分析和启发式位置特征,在不额外增加领域词典的前提下,提高了属性抽取的精度和性能。Chernyshevich 等人^[11]基于 IHS Goldfire 语言处理器,利用 CRF 模型的词汇、句法和统计特征,实现属性的跨域提取。Shu 等人^[12]应用终身学习(Lifelong Machine Learning, LML)^[13]思想提出 L-CRF 方法,在 CRF 之后额外增加终身提取阶段,传统 CRF 与 LML 结合可以逐步形成更优的属性判别模式。利用 CRF 实现属性抽取的研究阶段,是一个重要的过渡期,其出现代表了研究焦点正从“约定确定的模式”向“规划不确定的模式”进行转变^[14]。

近期,研究主要集中在以神经网络为工具的抽取任务上。Huang 等人^[15]在长短时记忆网络(Long-Short Term Memory, LSTM)的基础上提出 LSTM-CRF 模型,可以在不借助单词嵌入的情况下实现标注准确度。Xu 等人^[16]将卷积神经网络(Recurrent Neural Network, RNN)^[17]引入属性抽取任务,提出具有双重嵌入机制的轻量化模型,这种思路取得了不错的效果。Fan 等人^[18]将目标信息分别编码到左右上下文,在解码器中结合全局上下文提取评价对象。尉桢楷等人^[19]将 RNN 的思想应用到注意力机制(Attention)的计算上,在卷积层中植入加权因子,避免全局注意力加权的盲目性。Liu 等人^[20]提出原型引导的属性范围最小化方法,基于统一学习和动态范围最小化方案促进多模态属性值提取。

然而,上述方法大多将每个属性作为一个实体类型,并为每个属性建立一套命名实体识别标签(如 BIO),这种方法导致属性标签过多、存在模型爆炸的潜在风险,不适用规模大、属性类别多的数据集。同时,目前的属性抽取模型对上下文均分配一致的注意力权重因子,并未将上下文的变化考虑在内。

基于此,该文提出一种基于问答模式的属性抽取模型。(1)和以往的抽取模型不同,该文在 CRF 的基础上对所有目标属性使用统一的全局 BIO 标签;(2)明确地对属性和上下文的语义表示进行建模,提出采用动态 Attention 捕获两者间的语义互动关系、自适应地控制融入后一层神经元的属性信息;(3)受机器阅读理解的启发,将问答模式应用到属性抽取任务中,在给定文本和属性的前提下,期望模型返回预测的属性值标签。最后,将该模型与目前广泛应用的 BiLSTM

模型、BiLSTM-CRF 模型、OpenTag 模型和 Open Tagging 模型做对比实验,取得了很不错的效果。

1 问题描述

该文将属性抽取定义为序列标注任务。给定文本 X' 和属性 X^a ,目标是从 X' 中为 X^a 提取相应的属性值,例如文本:2019 夏季女装纽扣装饰印花连衣裙露肩派对沙滩太阳裙 Boho 意粉长裙加尺码 FICUSRONG;属性:季节、性别、领口。对于三个不同的属性,模型的目标是获得用于季节的“夏季”、用于性别的“女性”和用于领口的“None”,其中前两个属性在文本中描述,而后者没有在文本中呈现。

形式上,给定产品文本 $X' = \{x'_1, x'_2, \dots, x'_m\}$ 和属性 $X^a = \{x^a_1, x^a_2, \dots, x^a_n\}$,模型输出标签序列 $y = \{y_1, y_2, \dots, y_m\}$, $y_i \in \{B, I, O\}$,其中 B 和 I 分别表示提取的属性值的开始标记和内部标记, O 表示值的外部。

2 问答模式的属性抽取模型

2.1 总体结构

以往的属性抽取模型大多将每个属性作为一个实体类型,并为每个属性建立一套命名实体识别标签,这导致在抽取过程中由于属性标签过多时模型存在爆炸风险的问题,该文为所有目标属性建立统一的 BIO 标签,将实体属性抽取任务扩展到属性类别庞大的数据集上;同时,现有的抽取方法未将上下文的变化考虑在内,对上下文均分配一致的注意力因子,且忽略了属性的语义信息,该文提出一种动态 Attention,用于捕捉文本和属性间的语义交互、实现其信息融合与取舍,该动态 Attention 使得模型能够自适应地控制融入下一层神经元的属性信息。

