



(12) 发明专利

(10) 授权公告号 CN 114357156 B

(45) 授权公告日 2023. 02. 28

(21) 申请号 202111473889.1

(22) 申请日 2021.12.02

(65) 同一申请的已公布的文献号

申请公布号 CN 114357156 A

(43) 申请公布日 2022.04.15

(73) 专利权人 北京邮电大学

地址 100876 北京市海淀区西土城路10号

(72) 发明人 李睿凡 陈昊 翟泽鹏 冯方向  
王小捷

(74) 专利代理机构 北京金咨知识产权代理有限公司 11612

专利代理师 宋教花

(51) Int. Cl.

G06F 16/35 (2019.01)

G06F 40/211 (2020.01)

G06F 40/268 (2020.01)

G06F 40/284 (2020.01)

G06N 3/042 (2023.01)

G06N 3/0464 (2023.01)

G06N 3/047 (2023.01)

G06N 3/048 (2023.01)

G06N 3/08 (2023.01)

(56) 对比文件

CN 112883741 A, 2021.06.01

CN 113239186 A, 2021.08.10

CN 112163425 A, 2021.01.01

审查员 刘飞

权利要求书4页 说明书17页 附图4页

(54) 发明名称

一种多特征多通道图卷积网络模型训练方法及属性情感三元组抽取方法

(57) 摘要

本发明提供一种多特征多通道图卷积网络模型训练方法及属性情感三元组抽取方法,多特征多通道图卷积网络模型训练方法包括步骤:将第一语句输入预设的第一模型中,将第一语句划分为单词,编码得到第一词向量,得到隐藏状态序列,生成第一邻接张量;根据单词词性生成第二邻接张量,根据句法依存类型生成第三邻接张量,根据基于树的词对距离生成第四邻接张量,根据单词的相对距离生成第五邻接张量;隐藏状态序列分别与邻接张量图卷积,并平均池化,得到联合特征序列;将邻接张量进行拼接得到联合张量;根据联合张量和联合特征序列为每个词对生成第一词对向量,基于分类函数得到概率分布张量;计算总损失函数,根据总损失函数对第一模型进行训练。



1. 一种多特征多通道图卷积网络模型训练方法,其特征在于,该方法包括以下步骤:

将第一语句输入预设的第一模型中,将第一语句划分为多个单词,对每个所述单词进行编码得到第一词向量,组合多个所述第一词向量得到隐藏状态序列,基于双仿射注意力机制生成对应所述隐藏状态序列的第一邻接张量,基于双仿射注意力机制生成对应所述隐藏状态序列的第一邻接张量的步骤包括:利用多层感知机求得对应每个第一词向量的第二词向量;将第一语句中每两个单词组合为词对,对所述词对中分别对应两个单词的两个第二词向量进行注意力计算,得到对应词对的词组向量中每个维度的值,组合所述词组向量中每个维度的值,得到词组向量;组合所述第一语句中所有词对对应的词组向量得到第一邻接张量;

根据所述第一语句中单词的词性生成第二邻接张量,根据所述第一语句中单词之间的句法依存类型生成第三邻接张量,根据所述第一语句中单词之间基于树的词对距离生成第四邻接张量,根据所述第一语句中单词的相对距离生成第五邻接张量;

将所述隐藏状态序列分别与第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量进行图卷积,生成第一特征序列、第二特征序列、第三特征序列、第四特征序列和第五特征序列,对所述第一特征序列、第二特征序列、第三特征序列、第四特征序列和第五特征序列进行平均池化,得到联合特征序列;

将所述第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量进行拼接得到联合张量;

根据联合张量和联合特征序列为每个词对生成第一词对向量,基于分类函数对多个第一词对向量进行处理,得到概率分布张量;

根据所述概率分布张量计算差异损失函数,基于差异损失函数计算总损失函数,根据总损失函数对所述第一模型进行训练。

2. 根据权利要求1所述的方法,其特征在于,利用多层感知机求得对应每个第一词向量的第二词向量,根据如下公式:

$$h_i^a = \text{MLP}_a(h_i)$$

$$h_j^o = \text{MLP}_o(h_j)$$

$h_i$ 表示第*i*个单词对应的第一词向量, $h_j$ 表示第*j*个单词对应的第一词向量, $h_i^a$ 表示第*i*个单词对应的第二词向量, $h_j^o$ 表示第*j*个单词对应的第二词向量, $\text{MLP}_a$ 表示对第*i*个单词进行多层感知机处理, $\text{MLP}_o$ 表示对第*j*个单词进行多层感知机处理。

