

基于句法增强的细粒度情感三元组抽取方法

刘欣逸¹ 宁 博¹ 王 明¹ 杨 超² 商 迪¹ 李冠宇¹

¹(大连海事大学信息科学技术学院 辽宁大连 116026)

²(国网辽宁省电力有限公司信息通信分公司 沈阳 110006)

(lxy869281725@dlmu.edu.cn)

Fine-Grained Sentiment Triplet Extraction Method Based on Syntactic Enhancement

Liu Xinyi¹, Ning Bo¹, Wang Ming¹, Yang Chao², Shang Di¹, and Li Guanyu¹

¹(School of Information Science and Technology, Dalian Maritime University, Dalian, Liaoning 116026)

²(Information and Communication Branch, State Grid Liaoning Electric Power Supply Co., Ltd., Shenyang 110006)

Abstract The task of aspect sentiment triplet extraction (ASTE) mainly detects aspect terms and their corresponding opinion terms and sentiment polarities from sentences. However, when extracting multi-word aspect terms and opinion terms, it is impossible to accurately extract all words. The existence of repeated aspect terms and opinion terms makes it difficult for previous studies to capture all the correlations between aspect terms and opinion terms in word pairs. In response to these problems, we propose a framework based on syntactic enhanced multitasking learning to perform the task of end-to-end sentiment triplet extraction. The syntactic structure of a sentence reflects syntactic attributes and dependency or association information, therefore having a positive effect on the extraction task and the sentiment classification task. The proposed model utilizes dependency syntactic embedding graph convolutional network to fully mine syntactic features in sentences, and then transmits these features to 3 sub-tasks including aspect terms extraction, opinion terms extraction and sentiment analysis, thus realizing the effective fusion of syntactic information and multi-task joint learning framework. The model is evaluated in the sentiment analysis task upon 4 English datasets and 1 Chinese dataset. The experimental results show that the proposed model is effective and significantly better than other baseline models. At the same time, the results of specific case analysis prove that the method solves the problem of multiple words and repeated words to a certain extent.

Key words sentiment analysis; triplet extraction; syntactic dependency tree; graph convolutional networks; deep learning

摘 要 属性级情感三元组抽取 (aspect sentiment triplet extraction, ASTE) 任务主要是从句子中检测出属性词及其对应的评价词和情感倾向, 然而当抽取多词属性词和评价词时, 无法准确地抽取出全部的单词; 当面对重复的属性词和评价词时, 以往的研究很难学习到“属性词-评价词”词对之间所有的关联关系. 为解决这些问题, 提出了一种基于句法增强的多任务学习框架, 来解决端到端的情感三元组抽取任务. 句子中的句法结构反映的是句法属性和依赖关联信息, 这对抽取任务和情感分类任务有积极作用. 该方法是利用依存句法嵌入图卷积网络充分挖掘句法特征, 并将其传递到属性词抽取、评价词抽取和情感分析这 3 个子任务中, 实现了句法信息与多任务联合学习框架的融合. 在情感分析任务的 4 个英文数据集和 1 个

收稿日期: 2022-03-18; 修回日期: 2022-09-27

基金项目: 国家自然科学基金项目 (61976032, 62002039); 辽宁省教育厅科学研究面上项目 (LJKZ0063)

This work was supported by the National Natural Science Foundation of China (61976032, 62002039) and the Scientific Research General Program of the Education Department of Liaoning Province (LJKZ0063).

通信作者: 宁博 (ningbo@dlmu.edu.cn)

中文数据集上对模型进行了评估,实验结果表明,提出的方法是有效的且明显优于其他的基线模型,同时对具体案例进行分析,证明该方法一定程度上解决了多词和重复词的问题。

关键词 情感分析;三元组抽取;句法依存树;图卷积网络;深度学习

中图法分类号 TP391

文本情感分析是自然语言处理领域中的一个分支任务,主要是从大量且复杂的评价信息中抽取用户对各种产品的评价观点,进一步挖掘出这些评价信息深层次的情感倾向和含义。从电商平台或影评软件的评价信息提取出的评价观点不仅有利于用户快速了解产品信息以及评估产品购买价值,还可以为企业提供不同产品相应的使用反馈,从而帮助企业提升产品质量、跟进竞争策略和提高企业竞争力,因此文本情感分析有较高的研究价值。

传统的文本情感分析主要是针对文档或者句子进行情感倾向分析,得到整篇文档或者整个句子的情感倾向,如积极、消极或者中性,这是一个粗粒度的情感分析任务。如果一篇文档或一个句子中有多个属性词对应多个情感倾向时,这种粗粒度的分析无法抽取用户所评价的对象,也无法准确地分析出不同的评价对象所对应的情感倾向。随后,研究者们进一步研究细粒度的情感分析,其中属性级情感分析^[1-7]是分析已知的属性词对应的情感倾向,然而这种已知属性词的前提并不总是存在的,因此有人提出了属性词抽取子任务^[8-10],该子任务的目的是从句子中抽取需要的属性词。之后,研究者们^[11-17]将属性词抽取和情感分类2个子任务进行结合,抽取用户所评价的评价对象(也称属性词)和该属性词对应的情感倾向。然而,除了属性词和情感倾向,该属性词的评价词也是十分重要的要素,可以得到用户对该产品某一方面积极或者消极的态度,这可以给商家和消费者提供更加准确的数据。Peng等人^[18]认为评价词表示的是情感产生的原因,应该将评价词抽取看作一个重要的子任务而不是仅仅作辅助任务,因此提出了一个新的情感分析任务:属性级情感三元组抽取^[18-23](aspect sentiment triplet extraction, ASTE),该任务是在属性词抽取和情感分类联合学习的基础上将评价词也抽取出来,得到一个三元组(属性词,评价词,情感倾向)。

在上述细粒度情感三元组抽取任务中,研究者们没有考虑在联合框架中引入句法信息来共同增强多个子任务,然而句子之间的句法信息对属性词提取、评价词提取和情感分类任务有十分重要的影响。

例如,当属性词和评价词是由多个单词构成时,以往的研究很难准确地抽取全部的单词,存在跨度问题。本文认为如果利用句子之间的句法信息,并根据单词之间的依存关系,可以有效地提高抽取的准确性。当评价词与属性词相距较远时,之前的研究没有考虑二者之间的句法关系,可能不会将二者当作有关联的信息,无法提取出准确的三元组,但如果利用句法依存关系,就可以对评价词和属性词匹配有积极影响。因此针对上述问题,提出了一个基于句法增强的情感三元组抽取框架,该框架是联合学习框架,采用多头任务学习层对3个子任务共同学习,还可以充分利用句法信息增强多个子任务。设计了一个依存句法嵌入图卷积网络,充分地对句子中的每一个单词之间的句法依存关系进行建模。

如图1所示,“beef noodles were great but the service was dreadful!”中存在多个属性词“beef noodles”和“service”,也存在多个评价词和不同的情感倾向,表1所表示的是图1中的依存关系类型及其含义。其中“noodles”是属性词的一部分,根据依存关系“compound”可以得到“beef”与“noodles”有关联,组合为一个完整的属性词,又通过依存关系“nsubj”可以推断出该属性词对应的评价词为“great”,从而分析出其情感倾向为积极。整个句子是通过序列标注的方法抽取属性词和评价词,同时还需要识别出评价词和属性词之间的搭配关系以及所对应的情感倾向,最后获得细粒度的情感三元组分析结果:“beef noodles-great-positive”和“service-dreadful-negative”。