模型的总体结构如图 1 所示。为解决上述问题,提出一种问答模式下的属性抽取模型,将文本和属性分别视为上下文和查询,从上下文中提取的答案等同于期望的属性值。具体地,在给定实体与属性的前提下,可以看作询问实体属性的具体属性值,模型需要从文本中寻找其属性值答案。其结果可能存在以下两种情况:①若文本中不存在要找寻的属性值,模型的结果返回<None>;②若文本中存在要找寻的属性值,模型返回的是基于 BIO 标签预测的属性值答案。

2.2 算法流程

图 2 为该模型的算法流程,各模块的具体信息由后面的小节展开介绍。

2.3 语义信息的嵌入

采用中文预训练模型 BERT 进行训练,将文本和属性的每个字映射到高维向量空间。对于句子中的每一个字, BERT 根据具体上下文生成对应的向量表示。

给定输入文本 $X^t = \{x_1^t, x_2^t, \dots, x_m^t\}$ 、属性 $X^a = \{x_1^a, x_2^a, \dots, x_n^a\}$ ，分别经由 BERT 得到序列 $W^t = \{w_1^t, w_2^t, \dots, w_m^t\}$ 、序列 $W^a = \{w_1^a, w_2^a, \dots, w_n^a\}$ 。

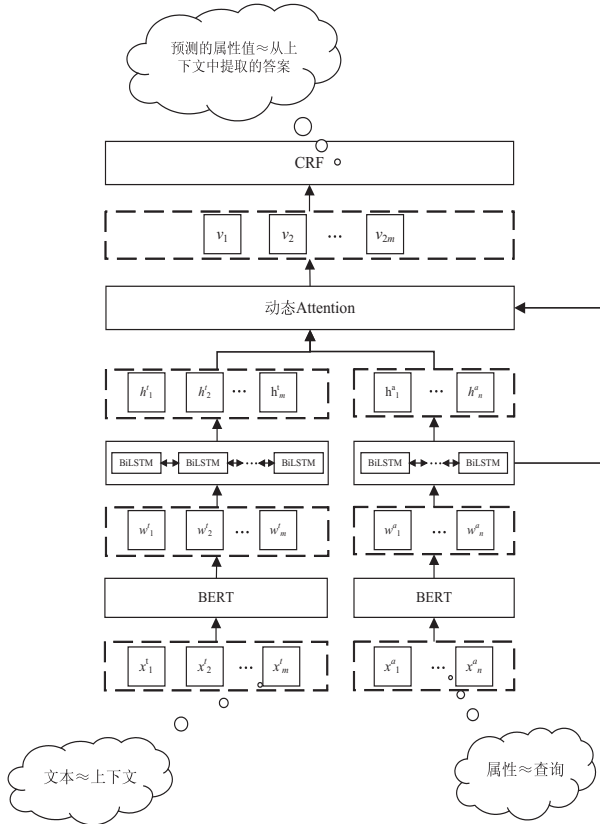


图1 基于问答模式的属性抽取总体结构

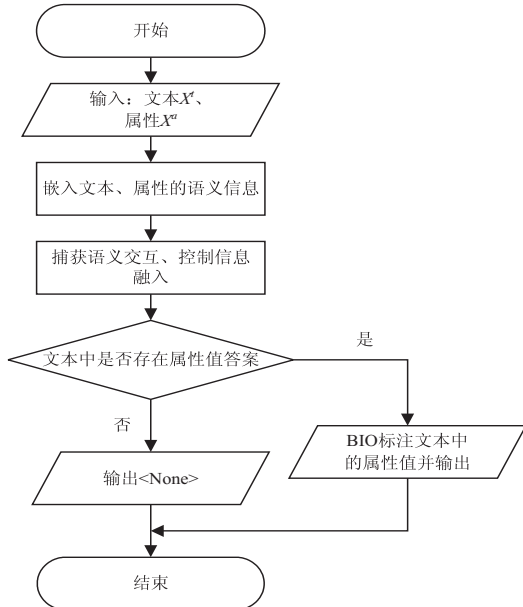


图2 算法流程

LSTM 能够处理梯度消失的问题，并且可以沿着序列模拟上下文信息。由于单向 LSTM 只考虑过去时间步的上下文特征，在获取信息上存在一定局限性。该文采用 BiLSTM 编码，从过去和未来的时间步中共同捕获上下文特征，用于获取序列的隐藏层表示。在