3. 根据权利要求1或2所述的方法,其特征在于,对所述词对中分别对应两个单词的两个第二词向量进行注意力计算,得到对应词对的词组向量中每个维度的值,根据如下公式:

$$g_{i,j} = h_i^{aT} U_1 h_j^o + U_2 (h_i^a \oplus h_j^o) + b_1$$

$g_{i,j}$ 表示由第*i*个单词和*j*个单词所组成词对的注意力计算的结果向量, $h_i^{aT}$ 表示第*i*个单词对应的第二词向量的转置, $h_j^o$ 表示第*j*个单词对应的第二词向量, $h_i^a$ 表示第*i*个单词对应的第二词向量, $U_1$ 和 $U_2$ 均表示注意力计算的权重参数, $b_1$ 表示注意力计算的偏置;

$$r_{i,j,k} = \frac{\exp(g_{i,j,k})}{\sum_{l=1}^m \exp(g_{i,j,l})}$$

$r_{i,j,k}$ 表示由第*i*个单词和*j*个单词所组成词对对应的词组向量中第*k*个维度的值, $g_{i,j,k}$ 表示结果向量第*k*个维度的值, $m$ 表示结果向量的维度数量, $l$ 表示结果向量的第1个维度。

4. 根据权利要求1所述的方法,其特征在于,根据所述第一语句中单词的词性生成第二邻接张量的步骤为,基于每个单词的词性对词对的词性进行标记,对标记相同的词对生成相同的词性向量,组合所有词对的词性向量得到第二邻接张量;

根据所述第一语句中单词之间的句法依存类型生成第三邻接张量的步骤为,基于每个词对中两个单词在第一语句中的句法依存类型对词对进行标记,对标记相同的词对生成相同的句法向量,组合所有词对的句法向量得到第三邻接张量;

根据所述第一语句中单词之间基于树的词对距离生成第四邻接张量的步骤为,基于每个词对中两个单词基于树的词对距离对词对进行标记,对标记距离相同的词对生成相同的距离向量,组合所有词对的距离向量得到第四邻接张量;

根据所述第一语句中单词的相对距离生成第五邻接张量的步骤为,基于每个词对中两个单词在第一语句中的相对距离对词对进行标记,对标记相对距离相同的词对生成相同的相对距离向量,组合所有词对的相对距离向量得到第五邻接张量。

5. 根据权利要求4所述的方法,其特征在于,将所述隐藏状态序列分别与第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量进行图卷积,生成第一特征序列、第二特征序列、第三特征序列、第四特征序列和第五特征序列的步骤包括:

分别提取第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量中每个通道的通道切片;

根据图卷积分别求出每个通道切片对应的切片特征序列;

收集第一邻接张量每个通道切片对应的切片特征序列,对第一邻接张量的所有切片特征序列进行平均池化,得到第一特征序列;

收集第二邻接张量每个通道切片对应的切片特征序列,对第二邻接张量的所有切片特征序列进行平均池化,得到第二特征序列;

收集第三邻接张量每个通道切片对应的切片特征序列,对第三邻接张量的所有切片特征序列进行平均池化,得到第三特征序列;

收集第四邻接张量每个通道切片对应的切片特征序列,对第四邻接张量的所有切片特征序列进行平均池化,得到第四特征序列;

收集第五邻接张量每个通道切片对应的切片特征序列,对第五邻接张量的所有切片特征序列进行平均池化,得到第五特征序列。

6. 根据权利要求5所述的方法,其特征在于,根据图卷积分别求出每个通道切片对应的切片特征序列,根据如下公式:

$$\widetilde{H}_k = \sigma(R_{:, :, k} H W_k + b_k)$$

$R$ 表示第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量的任一个, $R_{:, :, k}$ 表示第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量的任一个在第*k*个通道的通道切片, $\sigma$ 表示激活函数ReLU, $H$ 表示隐藏状态序列, $W_k$ 表示

图卷积的权重,  $b_k$  表示图卷积的偏置,  $\widetilde{H}_k$  表示第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量的任一个在第  $k$  个通道的通道切片所对应的切片特征序列。