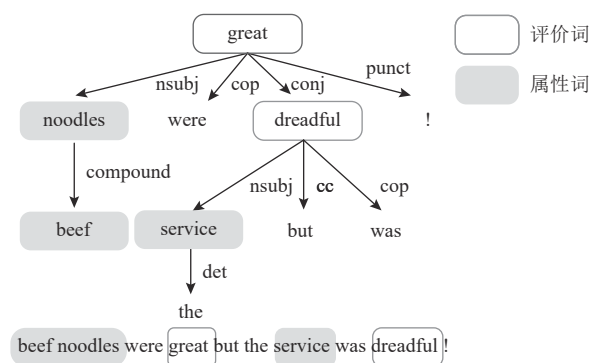


Fig. 1 Dependency parse tree case

图1 依存句法分析树案例

Table 1 Type of Dependency Syntactic Relationship

表1 依存句法关系类型

依赖类型	含义
conj	连接 2 个并列的词
nsubj	名词主语
compound	组合
cc	并列关系
cop	系动词
det	决定词
punct	标点符号

本文的贡献主要有 3 个方面:

1) 由于句子中存在重复或多词的属性词和评价词, 本文提出了一种基于句法增强的细粒度情感三元组抽取模型, 该模型利用句法信息来增强属性词抽取、评价词抽取和情感依赖分析这 3 个子任务, 从而提高了三元组抽取任务的准确性.

2) 使用图卷积网络(graph convolution network, GCN)与依存句法树来生成图节点的句法信息, 更好地保留了句子中的语法特征.

3) 在情感三元组任务的 4 个常用英文数据集和 1 个中文数据集上, 设计了本文模型与基准模型的对比实验, 并设计了子任务对比实验来进一步验证模型性能, 实验结果表明了本文模型的有效性.

1 相关工作

近年来在细粒度情感分析的研究中有多个子任务, 如属性级情感分类、属性词抽取、评价词抽取、情感分类联合学习和三元组任务等.

基于属性级的情感分类任务^[1-7], 是将已知的属性词融合到情感分类中. 例如, Dong 等人^[1]提出了自适应递归神经网络(adaptive recursive neural network, AdaRNN), 它根据每个单词之间的上下文和依存句法信息, 自适应地将单词的情感传递给目标. Yang 等人^[2]提出了一种以注意力为基础的双向 LSTM(long short-term memory)方法来进行基于目标的情感分类任务. Tang 等人^[4]针对注意力机制的缺点, 提出了一种渐进式的自监督注意力学习方法, 迭代地进行属性级情感分类预测, 完善注意力机制. 此外, Song 等人^[6]为解决传统循环神经网络(recurrent neural network, RNN)和注意力带来的问题, 提出了一种注意力编码器网络(attentional encoder network, AEN), 使用基于注意力的编码器来建模上下文和目标的关系, 从而

进行基于特定目标的情感分类任务. 而 Sun 等人^[7]将属性词构造为一个辅助句, 将属性级情感分类任务转化为句子对分类任务, 利用 BERT(bidirectional encoder representation from transformers)模型进行分类.

属性词抽取任务是抽取评价中的属性词, Yin 等人^[8]采用无监督的方式学习单词和依存路径的分布式表征, 并用条件随机场(conditional random field, CRF)模型进行属性词的抽取. Li 等人^[9]提出了一种新的属性词抽取框架, 利用评价词摘要和属性词检测历史信息来提高属性词抽取的准确性. 因为基于序列的方法不能充分利用整个句子的整体含义, 并且依赖关系存在局限性, 因此, Ma 等人^[10]提出了门控单元网络和位置感知注意力机制, 并将属性词抽取任务定义为序列到序列的任务.

之后, 研究者们^[11-17]将属性词抽取和情感分类 2 个子任务进行结合, Phan 等人^[11]采用流水线的方式先抽取属性词, 再利用情感分类器来得到情感倾向. 但由于子任务的分离, 模型的错误率会加大, 因此, Ma 等人^[12]提出联合学习属性词抽取和情感分类任务, 设计了分层堆栈双向门控循环单元(hierarchical multi-layer bidirectional gated recurrent units, HMBi-GRU)模型, 同时抽取目标并预测其情感极性. 该模型在情感分类任务中使用了评价词作为辅助任务, 没有更多地考虑评价词和分类任务之间的关联关系. 之后, 研究者对评价词有了更深入的研究, 例如 He 等人^[13]提出的交互式多任务学习网络(interactive multi-task learning network, IMN)模型可以同时进行多个任务的联合学习, 该模型中引入了一种消息传递机制, 可以通过加大评价词注意力权重来增强文本表示.

随后, Peng 等人^[18]提出了属性级情感三元组抽取方法^[18-23], 该方法采用流水线的方式先抽取“属性词-情感极性”和评价词, 之后通过分类器对属性词和评价词进行配对. 但是 Peng 等人^[18]的流水线方式还需要另外的配对任务, 有严重的局限性也增加了错误率. 因此, Xu 等人^[19]提出了一个基于位置感知标记的端到端联合模型, 实现了三元组的共同抽取, 很好地解决了流水线方法中存在的问题, 但是不能同时解决属性词重叠或者评价词重叠的问题. Wu 等人^[20]提出了网格标记方案(grid tagging scheme, GTS), 通过一个统一的网格标记任务以端到端的方式处理三元组任务或“属性词-评价词”词对抽取任务. 与此同时, Zhang 等人^[21]为解决统一标签和成对的“属性词-情感极性”带来的问题, 提出了一个多任务学习框架实现属性词和评价词的联合抽取以及二者之间的情

感关系解析,最终抽取出三元组.

2 细粒度情感三元组抽取模型

2.1 任务定义

本文属性级情感三元组抽取任务可以看作是序列标注任务和分类任务的结合. 给定输入文本序列 $X = \{x_i\}_{i=1}^{|X|}$, x_i 表示该句子中第 i 个单词, $|X|$ 表示整个句子的长度, 目标是输出一个三元组集合 $Y = \{y_j\}_{j=1}^{|Y|}$, y_j 表示第 j 个三元组元素, $|Y|$ 表示的是三元组集合的长度. 每一个三元组是由 3 个元素构成, 即 $y_j = ((sp_j^{(ap)}, ep_j^{(ap)}), (sp_j^{(op)}, ep_j^{(op)}), s_j^{(st)})$, 分别表示属性词跨度、评价词跨度和情感倾向, 其中 sp 和 ep 表示属性词的跨度和评价词的跨度, 即开始位置和结束位置. 因此, 本文的主要任务可以表达为将一个句子 $X = \{x_i\}_{i=1}^{|X|}$ 通过

模型学习转化为三元组集合:

$$Y = \{y_j | y_j = \{(sp_j^{(ap)}, ep_j^{(ap)}), (sp_j^{(op)}, ep_j^{(op)}), s_j^{(st)}\}_{j=1}^{|Y|}\}. \quad (1)$$

2.2 模型结构

细粒度情感三元组抽取框架是联合抽取框架, 其包括文本编码层、句法特征层、多头任务学习层和三元组解码层这 4 部分. 模型结构图如图 2 所示, 其中 n 表示输入句子的长度, m 表示输出三元组的个数, 将序列文本输入到由句子编码模块和句法关联模块组成的文本编码层, 得到句子上下文编码和依存句法分析树; 再进入带有依存关系的图卷积网络来充分学习句法信息, 得到上下文句法特征; 接着, 进入多头任务学习层, 同时进行属性词抽取、评价词抽取和情感依赖关系分析; 最后, 基于多头任务学习层获得的结果进行解码, 输出完整的三元组集合.