这项任务中，采用两组 BiLSTM 分别对文本和属性表示进行上下文嵌入。一组 BiLSTM 用于获取输入序列 W^t 的隐藏层表示 $H^t = \{h_1^t, h_2^t, \dots, h_m^t\}$ ：

$$H^t = [\vec{h}_i^t; \overleftarrow{h}_i^t] \quad (1)$$

其中，

$$\vec{h}_i^t = \text{BiLSTM}(\vec{h}_{i+1}^t, w_i^t), \quad \overleftarrow{h}_i^t = \text{BiLSTM}(\overleftarrow{h}_{i-1}^t, w_i^t) \quad (2)$$

另一组 BiLSTM 用于获取属性表示。属性的长度相较于文本要短很多，该文仅采用其最后隐藏状态作为属性表示：

$$H^a = [\vec{h}_n^a; \overleftarrow{h}_n^a] \quad (3)$$

类似地，

$$\vec{h}_n^a = \text{BiLSTM}(\vec{h}_{n+1}^a, w_n^a), \quad \overleftarrow{h}_n^a = \text{BiLSTM}(\overleftarrow{h}_{n-1}^a, w_n^a) \quad (4)$$

2.4 捕获语义交互、实现信息融合和取舍

Zheng 等人^[21]采用自注意力机制捕获重要信息，但为每个属性均引入一组标签，且忽略了属性的语义信息。为了能够更好地捕捉语义信息，文中模型通过计算文本和属性中每个词的相似度，来考虑二者间隐藏的语义交互。

具体地，如图3所示，节点 h_1^t 到 h_m^t 代表所有文本序列中的隐藏层节点，与属性序列中最后的隐藏层节点 h_n^a 进行注意力计算，得到 attention 向量 $S = \{s_1, s_2, \dots, s_m\}$ 。两个向量之间的相似性函数由余弦相似度来度量：

$$S = \text{cosine}(h_i^t, h_n^a) \quad (5)$$

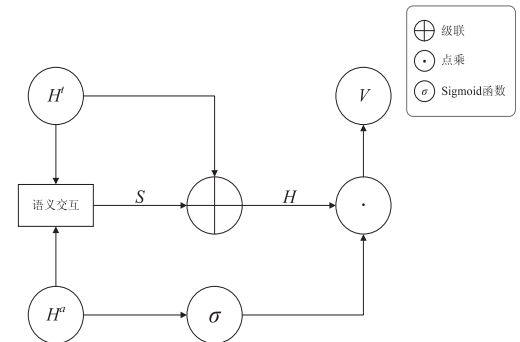


图3 结合语义信息的动态 Attention 网络结构

接下来，为了整合 H^t 和 S 的信息、提高模型的准确性和泛化能力，对二者分别进行线性变换后实施级联操作，得到融入属性信息的句子向量表示 H ：

$$H = (W_1 H^t + b_1) \oplus (W_2 S + b_2) \quad (6)$$

其中， W_1 、 W_2 、 b_1 和 b_2 均为可训练参数， \oplus 代表级联操作。

在完成两类隐藏层节点的初步信息融合后，引入门控机制控制属性信息融入文本向量的程度。如图3所示， H^a 为属性类型节点，对其进行线性变换和经过 Sigmoid 激活函数后，与向量 H 进行点乘操作得到向

量 $V = \{v_1, v_2, \dots, v_{2m}\}$ 。使用门控函数 Sigmoid 来衡量文本中的全局表示 X' 与属性表示 X^a 的信息重要程度,并根据其重要性进行取舍,进而得到最后的属性向量表示 V :

$$V = H \times \sigma(W_3 H_n^a + b_3) \quad (7)$$

其中, W_3 和 b_3 为可训练参数, \times 为点乘操作, σ 为 Sigmoid 函数, σ 属于 $0 \sim 1$ 的值。

2.5 输出预测的属性值

属性抽取任务的目的是预测和标记文本中属性值的位置,CRF 可以捕获邻域中输出标签间的依赖性。例如,如果已知一个 token 的标签是 I,那么下一个 token 的标签是 B 的概率会很低。

经由线性层和 dropout 降维后,向量 V 被传递到 CRF 层以预测标签序列。输出标签 y 的联合概率分布如下:

$$\Pr(y | X'; \psi) \propto \prod_{i=1}^m \exp\left(\sum_{k=1}^K \psi_k f_k(y_{i-1}, y_i, H_i)\right) \quad (8)$$