7. 根据权利要求1所述的方法, 其特征在于, 基于分类函数对多个第一词对向量进行处理, 得到概率分布张量的步骤还包括:

基于所述分类函数得出第一词对向量所对应的预测词向量, 组合所有预测词向量得到概率分布张量。

8. 根据权利要求7所述的方法, 其特征在于, 若分类函数为softmax函数, 则基于所述分类函数得出第一词对向量所对应的预测词向量, 根据如下公式:

$$p_{ad} = \text{softmax}(W_p s_{ab} + b_p)$$

$p_{ab}$  表示第一语句中第  $a$  个词和第  $b$  个词所组成词对所对应预测词向量,  $W_p$  表示权重参数,  $b_p$  表示偏置, softmax表示将softmax函数作为分类函数,  $s_{ab}$  表示由第  $a$  个词和第  $b$  个词所组成词对所对应的第一词对向量。

9. 根据权利要求7或8所述的方法, 其特征在于, 所述第一语句预设有实际标签张量, 所述实际标签张量与概率分布张量的通道数相等, 所述实际标签张量和概率分布张量的通道均与预设的词对关系表中的关系一一对应, 所述第一语句中的词对在实际标签张量或概率分布张量的每个通道的值对应该词对在该种关系的得分, 所述第一语句的每个词对均预设有关实际关系, 所述词对在实际关系对应通道的值大于其他通道的值, 根据所述概率分布张量计算差异损失函数, 根据如下公式:

$$\mathcal{L}_p = - \sum_a^n \sum_b^n \sum_{c \in \mathcal{C}} I(y_{ab} = c) \log(p_{ab|c})$$

$\mathcal{L}_p$  表示差异损失函数的值,  $a$  表示第一语句中的第  $a$  个单词,  $b$  表示第一语句中的第  $b$  个单词,  $n$  表示第一语句中的总的单词数,  $\mathcal{C}$  表示预设的关系表中所有的关系类型,  $c$  表示关系表中的任一关系类型,  $y_{ab}$  表示由第  $a$  个单词和第  $b$  个单词所组成词对的实际关系的关系类型,  $p_{ab|c}$  表示由第  $a$  个单词和第  $b$  个单词所组成词对所对应的预测词向量在概率分布张量的  $c$  关系类型对应的通道所对应的值,  $I(\cdot)$  表示指示函数。

10. 根据权利要求9所述的方法, 其特征在于, 基于差异损失函数计算总损失函数的步骤包括:

分别计算所述第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量与实际标签张量的第一损失函数、第二损失函数、第三损失函数、第四损失函数和第五损失函数;

联合差异损失函数、第一损失函数、第二损失函数、第三损失函数、第四损失函数和第五损失函数进行计算, 得到总损失函数。

11. 根据权利要求10所述的方法, 其特征在于, 基于差异损失函数计算总损失函数, 根据如下公式:

$$\mathcal{L} = \alpha \mathcal{L}_p + \beta \mathcal{L}_1 + \gamma (\mathcal{L}_2 + \mathcal{L}_3 + \mathcal{L}_4 + \mathcal{L}_5)$$

$\mathcal{L}$  表示总损失函数的值,  $\mathcal{L}_p$  表示差异损失函数值的值,  $\mathcal{L}_1$  表示第一损失函数的值,  $\mathcal{L}_2$  表示第二损失函数的值,  $\mathcal{L}_3$  表示第三损失函数的值,  $\mathcal{L}_4$  表示第四损失函数的值,  $\mathcal{L}_5$  表示第五

损失函数的值,  $\alpha, \beta, \gamma$  均表示用于调节对应损失影响的超参数。

12. 一种属性情感三元组抽取方法, 其特征在于, 该方法包括以下步骤:

将第二语句输入根据上述权利要求1-11任一项所述方法所训练的第一模型, 得到所述第二语句所对应的概率分布张量;