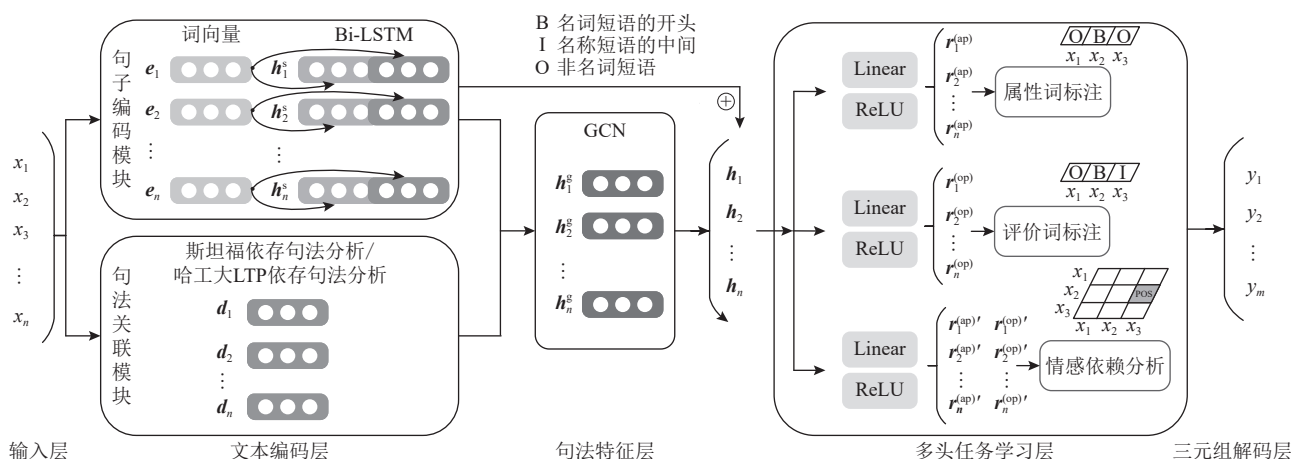


Fig. 2 Our model structure

图2 本文模型结构

2.3 文本编码层

学习语义信息和语境信息需要将句子转化为词向量, 这是文本编码层的关键. 文本编码层是由句子编码模块和句法关联模块组成的.

在句子编码模块中, 给定一个输入 $X = \{x_i\}_{i=1}^{|X|}$, 为了可以充分地学习句子中每个单词的上下文信息, 使用了一个双向长短期记忆网络 (bi-directional long short-term memory, Bi-LSTM). 将输入的句子转化为向量形式, 然后使用 Bi-LSTM 网络进一步充分学习句子中的上下文信息. Bi-LSTM 不同于 LSTM, 它由 2 个不同方向的 LSTM 组成, 因此 Bi-LSTM 不仅可以同 LSTM 那样学习到每一个单词在句子中的长期依赖关系, 而且可以从 2 个方向获取上下文信息, 保留了更多的重要信息. 输入的句子 $X = \{x_i\}_{i=1}^{|X|}$ 进行编码

后得到向量表示 $E_s = \{e_i | e_i \in \mathbb{R}^{d_e}\}_{i=1}^{|X|}$, 使用 Bi-LSTM 得到句子的上下文表示 $H_s = \{h_i^s | h_i^s \in \mathbb{R}^{2d_h}\}_{i=1}^{|X|}$, 其中, $h_i^s = (h_{LSTM}(e_i) \oplus h_{LSTM}(e_i))$, d_{emb} 和 d_{hid} 分别表示词向量的维度和 LSTM 的隐藏状态维度, $h_{LSTM}(e_i)$ 和 $h_{LSTM}(e_i)$ 分别表示正向和反向的 LSTM.

在句法关联模块中, 给定一个输入 $X = \{x_i\}_{i=1}^{|X|}$, 为学习到句子中的句法信息, 利用依存句法分析将每一个句子转化为依存句法树, 从而得到句子的依存关系. 输入的句子 X 利用斯坦福依存句法分析, 得到句子的依存句法表示 $D_{SP} = d_{SP}(X) = \{d_i\}_{i=1}^{|X|}$. 其中, SP 表示斯坦福依存句法分析器, d_i 表示第 i 个单词在依存句法树中相关联的父节点序号. 此外, 对于中文酒店数据集, 本文采用哈尔滨工业大学的中文自然语言处理工具 LTP 进行句法分析, 同样得到句子的依存

句法表示 $D_{LTP} = d_{LTP}(X) = \{d_i\}_{i=1}^{|X|}$.

2.4 句法特征层

该模型在句法特征层引入了图卷积网络(GCN)来学习单词之间的句法信息. 在上一层的句法关联模块中得到了依存句法分析树, 该依存分析树也可以表示为具有 n 个节点的图, 其中的节点就是每一个句子中的单词, 边表示的是单词之间的句法关联路径. 而依存分析树中的节点是经过上一层的句子编码模块中 Bi-LSTM 编码得到的向量表征结果, 上一层的这些输出是作为句法特征层的输入. 模型中采用 GCN 来建模句法特征, GCN 的主要思想是每一个节点的特征是由相邻节点不断更新得到的, 当存在多层的 GCN 时, 当前层节点的特征可以由上一层的相邻节点和自身节点特征更新获得.

将句子的上下文表示和依存分析树的邻接矩阵作为句法特征层中 GCN 的输入. 任意句子的 $n \times n$ 的邻接矩阵可以由其依存分析树转换而来, 在依存分析树中节点 i 和节点 j 之间存在一条依存路径, 则邻接矩阵的 $A_{ij}=1$, 否则 $A_{ij}=0$. 与此同时, 每一层的节点自身应当是有关联和影响的, 因此设计了一个自循环, 即 $A_{ii}=1$. GCN 是多层的结构, 每一层节点的信息是由上一层的节点通过其依存句法路径来传播和聚集特征信息, 并不断更新节点的特征. 因此, 将 Bi-LSTM 输出节点的上下文向量作为 GCN 的输入来进行句法关联信息的构建. GCN 中的第 $k+1$ 层的第 i 个节点更新为

$$\mathbf{h}_i^{k+1} = \sigma \left(\sum_{j=1}^n c^j (A_{ij} \mathbf{W}^k \mathbf{h}_j^k + \mathbf{b}^k) \right). \quad (2)$$

其中 \mathbf{h}_i^{k+1} 是 GCN 的第 $k+1$ 层中节点 i 的特征; \mathbf{h}_j^k 表示由 Bi-LSTM 建模得到的初始化向量, 是 GCN 的第 k 层中节点 j 的特征; $\sigma(\cdot)$ 是一个非线性激活函数, 本文使用的是 ReLU 激活函数, \mathbf{W}^k 和 \mathbf{b}^k 是可训练的权重和偏差, c^i 表示图中节点 i 的度数的倒数, 为归一化常数, 表示为

$$c^i = 1/d^i = 1 / \left(\sum_{j=1}^n A_{ij} \right). \quad (3)$$

最后通过 GCN 层得到带有依存句法关联的向量表示:

$$\mathbf{H}_{GCN} = \{\mathbf{h}_i^g | \mathbf{h}_i^g \in \mathbb{R}^{2d_h}\}_{i=1}^{|X|}, \quad (4)$$

其中 \mathbf{h}_i^g 表示最后一层 GCN 得到的依存句法特征.