表1 数据集 AE-650K 中不同频率的 8 906 个属性分布统计和例子

Groups	Occurrence	# of Attributes	Example of attributes
High	[10 000, ∞)	10	Gender, Brand Name, Model Number, Type, Material
Sub-high	[1 000, 10 000]	60	Feature, Color, Category, Fit, Capacity
Medium	[100, 1 000]	248	Lenses Color, Pattern, Fuel, Design, Application
Low	[10, 100]	938	Heel, Shaft, Sleeve Style, Speed, Carbon Yarn
Rare	[1, 10]	7 650	Tension, Astronomy, Helmet Light, Flashlight Pouch

表2 数据集 AE-110K 的统计数据

Attributes	Train	Dev	Test
Brand Name	50 413	5 601	14 055
Material	22 814	2 534	6 355
Color	5 594	621	1 649
Category	5 906	590	1 462
Total	84 727	9 346	23 521

3.2 实验设置及评价指标

实验基于 PyTorch 实现,所有模型的词嵌入通过 BERT 预训练,维度为 768, BiLSTM 中隐藏状态的维度被设置为 512,其余参数设置如表 3 所示。

表3 参数设置

参数	参数值
Batch_size	32
Max_epoch	20
Learning_rate	2e-5
Decay_rate	0.5
Dropout	0.2

该文采用 BIO 标签策略,对任何属性都仅使用一

其中, ψ_k 代表相应的权重, f_k 是特征函数, K 为特征数量。

最终输出具有最高条件概率的最佳标签序列 y^* :

$$y^* = \operatorname{argmax}_y \Pr(y | X'; \psi) \quad (9)$$

3 实验分析

3.1 数据集

为了验证文中模型的有效性,选用公开数据集 AE-650K 和 AE-110K 进行实验。AE-650K 选用三元组 {文本, 属性, 值} 作为初始数据,包含 8 906 种属性,根据出现频率将其分为五组,表 1 显示了每个频率组中属性的数量以及一些例子。数据集 AE-110K 由四个经常出现的属性组成,共 117 594 个三元组,表 2 显示了 AE-110K 中属性的统计和分布情况。

套全局的 BIO 标签。使用准确度 (P)、召回率 (R) 和 F1 值 ($F1$) 作为评价指标, P , R 和 $F1$ 越高表示模型性能越好,各评价指标的计算方式如下。评价过程中遵循精确匹配的标准,只有当每个字的预测标签都正确时,才会将该结果当成正确答案。

$$P = TP / (TP + FP) \quad (10)$$

$$R = TP / (TP + FN) \quad (11)$$

$$F1 = (2P * R) / (P + R) \quad (12)$$

式中, TP 代表实际为正例,预测也为正例; FP 代表实际为负例,预测为正例; FN 代表实际为正例,预测为负例。

3.3 实验结果与分析

3.3.1 基线

为了验证文中方法的可靠性,选用四个序列标签模型作为基线。

BiLSTM 模型使用预训练模型 BERT 表示文本中的每个字,由 BiLSTM 产生文本上下文嵌入,最后选用 Softmax 激活函数预测标签; BiLSTM-CRF^[15] 作为命名实体识别的先驱和最先进的序列标签模型,使用 CRF 预测输出标签,由 BiLSTM 产生的隐藏状态被用

作 CRF 层的输入特征;OpenTag^[21]在 CRF 层前加入自注意力机制以突出重要信息;Open Tagging^[22]采用 BiLSTM 和 Attention 生成属性理解标题,之后传递给 CRF 层进行最终标注。

3.3.2 对比实验

由于基线模型不能扩展到属性标签庞大的数据集,表 4 仅报告其在数据集 AE-110K 的性能,而对于文中模型,报告了在 AE-110K 和 AE-650K 的表现。