对所述第二语句所对应的概率分布张量进行三元组解码, 得到第二语句中的属性情感三元组。

## 一种多特征多通道图卷积网络模型训练方法及属性情感三元组抽取方法

### 技术领域

[0001] 本发明涉及自然语言处理技术领域,尤其涉及一种多特征多通道图卷积网络模型训练方法及属性情感三元组抽取方法。

### 背景技术

[0002] 属性级情感分析 (Aspect-based Sentiment Analysis, ABSA) 是一个细粒度情感分析和观点挖掘任务,在自然语言处理领域受到了广泛的关注和研究。通常,ABSA 由三个最基本的任务组成,分别是属性词抽取 (Aspect Term Extraction, ATE)、属性情感分类 (Aspect Sentiment Classification, ASC) 和观点词抽取 (Opinion Term Extraction, OTE)。ATE 旨在识别出句子中的属性词,而 ASC 则去判断属性词的情感极性,OTE 任务目标是抽取出句子中所有的观点词。然而,在实际的应用中,我们需要同时获得属性词、描述该属性词的观点词以及属性词的情感极性。

[0003] 属性情感三元组旨在从句子评论中抽取出所有的三元组,每个三元组包含三个元素,分别是属性词、观点词和对应的情感极性。大多数的研究聚焦于设计一种可以实现端到端的属性情感三元组抽取的标注方法。然而,这些方法忽略了在该任务中句子的词与词之间的关系。

### 发明内容

[0004] 鉴于此,本发明实施例提供了一种多特征多通道图卷积网络模型训练方法及属性情感三元组抽取方法,以消除或改善现有技术中存在的一个或多个缺陷。

[0005] 本发明的一个方面提供了一种多特征多通道图卷积网络模型训练方法,该方法包括以下步骤:

[0006] 将第一语句输入预设的第一模型中,将第一语句划分为多个单词,对每个所述单词进行编码得到第一词向量,组合多个所述第一词向量得到隐藏状态序列,基于双仿射注意力机制生成对应所述隐藏状态序列的第一邻接张量;

[0007] 根据所述第一语句中单词的词性生成第二邻接张量,根据所述第一语句中单词之间的句法依存类型生成第三邻接张量,根据所述第一语句中单词之间基于树的词对距离生成第四邻接张量,根据所述第一语句中单词的相对距离生成第五邻接张量;

[0008] 将所述隐藏状态序列分别与第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量进行图卷积,生成第一特征序列、第二特征序列、第三特征序列、第四特征序列和第五特征序列,对所述第一特征序列、第二特征序列、第三特征序列、第四特征序列和第五特征序列进行平均池化,得到联合特征序列;

[0009] 将所述第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量进行拼接得到联合张量;

[0010] 将第一语句中每两个单词组合为词对,根据联合张量和联合特征序列为每个词对

生成第一词对向量,基于分类函数对多个第一词对向量进行处理,得到概率分布张量;

[0011] 根据所述概率分布张量计算差异损失函数,基于差异损失函数计算总损失函数,根据总损失函数对所述第一模型进行训练。

[0012] 本发明的多特征多通道图卷积网络模型训练方法,我们首先为属性情感三元组任务定义多种关系类型,然后采用双仿射注意力机制(Biaffine)构建第一邻接张量表示词之间的关系,多特征多通道图卷积网络(Multi-feature Multi-channel Graph Convolutional Network,MMGCN)可以将句子转换成一个多通道图结构,如本申请中的各个邻接张量,所述关系类型的数量与各个邻接张量的通道数相等,因此,MMGCN可以学习到关系意识的节点表示。此外,我们引入了词性、句法、基于树的词对距离和单词的相对距离,作为多种特征以增强MMGCN模型,充分考虑了句子中词与词之间的关系,提高三元组提取的准确度。

[0013] 在本发明的一些实施方式中,基于双仿射注意力机制生成对应所述隐藏状态序列的第一邻接张量的步骤包括:

[0014] 利用多层感知机求得对应每个第一词向量的第二词向量;

[0015] 对所述词对中分别对应两个单词的两个第二词向量进行注意力计算,得到对应词对的词组向量中每个维度的值,组合所述词组向量中每个维度的值,得到词组向量;

[0016] 组合所述第一语句中所有词对对应的词组向量得到第一邻接张量。

[0017] 在本发明的一些实施方式中,利用多层感知机求得对应每个第一词向量的第二词向量,根据如下公式:

$$[0018] \quad h_i^a = \text{MLP}_a(h_i)$$

$$[0019] \quad h_j^o = \text{MLP}_o(h_j)$$

[0020]  $h_i$ 表示第i个单词对应的第一词向量, $h_j$ 表示第j个单词对应的第一词向量, $h_i^a$ 表示第i个单词对应的第二词向量, $h_j^o$ 表示第j个单词对应的第二词向量, $\text{MLP}_a$ 表示对第i个单词进行多层感知机处理, $\text{MLP}_o$ 表示对第j个单词进行多层感知机处理。