此外, 还需要建立残差连接, 将 Bi-LSTM 输出的向量和 GCN 层输出的特征进行相加和残差连接, 得到多头任务层的输入, 这样训练的时候可以使梯度

直接走捷径反传给最初始层. 句法特征层的最终输出为

$$\mathbf{H} = \mathbf{H}_s + \mathbf{H}_{GCN} = \{\mathbf{h}_i = \mathbf{h}_i^s + \mathbf{h}_i^g | \mathbf{h}_i \in \mathbb{R}^{2d_h}\}_{i=1}^{|X|}. \quad (5)$$

2.5 多头任务学习层和解码层

2.5.1 多头任务学习层

多头任务学习层包括属性词抽取、评价词抽取和情感依赖分析 3 个部分, 下面介绍这 3 个部分.

1) 属性词和评价词抽取

采用序列标注的方法来进行属性词和评价词的抽取任务, 将句子中的单词标记为属性词 ap 和评价词 op. 句法特征层的隐藏状态包含了一些多余的信息, 可能会导致过拟合的风险. 因此, 首先使用简单的线性层和非线性激活函数从句法特征层输出的特征中分别抽取出属性词特征和评价词特征, 并将句法中的无关特性去除掉. 具体方法可由式(6)(7)所示:

$$\mathbf{r}_i^{(ap)} = \sigma(\mathbf{W}_{rep}^{(ap)} \mathbf{h}_i + \mathbf{b}_{rep}^{(ap)}), \quad (6)$$

$$\mathbf{r}_i^{(op)} = \sigma(\mathbf{W}_{rep}^{(op)} \mathbf{h}_i + \mathbf{b}_{rep}^{(op)}), \quad (7)$$

其中 $\mathbf{r}_i^{(ap)}$ 和 $\mathbf{r}_i^{(op)}$ 分别为属性词和评价词的特征表示, $\mathbf{W}_{rep}^{(ap)}$, $\mathbf{W}_{rep}^{(op)}$, $\mathbf{b}_{rep}^{(ap)}$, $\mathbf{b}_{rep}^{(op)}$ 是可学习的权重和偏差.

在得到属性词和评价词的特征表示后, 使用(B,I,O)序列标注方法对特征进行标注. 通过线性层和 softmax 函数得到属性词和评价词的分布序列 $\mathbf{p}_i^{(ap)}$ 和 $\mathbf{p}_i^{(op)}$:

$$\mathbf{p}_i^{(ap)} = \text{softmax}(\mathbf{W}_{tag}^{(ap)} \mathbf{r}_i^{(ap)} + \mathbf{b}_{tag}^{(ap)}), \quad (8)$$

$$\mathbf{p}_i^{(op)} = \text{softmax}(\mathbf{W}_{tag}^{(op)} \mathbf{r}_i^{(op)} + \mathbf{b}_{tag}^{(op)}). \quad (9)$$

因此, 属性词和评价词抽取任务的损失为

$$\mathcal{L}_{tag} = -\frac{1}{|X|} \sum_i \sum_k \hat{\mathbf{p}}_{i,k}^{(ap)} \ln(\mathbf{p}_{i,k}^{(ap)}) - \frac{1}{|X|} \sum_i \sum_k \hat{\mathbf{p}}_{i,k}^{(op)} \ln(\mathbf{p}_{i,k}^{(op)}), \quad (10)$$

其中 $\hat{\mathbf{p}}_{i,k}^{(ap)}$ 和 $\hat{\mathbf{p}}_{i,k}^{(op)}$ 分别表示每个单词的真实属性词和评价词的分布情况, $\mathbf{p}_{i,k}^{(ap)}$ 和 $\mathbf{p}_{i,k}^{(op)}$ 分别表示模型预测的属性词和评价词的分布情况, i 表示一句话中的第 i 个单词, k 表示第 i 个单词序列的标注类型.

2) 情感依赖分析

定义了 4 种情感依赖类型 POS, NEG, NEU, NO-DEP, 其中 POS 表示积极, NEG 表示消极, NEU 表示中性, NO-DEP 表示没有依赖关系.

与属性词抽取和评价词抽取 2 个子任务一样, 先用不同的参数抽取到属性词和评价词的特征, 如式(11)(12)所示, 然后用双仿射评分器来获取每个词对的依赖.

$$\mathbf{r}_i^{(ap')} = \sigma(\mathbf{W}_{rep}^{(ap')} \mathbf{h}_i + \mathbf{b}_{rep}^{(ap')}), \quad (11)$$

$$\mathbf{r}_i^{(op')} = \sigma(\mathbf{W}_{rep}^{(op')} \mathbf{h}_i + \mathbf{b}_{rep}^{(op')}). \quad (12)$$

句子中存在 $|X|^2$ 个由属性词和评价词组成的词对,因此期望可以对每一词对进行分析.为减少学习的冗余度,面对多词构成的属性词和评价词时,先只考虑最后一个词进行分析.每一个单词对依赖类型的评分可由式(13)计算得出:

$$\tilde{s}_{i,j,k} = [W^{(k)}r_i^{(ap)} + b^{(k)}]r_j^{(op)}, \quad (13)$$

其中 $\tilde{s}_{i,j,k}$ 表示单词对 $[x_i, y_j]$ 的第 k 个依赖类型得分, $W^{(k)}$ 和 $b^{(k)}$ 表示第 k 个依赖类型的可学习权重和偏差.

然后,使用 softmax 函数来进行归一化, $s_{i,j,k}$ 表示单词对 $[x_i, y_j]$ 所有依赖关系类型的概率.

$$s_{i,j,k} = \text{softmax}(\tilde{s}_{i,j,k}). \quad (14)$$

因此,情感依赖分析的损失由式(15)所示.其中 $\hat{s}_{i,j,k}$ 为每个词对的真实情感依赖概率, k 表示第 k 个依赖关系类型.

$$\mathcal{L}_{\text{dep}} = -\frac{1}{|X|^2} \sum_{(i,j)} \sum_k \hat{s}_{i,j,k} \ln(s_{i,j,k}). \quad (15)$$

最后,采用多头任务学习框架的联合训练方式,训练目标如式(16)所示,其中, α 表示权衡参数,用于平衡抽取任务和情感依赖分析任务, θ 表示可训练参数, γ 表示 L_2 正则化的权重值.

$$\min_{\theta} \mathcal{L} = \min_{\theta} \mathcal{L}_{\text{tag}} + \alpha \mathcal{L}_{\text{dep}} + \gamma \|\theta\|^2. \quad (16)$$

2.5.2 三元组解码层

在获得属性词、评价词和二者的情感依赖关系后,以评分器得到的情感依赖为基础,进行三元组的解码.如图1所示,句子“beef noodles were great but the service was dreadful !”,属性词抽取结果为(B,I,O,O,O,O,B,O,O,O,O),评价词抽取结果为(O,O,O,B,O,O,O,O,

B,O),情感依赖分析结果为(1,3,POS),(6,8,NEG),其中数字表示属性词和评价词中最后一个单词的下标.因此,通过情感依赖分析结果反向遍历属性词标签和评价词标签,解析出完整的三元组,最终的情感三元组为((0,1),(3,3),POS)和((6,6),(8,8),NEG),其中(0,1)表示的是属性词“beef noodles”的开始位置和结束位置,(3,3)表示的是评价词“great”的开始位置和结束位置,POS表示该属性词的情感倾向为积极.