表 4 对比实验结果 %

Attributes	Models	P	R	F1
Brand Name	BiLSTM	95.08	96.81	95.94
	BiLSTM-CRF	95.45	97.17	96.30
	OpenTag	95.18	97.15	96.35
	Open Tagging	96.65	97.21	96.39
	Our model-110K	97.12	97.38	97.25
	Our model-650K	96.89	97.51	97.63
Material	BiLSTM	78.26	78.54	78.40
	BiLSTM-CRF	77.15	78.12	77.63
	OpenTag	78.69	78.62	77.63
	Open Tagging	82.76	83.57	83.16
	Our model-110K	83.30	83.98	83.54
	Our model-650K	84.23	84.17	83.05
Color	BiLSTM	68.06	68.00	68.04
	BiLSTM-CRF	68.13	67.46	67.79
	OpenTag	71.19	70.50	70.84
	Open Tagging	75.11	72.61	73.84
	Our model-110K	77.23	74.73	75.20
	Our model-650K	79.10	74.85	76.96
Category	BiLSTM	82.74	78.40	80.51
	BiLSTM-CRF	81.57	79.94	80.75
	OpenTag	82.74	80.63	81.67
	Open Tagging	84.11	80.80	82.84
	Our model-110K	86.36	81.27	83.49
	Our model-650K	89.34	81.85	85.70

根据表 4 可以得出以下结论:

(1)在 AE-110K 数据集上,文中模型利用属性和文本间的语义信息,相较于其他四个基线, P 提高 0.47% ~ 9.17%, R 提高 0.18% ~ 6.73%, $F1$ 值提高 0.38% ~ 7.41%,可以有效地解决实体属性抽取问题。

(2)BiLSTM、BiLSTM-CRF、OpenTag 仅将属性抽取解释为序列标注任务,Open Tagging 模型结合属性语义生成属性理解标题后,性能优于同类其他基线,证明结合属性语义的有效性和合理性。

(3) 选用 Attention 的模型 OpenTag 和 Open

Tagging 仅关注文本的重要信息,该文采用动态注意力机制后,实验效果超过了 OpenTag 和 Open Tagging,证明考虑文本上下文变化对属性抽取的重要性。

(4)实验基于 AE-650K 中的四个高频属性,文中模型依旧表现出良好的稳健性,证明文中模型可以将属性抽取任务可以扩展到属性标签庞大的数据集上。

3.3.3 消融实验

为探究结合属性语义和动态 Attention 对实体属性抽取的影响,设计了以下消融实验。

(1)不结合属性语义,且采用静态 Attention 进行实体属性抽取。

(2)结合属性语义,仍采用静态 Attention 完成属性抽取。

(3)不结合属性语义,但采用动态 Attention 完成实体属性抽取任务。

三组实验均在数据集 AE-110K 上进行,选用属性“品牌名称”获取实验结果,如表 5 所示。

表 5 消融实验结果 %

实验内容	P	R	F1
不结合属性语义+静态 Attention	95.18	97.15	96.35
结合属性语义+静态 Attention	96.65	97.21	96.93
不结合属性语义+动态 Attention	95.76	97.33	96.89
结合属性语义+动态 Attention	97.12	97.38	97.25

表 5 说明,模型结合属性的语义信息但采用静态 Attention, P , R 和 $F1$ 均比仅采用常规策略的抽取模型的高。这是因为,将属性视为问答模式下的查询可以模拟嵌入在属性中的语义信息,结果表明结合属性语义完成实体属性抽取任务是很有必要的。当模型不结合属性语义但引入动态 Attention 后, P 提升 0.58%, R 提升 0.18%, $F1$ 提升 0.54%,验证了动态 Attention 的有效性。

以上实验表明,文中模型采用动态 Attention 且结合属性语义信息可以更好地捕获文本和属性间隐藏的语义交互,获得相较于其他大部分主流抽取模型更好的效果。

4 结束语

实体属性抽取任务常面临属性标签过多时模型存在爆炸风险的问题,且大多数抽取模型对 Attention 均分配一致的注意力因子,忽略了文本上下文的变化。为解决上述问题,该文提出一种基于问答模式下的结合属性语义的实体属性抽取模型,采用全局 BIO 标签预测属性值。将属性视为问答模式下的查询用来模拟嵌入在属性中的语义信息,采用动态 Attention 用于捕捉属性和文本间隐藏的语义交互,并实现其的融合和

取舍。相较于大多数属性抽取模型,该模型的效果更好,但在信息取舍时的抗干扰能力仍存在一定的改进空间,将在后续的研究过程中进一步深化和推进。

参考文献:

- [1] LIN Y, JI H, HUANG F, et al. A joint neural model for information extraction with global features [C]//Proc of the 58th annual meeting of the association for computational linguistics. Stroudsburg: Association for Computational Linguistics, 2020: 7999–8009.
- [2] LUAN Y, WADDEN D, HE L, et al. A general framework for information extraction using dynamic span graphs [C]//Proc of the 2019 annual conference of the North American chapter of the association for computational linguistics. Stroudsburg: Association for Computational Linguistics, 2019: 3036–3046.
- [3] 马进, 杨一帆, 陈文亮. 基于远程监督的人物属性抽取研究[J]. 中文信息学报, 2020, 34(6): 64–72.
- [4] HU M Q, LIU B. Mining and summarizing customer reviews [C]//Proc of the 10th ACM SIGKDD international conference on knowledge discovery and data mining. Seattle: ACM, 2004: 168–177.
- [5] MILLER G A. WordNet: a lexical database for English [J]. Communications of the ACM, 1995, 38(11): 39–41.
- [6] ZHUANG L, JING F, ZHU X Y. Movie review mining and summarization [C]//Proc of the 15th ACM international conference on information and knowledge management. Arlington: ACM, 2006: 43–50.
- [7] QIU G, LIU B, BU J, et al. Opinion word expansion and target extraction through double propagation [J]. Computational Linguistics, 2011, 37(1): 9–27.
- [8] QIAN L, GAO Z, LIU B, et al. Automated rule selection for aspect extraction in opinion mining [C]//Proc of the 24th international conference on artificial intelligence. CA: AAAI, 2015: 1291–1297.
- [9] JAKOB N, GUREVYCH I. Extracting opinion targets in a single and cross-domain setting with conditional random fields [C]//Proc of the 2010 conference on empirical methods in natural language processing. Stroudsburg: Association for Computational Linguistics, 2010: 1035–1045.
- [10] 徐冰, 赵铁军, 王山雨, 等. 基于浅层句法特征的评价对象抽取研究[J]. 自动化学报, 2011, 37(10): 1241–1247.
- [11] CHERNYSHEVICH M. IHS R&D Belarus: cross-domain extraction of product features using conditional random fields [C]//Proc of the 8th international conference on computational linguistics. Stroudsburg: Association for Computational Linguistics, 2014: 309–313.
- [12] SHU L, XU H, LIU B. Lifelong learning CRF for supervised aspect extraction [C]//Proc of the 55th annual meeting of the association for computational linguistics. Stroudsburg: Association for Computational Linguistics, 2017: 148–154.
- [13] CHEN Z, LIU B. Lifelong machine learning [J]. Synthesis Lectures on Artificial Intelligence and Machine Learning, 2018, 12(3): 1–207.
- [14] 徐庆婷, 洪宇, 潘雨晨, 等. 属性抽取研究综述[J]. 软件学报, 2023, 34(2): 690–711.
- [15] HUANG Z, XU W, YU K. Bidirectional LSTM-CRF models for sequence tagging [J]. arXiv: 1508.01991, 2015.
- [16] XU H, LIU B, SHU L, et al. Double embeddings and CNN-based sequence labeling for aspect extraction [C]//Proc of the 56th annual meeting of the association for computational linguistics. Stroudsburg: Association for Computational Linguistics, 2018: 592–598.
- [17] PORIA S, CAMBRIA E, GELBUKH A. Deep convolutional neural network textual features and multiple kernel learning for utterance-level multimodal sentiment analysis [C]//Proc of the 2015 conference on empirical methods in natural language processing. Stroudsburg: Association for Computational Linguistics, 2015: 2539–2544.
- [18] FAN Z, WU Z, DAI X Y, et al. Target-oriented opinion words extraction with target-fused neural sequence labeling [C]//North American chapter of the association for computational linguistics. Stroudsburg: Association for Computational Linguistics, 2019: 2509–2518.
- [19] 尉桢楷, 程梦, 周夏冰, 等. 基于类卷积交互注意力机制的属性抽取研究[J]. 计算机研究与发展, 2020, 57(11): 2456–2466.
- [20] LIU M, ZHU C, GAO H, et al. Boosting multi-modal e-commerce attribute value extraction via unified learning scheme and dynamic range minimization [J]. arXiv: 2207.07278, 2022.
- [21] ZHENG G, MUKHERJEE S, DONG X L, et al. OpenTag: open attribute value extraction from product profiles [C]//Proc of the 24th ACM SIGKDD international conference on knowledge discovery & data mining. New York: ACM, 2018: 1049–1058.
- [22] XU H, WANG W, MAO X, et al. Scaling up open tagging from tens to thousands: comprehension empowered attribute value extraction from product title [C]//Proc of the 57th annual meeting of the association for computational linguistics. Stroudsburg: Association for Computational Linguistics, 2019: 5214–5223.