[0021] 在本发明的一些实施方式中,对所述词对中分别对应两个单词的两个第二词向量进行注意力计算,得到对应词对的词组向量中每个维度的值,根据如下公式:

$$[0022] \quad g_{ij} = h_i^{aT} U_1 h_j^o + U_2 (h_i^a \oplus h_j^o) + b_1$$

[0023]  $g_{i,j}$ 表示由第i个单词和j个单词所组成词对的注意力计算的结果向量, $h_i^{aT}$ 表示第i个单词对应的第二词向量的转置, $h_j^o$ 表示第j个单词对应的第二词向量, $h_i^a$ 表示第i个单词对应的第二词向量, $U_1$ 和 $U_2$ 均表示注意力计算的权重参数, $b_1$ 表示注意力计算的偏置;

$$[0024] \quad r_{ij,k} = \frac{\exp(g_{ij,k})}{\sum_{l=1}^m \exp(g_{ij,l})}$$

[0025]  $r_{i,j,k}$ 表示由第i个单词和j个单词所组成词对对应的词组向量中第k个维度的值, $g_{i,j,k}$ 表示结果向量第k个维度的值,m表示结果向量的维度数量,l表示结果向量的第l个维度。

[0026] 在本发明的一些实施方式中,根据所述第一语句中单词的词性生成第二邻接张量

的步骤为,基于每个单词的词性对词对的词性进行标记,对标记相同的词对生成相同的词性向量,组合所有词对的词性向量得到第二邻接张量;

[0027] 根据所述第一语句中单词之间的句法依存类型生成第三邻接张量的步骤为,基于每个词对中两个单词在第一语句中的句法依存类型对词对进行标记,对标记相同的词对生成相同的句法向量,组合所有词对的句法向量得到第三邻接张量;

[0028] 根据所述第一语句中单词之间基于树的词对距离生成第四邻接张量的步骤为,基于每个词对中两个单词基于树的词对距离对词对进行标记,对标记距离相同的词对生成相同的距离向量,组合所有词对的距离向量得到第四邻接张量;

[0029] 根据所述第一语句中单词的相对距离生成第五邻接张量的步骤为,基于每个词对中两个单词在第一语句中的相对距离对词对进行标记,对标记相对距离相同的词对生成相同的相对距离向量,组合所有词对的相对距离向量得到第五邻接张量。

[0030] 采用上述方案,本申请能够充分地利用词之间的关系,有助于属性情感三元组的抽取。首先为属性情感三元组任务定义十种词之间关系类型,然后采用一个双仿射(Biaffine)构建第一邻接张量表示词之间的关系。之后,MMGCN可以将句子转换成一个多通道图结构,MMGCN可以学习到关系意识的节点表示。我们引入了词法和句法信息作为多种特征以增强MMGCN模型。我们设计了一种有效的用于抽取三元组的细化策略(refining strategy),当判断词对是否匹配时可以考虑到属性词和观点词抽取的隐式结果。

[0031] 在本发明的一些实施方式中,将所述隐藏状态序列分别与第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量进行图卷积,生成第一特征序列、第二特征序列、第三特征序列、第四特征序列和第五特征序列的步骤包括:

[0032] 分别提取第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量中每个通道的通道切片;

[0033] 根据图卷积分别求出每个通道切片对应的切片特征序列;

[0034] 收集第一邻接张量每个通道切片对应的切片特征序列,对第一邻接张量的所有切片特征序列进行平均池化,得到第一特征序列;

[0035] 收集第二邻接张量每个通道切片对应的切片特征序列,对第二邻接张量的所有切片特征序列进行平均池化,得到第二特征序列;

[0036] 收集第三邻接张量每个通道切片对应的切片特征序列,对第三邻接张量的所有切片特征序列进行平均池化,得到第三特征序列;

[0037] 收集第四邻接张量每个通道切片对应的切片特征序列,对第四邻接张量的所有切片特征序列进行平均池化,得到第四特征序列;