3 实验

3.1 实验数据集

在Peng等人^[18]提出的ASTE-DATA-V1数据集和Xu等人^[19]提出的ASTE-DATA-V2数据集上对设计的模型进行评估,这2个数据集都包括了3个餐厅领域的数据集和1个笔记本领域的数据集,具体为:lap14,rest14,rest15,rest16.这2个数据集的具体情况如表2和表3所示,这2个数据集都有3种情感倾向:积极、消极、中性.ASTE-DATA-V1数据集中缺失了一些冲突情感倾向的三元组,其中包括重复评价词的三元组,例如,“good food and service”这句话中评价词“good”对应了2个属性词“food”和“service”,应该得到2个三元组,但是数据集ASTE-DATA-V1没有标注出这部分.而Xu等人^[19]补充了这部分缺失,进一步对数据集进行完善,得到一个新的数据集ASTE-DATA-V2.这2个数据集的4个领域数据集都来自SemEval-2014 task 4^[24], SemEval-2015 task 12^[25], SemEval-

Table 2 Statistics of Four Datasets from ASTE-DATA-V1

表2 ASTE-DATA-V1中的4个数据集统计

数据集	lap14				rest14				rest15				rest16			
	#S	#+	#0	#-	#S	#+	#0	#-	#S	#+	#0	#-	#S	#+	#0	#-
Train	920	664	117	484	1300	1575	143	427	593	703	25	195	842	933	49	307
Dev	228	207	16	114	323	377	32	115	148	179	9	50	210	225	10	81
Test	339	335	50	105	496	675	45	142	318	291	25	139	320	362	27	76

注: #S表示句子数量; #+, #0, #-分别表示积极、中性、消极的三元组数量.

Table 3 Statistics of Four Datasets from ASTE-DATA-V2

表3 ASTE-DATA-V2中的4个数据集统计

数据集	lap14				rest14				rest15				rest16			
	#S	#+	#0	#-	#S	#+	#0	#-	#S	#+	#0	#-	#S	#+	#0	#-
Train	906	817	126	517	1266	1692	166	480	605	783	25	205	857	1015	50	329
Dev	219	169	36	141	310	404	54	119	148	185	11	53	210	252	11	76
Test	328	364	63	116	492	773	66	155	322	317	25	143	326	407	29	78

注: #S表示句子数量; #+, #0, #-分别表示积极、中性、消极的三元组数量.

2016 task 5^[26].

此外,为进一步验证模型的有效性,本文使用谭松波酒店评论语料作为中文情感分析数据集,目前还没有中文的情感三元组数据集,因此,采用手工标注的方式对中文数据集进行处理.数据集具体细节如表4所示.

Table 4 Statistics of Chinese Datasets

表4 中文数据集统计

数据集	谭松波酒店评论				
	#S	#T	#+	#0	#-
Train	411	1 568	832	67	669
Dev	82	336	249	8	79
Test	114	422	305	5	112

注: #S表示句子数量; #T表示三元组总数量; #+, #0, #-分别表示积极、中性、消极的三元组数量.

3.2 实验参数设置

在本文的实验中,词嵌入使用预训练的 GloVe 词向量进行初始化,词嵌入维度为 300,隐藏向量维度为 300,设置联合训练的平衡参数 $\alpha=1$,正则化参数 $\gamma=10^{-5}$,为了缓解过拟合,将 dropout 设置为 0.5,训练过程中的学习率设置为 0.003,每次训练的样本数量 BatchSize 设置为 32,模型训练的轮数设置为 10.

3.3 对比实验

为了验证本文模型的有效性,使用 5 个基线模型进行对比实验.

1) Pipeline^[18]采用流水线的方式进行三元组抽取,并分为 2 个阶段:第 1 个阶段是采用 GCN 进行属性词的情感分类和评价词抽取;第 2 个阶段是将评价词与属性词进行配对.

2) CMLA^[27]是由 CMLA^[27](coupled multi-layer attentions)模型改良得到的,CMLA^[27]采用注意力机制学习属性词和评价词之间的相互关联影响,从而抽取目标属性词和评价词.CMLA^[27]是在第 1 阶段后加入了一个分类器来抽取出三元组.

3) JET^[19](jointly extract the triplets)是把 ASTE 任务定义为一个基于统一标签的序列标注任务,采用位置感知标记方法来实现端到端的联合三元组抽取.

4) GTS^[20]是把 ASTE 任务定义为一个统一的网格标注任务,采用网格标记方法实现端到端的三元组抽取任务,先抽取每个单词的情感倾向特征,然后利用这些特征得出每个“属性词-评价词”对的初始预测概率,最后设计了一个基于网格的解码策略来得到完整的三元组.

5) OTE-MTL^[21](multi-task learning framework for opinion triplet extraction)采用联合抽取方式,该框架先使用共享的编码器来学习 3 个子任务的联合特征,然后提出了一个多头任务学习框架共同实现属性词和评价词的联合抽取以及二者的情感关联解析,最终使用解码器来获得完整的三元组.

3.4 实验结果和分析

3.4.1 实验结果

在相同的数据集和实验环境下,本文模型与基线模型所运行的时间对比结果如表5所示,本文模型相比基线模型较为复杂,GCN 在训练过程中需要消耗部分时间,因此耗时较多,但是本文模型有更高的精度,整体性能更好.

Table 5 Total Training Time Results Compared with the Baseline Models

表5 与基线模型的训练总时间的对比结果

模型	各模型训练总时间/s
CMLA+	8 748
Pipeline	11 232
JET	9 539
GTS	10 620
OTE-MTL	9 288
本文模型	12 816

与基线模型的对比实验结果如表6所示,在精确率 P 、召回率 R 和 $F1$ 分数方面比较了基线模型和本文模型,实验结果都是取的最佳值.

在 ASTE-DATA-V1 数据集上本文模型整体性能优于 OTE-MTL^[21]模型,在精确率(P)、召回率(R)和 $F1$ 分数这 3 个性能上大多高于 OTE-MTL^[21]模型,结果表明本文模型能够有效地利用依存句法分析树和 GCN 来获取句子之间的句法特征,提高三元组的抽取效率.对于 ASTE-DATA-V1 数据集,本文的方法在 ASTE-DATA-V1 数据集的 4 个数据集上的 $F1$ 分数比 OTE-MTL^[21]模型分别提升了 0.22, 0.44, 1.74, 1.07 个百分点.但由于 ASTE-DATA-V1 数据集中缺失了部分数据,所以又在 ASTE-DATA-V2 数据集上和相同的基线模型进行对比实验.在 $F1$ 分数性能下,本文模型均优于基线模型,在性能提升最高的 lap14 数据集上, $F1$ 值达到 48.70%,比表现最好的模型 OTE-MTL^[21]提升了 1.95 个百分点,而在 ASTE-DATA-V1 数据集的 lap14 数据集上本文模型只提高了 0.22 个百分点,可以看出本文模型对 ASTE-DATA-V2 数据集中补充的重复评价词有积极影响;相比 rest14,

Table 6 Experimental Results of Different Models for Triplet Extraction Tasks

表 6 三元组抽取任务中不同模型的实验结果对比

%

模型	lap14 数据集				rest14 数据集				rest15 数据集				rest16 数据集			
	<i>P</i>	<i>R</i>	<i>F1</i>	<i>Dif</i>	<i>P</i>	<i>R</i>	<i>F1</i>	<i>Dif</i>	<i>P</i>	<i>R</i>	<i>F1</i>	<i>Dif</i>	<i>P</i>	<i>R</i>	<i>F1</i>	<i>Dif</i>
CMLA+*	32.60	35.20	33.85	+13.6	39.13	47.86	43.05	+18.5	34.50	37.79	36.07	+17.4	43.74	39.91	41.73	+17.9
Pipeline*	41.50	47.69	44.38	+3.12	43.34	64.80	51.94	+9.60	43.90	51.56	47.42	+6.08	47.30	63.84	54.33	+5.37
JET *	54.33	35.80	43.16	+4.34	68.09	47.94	56.26	+5.28	59.32	44.81	51.05	+2.45	66.24	50.40	57.24	+2.46
GTS*	51.60	43.28	47.08	+0.42	75.72	51.39	61.22	+0.32	64.83	46.40	54.08	-0.58	60.70	59.33	60.01	-0.31
OTE-MTL*	45.90	48.73	47.28	+0.22	69.47	54.54	61.10	+0.44	59.30	45.93	51.76	+1.74	63.81	54.24	58.63	+1.07
本文模型*	52.89	43.12	47.50	0	70.11	54.83	61.54	0	57.11	50.31	53.50	0	60.24	59.17	59.70	0
CMLA+	30.09	36.92	33.16	+15.5	39.18	47.13	42.79	+18.8	34.56	39.84	37.01	+15.8	41.34	42.10	41.72	+20.7
Pipeline	37.38	50.38	42.92	+5.78	43.24	63.66	51.50	+10.1	48.07	57.51	52.37	+0.43	46.96	64.24	54.26	+8.18
JET	54.84	34.44	42.31	+6.39	66.76	49.09	56.58	+4.96	59.77	42.27	49.52	+3.28	63.59	50.97	56.59	+5.85
GTS	58.02	40.11	47.43	+1.27	71.41	53.00	60.84	+0.70	64.57	44.33	52.57	+0.23	70.17	55.95	62.26	+0.18
OTE-MTL	54.26	41.07	46.75	+1.95	60.90	58.45	59.65	+1.89	61.76	43.30	50.91	+1.89	66.52	55.87	60.73	+1.71
本文模型	55.85	43.17	48.70	0	65.45	58.06	61.54	0	59.23	47.63	52.80	0	68.83	57.13	62.44	0

注: *P* 为精确率, *R* 为召回率, *Dif* 为本文模型相比基线模型 *F1* 分数的提升度; 带*号的模型使用的是 ASTE-DATA-V1 数据集, 不带*号的模型使用的是 ASTE-DATA-V2 数据集; 黑体数值表示最好的性能。

rest15, rest16 数据集, 本文模型都有所提升, 相比 OTE-MTL^[21] 模型, 本文模型提升了近 2 个百分点。大多数情况下, 本文模型在精确率和召回率都明显优于其他基线模型。由表 6 还发现, 流水线模型 Pipeline 和 CMLA+ 在召回率上取得较好的结果, 但由于精确率较低, 因此模型的整体性能较差。对于 ASTE-DATA-V2 数据集, 本文模型在 4 个数据集上的性能比 OTE-MTL 模型分别提升了 1.95, 1.89, 1.89, 1.71 个百分点, 比在 ASTE-DATA-V1 数据集上的提升度 *Dif* 都高。此外, 在大多数情况下, 相比其他基线模型, 本文模型在 ASTE-DATA-V2 数据集上的提升度 *Dif* 都明显高于 ASTE-DATA-V1 数据集上的提升度 *Dif*, 因此本文模型对重复评价词有一定帮助。

为进一步研究句法依存嵌入图卷积网络对整个模型的作用, 进行子任务的对比实验, 使用 ASTE-DATA-V2 数据集来分析属性词和评价词抽取任务的性能, 对比实验结果如表 7 所示。与 OTE-MTL^[21] 模型对比可得出, 在属性词抽取任务(ap)中, 本文模型在 ASTE-DATA-V2 数据集上都优于 OTE-MTL 模型, 性能提升最高的是在 lap14 数据集上 *F1* 分数达到 76.41%, 提升了 2.95%。在评价词抽取任务(op)中, 本文模型也都优于 OTE-MTL 模型, 性能提升最高的是在 lap14 数据集上 *F1* 分数达到 73.69%, 提升了 1.7%。在召回率性能上, 本文模型大多高于 OTE-MTL 模型, 虽然有些数据集的精确率没有高于 OTE-MTL 模型, 但是在整体性能 *F1* 分数上本文模型均取得较好的

Table 7 Experimental Results of Aspect and Opinion Extraction Tasks

表 7 属性词和评价词抽取任务的实验结果对比

%

数据集	任务	OTE-MTL			本文模型		
		<i>P</i>	<i>R</i>	<i>F1</i>	<i>P</i>	<i>R</i>	<i>F1</i>
lap14	ap	73.70	73.22	73.46	73.79	79.22	76.41
	op	74.77	69.41	71.99	71.43	76.11	73.69
rest14	ap	75.03	80.78	77.80	76.75	81.56	79.08
	op	77.28	83.26	80.16	82.50	79.02	80.72
rest15	ap	76.81	73.61	75.18	75.99	79.86	77.88
	op	72.59	71.80	72.19	71.52	75.70	73.55
rest16	ap	68.93	78.54	73.42	67.96	81.19	73.99
	op	80.66	82.53	81.58	80.28	85.68	82.89

注: ap 为属性词抽取任务, op 为评价词抽取任务, 黑体数值表示最好的性能。

结果。结果表明引入单词之间的句法依存关系对属性词抽取和评价词抽取任务均有很大提升, 猜测是因为当面对多词组成的属性词和评价词时, 若引入句法信息, 本文模型相比别的模型能得到更准确的抽取结果。

在谭松波酒店评论中文数据集中, 本文模型与基线模型的对比实验结果如表 8 所示。在三元组抽取任务(triplet)中, 本文模型相比 CMLA^[27] 模型提升很大, 相比 OTE-MTL^[21] 模型提升了 1.2 个百分点; 在属性词抽取任务(ap)中, 本文模型相比 OTE-MTL^[21] 模型提升了 3.23 个百分点; 在评价词抽取任务(op)中,

本文模型相比 OTE-MTL^[21]模型提升了 0.47 个百分点. 实验结果表明本文模型对中文数据集有积极影响, 但所有三元组抽取结果的 *F1* 值最高只有 43.29%, 效果一般. 猜测是因为中文数据集有较多口语化表达, 中文相比英文有更多非正式的语义和结构, 因此情感三元组抽取任务在中文数据集上还面临较大挑战.

Table 8 Experimental Results of Different Models on Chinese Datasets for Each Task

表 8 各任务中不同模型在中文数据集上的实验结果对比 %

任务	模型	<i>P</i>	<i>R</i>	<i>F1</i>
triplet	CMLA+	14.93	28.91	19.69
	OTE-MTL	43.90	40.43	42.09
	本文	45.97	40.91	43.29
ap	CMLA+	64.80	62.15	63.45
	OTE-MTL	66.18	69.59	67.84
	本文	76.33	66.49	71.07
op	CMLA+	58.87	55.18	56.96
	OTE-MTL	62.16	59.76	60.93
	本文	60.32	62.52	61.40

注: triplet 为三元组抽取任务, ap 为属性词抽取任务, op 为评价词抽取任务, 黑体数值表示最好的性能.