[0038] 收集第五邻接张量每个通道切片对应的切片特征序列,对第五邻接张量的所有切片特征序列进行平均池化,得到第五特征序列。

[0039] 在本发明的一些实施方式中,根据图卷积分别求出每个通道切片对应的切片特征序列,根据如下公式:

$$[0040] \quad \widetilde{H}_k = \sigma(R_{:, :, k} H W_k + b_k)$$

[0041] R表示第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量的任一个, $R_{:, :, k}$ 表示第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量的任一个在第k个通道的通道切片, $\sigma$ 表示激活函数ReLU,H表示隐藏状态序列, $W_k$



表示图卷积的权重,  $b_k$  表示图卷积的偏置,  $\widetilde{H}_k$  表示第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量的任一个在第k个通道的通道切片所对应的切片特征序列。

[0042] 在本发明的一些实施方式中, 对第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量、第五邻接张量的所有切片特征序列进行平均池化, 根据如下公式:

$$[0043] \quad \widehat{H} = f(\widetilde{H}_1, \widetilde{H}_2, \dots, \widetilde{H}_m)$$

[0044]  $\widetilde{H}_1, \widetilde{H}_2, \dots, \widetilde{H}_m$ : 分别表示第1, 2...m通道的通道切片所对应的切片特征序列,  $\widehat{H}$  表示第一特征序列、第二特征序列、第三特征序列、第四特征序列和第五特征序列中的任一个,  $f$  表示平均池化。

[0045] 在本发明的一些实施方式中, 对所述第一特征序列、第二特征序列、第三特征序列、第四特征序列和第五特征序列进行平均池化, 得到联合特征序列, 根据如下公式:

$$[0046] \quad Ha = f(\widehat{H}^1, \widehat{H}^2, \widehat{H}^3, \widehat{H}^4, \widehat{H}^5)$$

[0047]  $Ha$  表示联合特征序列,  $\widehat{H}^1$  表示第一特征序列,  $\widehat{H}^2$  表示第二特征序列,  $\widehat{H}^3$  表示第三特征序列,  $\widehat{H}^4$  表示第四特征序列,  $\widehat{H}^5$  表示第五特征序列,  $f$  表示平均池化。

[0048] 在本发明的一些实施方式中, 将所述第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量进行拼接得到联合张量, 根据如下公式:

$$[0049] \quad Ra = R^1 \oplus R^2 \oplus R^3 \oplus R^4 \oplus R^5$$

[0050]  $Ra$  表示联合张量,  $R^1$  表示第一邻接张量,  $R^2$  表示第二邻接张量,  $R^3$  表示第三邻接张量,  $R^4$  表示第四邻接张量,  $R^5$  表示第五邻接张量,  $\oplus$  表示拼接。

[0051] 在本发明的一些实施方式中, 所述联合特征序列包括对应第一语句中每个单词的第三词向量, 所述联合张量包括对应每个词对的第二词对向量, 根据联合张量和联合特征序列为每个词对生成第一词对向量, 根据如下公式:

$$[0052] \quad s_{ab} = h_a \oplus h_b \oplus r_{ab} \oplus r_{aa} \oplus r_{bb}$$

[0053]  $h_a$  表示第一语句中第a个词对应的第三词向量,  $h_b$  表示第一语句中第b个词对应的第三词向量,  $r_{ab}$  表示所述联合张量中第a个词和第b个词所组成词对所对应的第二词对向量,  $r_{aa}$  表示所述联合张量中两个第a个词所组成词对所对应的第二词对向量;  $r_{bb}$  表示所述联合张量中两个第b个词所组成词对所对应的第二词对向量,  $s_{ab}$  表示由第a个词和第b个词所组成词对所对应的第一词对向量。

[0054] 在本发明的一些实施方式中, 基于分类函数对多个第一词对向量进行处理, 得到概率分布张量的步骤还包括:

[0055] 基于所述分类函数得出第一词对向量所对应的预测词向量, 组合所有预测词向量得到概率分布张量。

[0056] 在本发明的一些实施方式中, 若分类函数为softmax函数, 则基于所述分类函数得出第一词对向量所对应的预测词向量, 根据如下公式:

$$[0057] \quad p_{ab} = \text{softmax}(W_p s_{ab} + b_p)$$

[0058]  $p_{ab}$  表示第一语句中第a个词和第b个词所组成词对所对应预测词向量,  $W_p$  表示权重

参数,  $b_p$  表示偏置, softmax表示将softmax函数作为分类函数,  $s_{ab}$  表示由第a个词和第b个词所组成词对所对应的第一词对向量。