3.4.2 实验分析

根据表 6 和表 7 的实验结果可以得出, 本文模型在不同的数据集上均优于其他基线模型. CMLA+和 Pipeline 模型采用的是流水线的方式, 由表 6 可得流

水线的方法虽然在召回率上有较好性能, 但是它们的精确率较低, 因为流水线方式使三元组任务分离为 2 个步骤, 整个模型在训练过程中有局限性, 同时还加大了误差, 导致最终的结果较低. JET, GTS, OTE-MTL 模型是采用联合训练的方式, 很好地解决了流水线方式带来的问题, 取得较好的结果. 本文模型也采用联合训练的方式, 并且还引入了句法信息来提升句子特征, 在面对重复评价词的情况时, 本文模型在属性词抽取、评价词抽取和三元组抽取任务均可以取得更好结果. 因此可以得出, 本文提出的句法依存嵌入图卷积网络对三元组任务有一定提升. 猜测在属性词和评价词抽取任务中, 当面对多词构成的属性词和评价词时, 利用句法信息可以更加准确地获得它们的跨度情况. 在属性词和评价词相匹配的过程中, 本文模型利用句法依存信息可以去除无关的配对结果, 从而提升模型的准确性.

3.5 案例分析

为了说明本文提出的细粒度情感三元组抽取任务的粒度范围, 本文举例说明了粗粒度情感分析和细粒度情感分析结果. 如表 9 所示, 粗粒度情感分析是分析整个句子的情感倾向, 如第 2 个和第 4 个例句所示, 当句子中有多种情感倾向时, 粗粒度情感分析不能准确地得到每个属性词对应的情感倾向, 也不能得到该属性词对应的评价词, 而本文的细粒度三元组抽取模型可以更加直观地反馈句子中的属性

Table 9 Case Analysis

表 9 案例分析

案例情况	例句	真实值		各模型预测值		
		粗粒度情感分析	细粒度情感分析	CMLA+ ^[27]	OTE-MTL ^[21]	本文模型
多词属性词	Try green curry with vegetables .	POS	(green curry with vegetables, Try, POS)	(green curry, Try, POS) ×	(green curry with vegetables, Try, POS) .	(green curry with vegetables, Try, POS)
重复评价词	After really enjoying ourselves at the bar we sat down at a table and had dinner.	POS	(bar, enjoying, POS) , (table, enjoying, NEU) , (dinner, enjoying, NEU) .	(bar, enjoying, NEU×) , (dinner, enjoying, POS×) , (×) .	(dinner, enjoying, POS×) , (×) , (×) .	(bar, enjoying, POS) , (table, enjoying, NEU) , (dinner, enjoying, NEU) .
重复属性词	The owners and employees are friendly and their pizza is fantastic .	POS	(owners, friendly, POS) , (employees, friendly, POS) , (pizza, fantastic, POS) .	(owners, friendly, POS) , (owners, fantastic, POS) ×, (employees, friendly, POS) , (employees, fantastic, POS) ×, (pizza, friendly, POS) ×, (pizza, fantastic, POS) .	(owners, friendly, POS) , (employees, friendly, POS) , (employees, fantastic, POS) ×, (pizza, fantastic, POS) .	(owners, friendly, POS) , (employees, friendly, POS) , (pizza, fantastic, POS) .
多情感、重复属性词	The pizza is delicious - they use fresh mozzarella instead of the cheap , frozen , shredded cheese common to most pizzeria 's .	POS	(pizza, delicious, POS) , (fresh mozzarella, fresh, POS) , (cheese, cheap, NEG) , (cheese, frozen, NEG) , (cheese, shredded, NEG) .	(pizza, delicious, POS) , (pizza, shredded, NEG) ×, (×) , (×) , (×) .	(Pizza, delicious, POS) (fresh mozzarella, delicious, POS) ×, (×) , (×) , (×) .	(pizza, delicious, POS) , (cheese, cheap, NEG) , (cheese, frozen, NEG) , (cheese, shredded, NEG) , (×) .

注: ×表示抽取错误的三元组.

词及其相应的多种情感倾向和评价对象。

为了进一步研究模型的性能,选取 $F1$ 值提升相对较小的 rest16 数据集中的部分样本进行案例分析,结果如表 9 所示。

当面对多词属性词这种情况时,CMLA+不能准确地抽取出全部属性词,而 OTE-MTL 和本文模型可以抽取出准确结果。

当多个三元组有重复评价词时,CMLA+和 OTE-MTL 都不能抽取到全部的三元组,会丢失一部分,而本文模型可以抽取出全部三元组。猜测是因为别的模型没有句法依存关系,在学习过程中评价词 enjoying 和属性词 table 之间的句法关系联系不起来。

当存在重复属性词时,CMLA+将所有可能的配对情况都抽取出来,OTE-MTL 尽管错误配对的情况减少很多,但是依旧有 1 个错误的三元组;而本文模型能抽取出全部的正确三元组且没有多余的错误配对。猜测 CMLA+因为没有学习属性词和评价词之间的依赖关系,导致匹配结果误差较大;而 OTE-MTL 模型没有考虑句法信息,导致将无关联的“属性词-评价词”对抽取出来。而本文模型不仅学习了子任务之间的依赖关系,还引入了句法特征来辅助三元组的抽取,将句法关联的词对排除出去。

当面对一句话中存在多个情感倾向且有重复的属性词的情况时,CMLA+和 OTE-MTL 模型都没有抽取出全部的三元组,其中部分三元组的情感倾向也有错误。猜测 CMLA+和 OTE-MTL 模型在面对复杂的多情感句子时,很难有效地学习到词对之间的情感倾向的关联。而本文模型取得了一个较好的结果是因为模型在句法特征层有效地将句子特征融入上下文特征,因此在多头任务层时可以利用句法特征学习到词对和情感的依赖关系。本文模型除了一个三元组没有抽取到,剩下的三元组都得到了准确的结果。这个错误的三元组猜测是因为该三元组的评价词是属性词的一部分,而本文模型没有考虑到这种情况。

4 结 论

对于属性级情感三元组分析,单词之间的句法信息有助于抽取情感特征和目标词语。本文提出了基于句法增强的细粒度情感三元组抽取模型。由于多跨度的属性词和评价词抽取任务的准确性低,属性词和评价词配对研究还有很多不足,因此,提出了一种基于依存句法增强的多任务学习框架,通过依

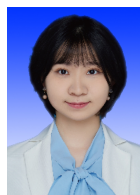
存句法信息嵌入图卷积网络来充分利用单词之间的句法信息挖掘语法知识和句法特征。然后将句法特征作用在多头子任务中,3 个子任务进行联合学习解析出完整三元组。在 4 个英文数据集和 1 个中文数据集上设计了对比实验,从而验证了本文模型的有效性。实验结果表明,该模型在细粒度情感三元组分析上可取得较好的结果,同时表明依存句法信息有助于属性词抽取、评价词抽取和三元组抽取任务。之后还考虑学习评价词抽取任务和情感依赖分析任务之间的关联作用,例如利用消息传递机制将抽取结果的特征传递到情感依赖分析任务中,实现子任务之间的交互。

作者贡献声明:刘欣逸提出算法思路和实验方案,完成实验并撰写和修改论文;宁博提出指导意见并修改完善论文;王明负责数据集收集和整理、论文校对工作;杨超提出实验改进建议;商迪负责论文审查;李冠宇负责理论指导。

参 考 文 献

- [1] Dong Li, Wei Furu, Tan Chuanqi, et al. Adaptive recursive neural network for target-dependent Twitter sentiment classification[C] // Proc of the 52nd Annual Meeting of the Association for Computational Linguistics, Volume 2: Short Papers. Stroudsburg, PA: ACL, 2014: 49-54
- [2] Yang Min, Tu Wenting, Wang Jingxuan, et al. Attention based LSTM for target dependent sentiment classification[C] //Proc of the 31st AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2017: 5013-5014
- [3] Li Xin, Bing Lidong, Lam Wai, et al. Transformation networks for target-oriented sentiment classification[C] //Proc of the 56th Annual Meeting of the Association for Computational Linguistics, Volume 1: Long Papers. Stroudsburg, PA: ACL, 2018: 946-946
- [4] Tang Duyu, Qin Bing, Liu Ting. Aspect level sentiment classification with deep memory network[C] //Proc of the 2016 Conf on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2016: 214-224
- [5] Tang Jialong, Lu Ziyao, Su Jinsong, et al. Progressive self-supervised attention learning for aspect-level sentiment analysis[C] //Proc of the 57th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2019: 557-566
- [6] Song Youwei, Wang Jiahai, Jiang Tao, et al. Attentional encoder network for targeted sentiment classification[J]. arXiv preprint, arXiv: 1902.09314, 2019
- [7] Sun Chi, Huang Luyao, Qiu Xipeng. Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence[C] //Proc of the 2019 Conf of the North American Chapter of the Association for

- Computational Linguistics: Human Language Technologies, Volume 1: Long and Short Papers. Stroudsburg, PA: ACL, 2019: 380–385
- [8] Yin Yichun, Wei Furu, Dong Li, et al. Unsupervised word and dependency path embeddings for aspect term extraction[C] // Proc of the 25th Int Joint Conf on Artificial Intelligence. New York: ACM, 2016: 380–385
- [9] Li Xin, Bing Lidong, Li Piji, et al. Aspect term extraction with history attention and selective transformation[C] //Proc of the 27th Int Joint Conf on Artificial Intelligence. New York: ACM, 2018: 4194–4200
- [10] Ma Dehong, Li Sujian, Wu Fangzhao, et al. Exploring sequence-to-sequence learning in aspect term extraction[C] //Proc of the 57th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2019: 3538–3547
- [11] Phan M H, Ogunbona P O. Modelling context and syntactical features for aspect-based sentiment analysis[C] //Proc of the 58th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2020: 3211–3220
- [12] Ma Dehong, Li Sujian, Wang Houfeng. Joint learning for targeted sentiment analysis[C] //Proc of the 2018 Conf on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2018: 4737–4742
- [13] He R, Lee W S, Ng H T, et al. An interactive multi-task learning network for end-to-end aspect-based sentiment analysis[C] //Proc of the 57th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2019: 504–515
- [14] Li Yuncong, Yin Cunxiang, Zhong Shenghua. Sentence constituent-aware aspect-category sentiment analysis with graph attention networks[C] // Proc of the 9th CCF Int Conf on Natural Language Processing and Chinese Computing. Berlin: Springer, 2020: 815–827
- [15] Li Xin, Bing Lidong, Zhang Wenxuan, et al. Exploiting BERT for end-to-end aspect-based sentiment analysis[C] //Proc of the 5th Workshop on Noisy User-Generated Text (W-NUT 2019). Stroudsburg, PA: ACL, 2019: 34–41
- [16] Ali W, Yang Yuwang, Qiu Xiulin, et al. Aspect-level sentiment analysis based on bidirectional-GRU in SLoT[J]. *IEEE Access*, 2021, 9: 69938–69950
- [17] Oh S, Lee D, Whang T, et al. Deep context-and relation-aware learning for aspect-based sentiment analysis[C] //Proc of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th Int Joint Conf on Natural Language Processing, Volume 2: Short Papers. Stroudsburg, PA: ACL, 2021: 495–503
- [18] Peng Haiyun, Xu Lu, Bing Lidong, et al. Knowing what, how and why: A near complete solution for aspect-based sentiment analysis[C] // Proc of the 34th AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2020: 8600–8607
- [19] Xu Lu, Li Hao, Lu Wei, et al. Position-aware tagging for aspect sentiment triplet extraction[C] //Proc of the 2020 Conf on Empirical Methods in Natural Language Processing (EMNLP). Stroudsburg, PA: ACL, 2020: 2339–2349
- [20] Wu Zhen, Ying Chengcan, Zhao Fei, et al. Grid tagging scheme for aspect-oriented fine-grained opinion extraction[C] //Proc of the Findings of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2020: 2576–2585
- [21] Zhang Chen, Li Qiuchi, Song Dawei, et al. A multi-task learning framework for opinion triplet extraction[C] // Proc of the Findings of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2020: 819–828
- [22] Wang Fang, Li Yuncong, Zhong Shenghua et al. Aspect-sentiment-multiple-opinion triplet extraction[C] // Proc of the 10th CCF Int Conf on Natural Language Processing and Chinese Computing. Berlin: Springer, 2021: 583–594
- [23] Xu Lu, Chia Y K, Bing Lidong. Learning span-level interactions for aspect sentiment triplet extraction[C] //Proc of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th Int Joint Conf on Natural Language Processing, Volume 1: Long Papers. Stroudsburg, PA: ACL, 2021: 4755–4766
- [24] Pontiki M, Galanis D, Pavlopoulos J, et al. SemEval-2014 task 4: Aspect based sentiment analysis[C] //Proc of the 8th Int Workshop on Semantic Evaluation. Stroudsburg, PA: ACL, 2014: 27–35
- [25] Pontiki M, Galanis D, Papageorgiou H, et al. SemEval-2015 task 12: Aspect based sentiment analysis[C] //Proc of the 9th Int Workshop on Semantic Evaluation. Stroudsburg, PA: ACL, 2015: 486–495
- [26] Pontiki M, Galanis D, Papageorgiou H, et al. SemEval-2016 task 5: Aspect based sentiment analysis[C] //Proc of the 10th Int Workshop on Semantic Evaluation. Stroudsburg, PA: ACL, 2016: 19–30
- [27] Wang Wenya, Pan Jialin, Dahlmeier D, et al. Coupled multi-layer attentions for co-extraction of aspect and opinion terms[C] // Proc of the 31st AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2017: 3316–3322



Liu Xinyi, born in 1997. Master candidate. Student member of CCF. Her main research interests include sentiment analysis and deep learning.

刘欣逸, 1997年生. 硕士研究生. CCF 学生会员. 主要研究方向为情感分析和深度学习.



Ning Bo, born in 1981. PhD, associate professor. Member of CCF. His main research interests include data management, data mining and privacy protection.

宁 博, 1981年生. 博士, 副教授. CCF 会员. 主要研究方向为数据管理、数据挖掘、隐私保护.



Wang Ming, born in 1999. Master candidate. His main research interests include multi-turn dialogue system and deep learning.

王 明, 1999年生. 硕士研究生. 主要研究方向为多轮问答系统和深度学习.



Yang Chao, born in 1988. PhD, senior engineer. His main research interests include data mining and artificial intelligence.

杨 超, 1988年生. 博士, 高级工程师. 主要研究方向为数据挖掘和人工智能.



Shang Di, born in 1997. Master candidate. Her main research interests include multi-turn dialogue system and deep learning.

商迪, 1997年生. 硕士研究生. 主要研究方向为多轮问答系统和深度学习.



Li Guanyu, born in 1963. PhD, professor. Member of CCF. His main research interests include intelligent information processing and semantic Internet.

李冠宇, 1963年生. 博士, 教授. CCF 会员. 主要研究方向为智能信息处理和语义互联网.