[0059] 在本发明的一些实施方式中, 所述第一语句预设具有实际标签张量, 所述实际标签张量与概率分布张量的通道数相等, 所述实际标签张量和概率分布张量的通道均与预设的词对关系表中的关系一一对应, 所述第一语句中的词对在实际标签张量或概率分布张量的每个通道的值对应该词对在该种关系的得分, 所述第一语句的每个词对均预设具有实际关系, 所述词对在实际关系对应通道的值大于其他通道的值, 根据所述概率分布张量计算差异损失函数, 根据如下公式:

$$[0060] \quad \mathcal{L}_p = - \sum_a^n \sum_b^n \sum_{c \in \mathcal{C}} I(y_{ab} = c) \log(p_{ab|c})$$

[0061]  $\mathcal{L}_p$  表示差异损失函数的值, a表示第一语句中的第a个单词, b表示第一语句中的第b个单词, n表示第一语句中的总的单词数,  $\mathcal{C}$  表示预设的关系表中所有的关系类型, c表示关系表中的任一关系类型,  $y_{ab}$  表示由第a个单词和第b个单词所组成词对的实际关系的关系类型,  $p_{ab|c}$  表示由第a个单词和第b个单词所组成词对所对应的预测词向量在概率分布张量的c关系类型对应的通道所对应的值。

[0062] 在本发明的一些实施方式中, 基于差异损失函数计算总损失函数的步骤包括:

[0063] 分别计算所述第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量与实际标签张量的第一损失函数、第二损失函数、第三损失函数、第四损失函数和第五损失函数;

[0064] 联合差异损失函数、第一损失函数、第二损失函数、第三损失函数、第四损失函数和第五损失函数进行计算, 得到总损失函数。

[0065] 在本发明的一些实施方式中, 分别计算所述第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量与实际标签张量的第一损失函数、第二损失函数、第三损失函数、第四损失函数和第五损失函数, 根据如下公式:

$$[0066] \quad \mathcal{L}_q = - \sum_a^n \sum_b^n \sum_{c \in \mathcal{C}} I(y_{ab} = c) \log(p_{ab|c}^q)$$

[0067]  $\mathcal{L}_q$  表示第一损失函数、第二损失函数、第三损失函数、第四损失函数和第五损失函数的任一个的值, 当q=1时,  $\mathcal{L}_1$  表示第一损失函数的值, 当q=2时,  $\mathcal{L}_2$  表示第二损失函数的值, 当q=3时,  $\mathcal{L}_3$  表示第三损失函数的值, 当q=4时,  $\mathcal{L}_4$  表示第四损失函数的值, 当q=5时,  $\mathcal{L}_5$  表示第五损失函数的值, a表示第一语句中的第a个单词, b表示第一语句中的第b个单词, n表示第一语句中的总的单词数,  $\mathcal{C}$  表示预设的关系表中所有的关系类型, c表示关系表中的任一关系类型,  $y_{ab}$  表示由第a个单词和第b个单词所组成词对的实际关系的关系类型, 当q=1时,  $p_{ab|c}^q$  为  $p_{ab|c}^1$ ,  $p_{ab|c}^1$  表示由第a个单词和第b个单词所组成词对所对应的词组向量在第一邻接张量的c关系类型对应的通道所对应的值, 当q=2时,  $p_{ab|c}^q$  为  $p_{ab|c}^2$ ,  $p_{ab|c}^2$  表示由第a个单词和第b个单词所组成词对所对应的词性向量在第二邻接张量的c关系类型对应的通道

所对应的值,当 $q=3$ 时, $p_{ab|c}^q$ 为 $p_{ab|c}^3$ , $p_{ab|c}^3$ 表示由第a个单词和第b个单词所组成词对所对应的句法向量在第三邻接张量的c关系类型对应的通道所对应的值,当 $q=4$ 时, $p_{ab|c}^q$ 为 $p_{ab|c}^4$ , $p_{ab|c}^4$ 表示由第a个单词和第b个单词所组成词对所对应的距离向量在第四邻接张量的c关系类型对应的通道所对应的值,当 $q=5$ 时, $p_{ab|c}^q$ 为 $p_{ab|c}^5$ , $p_{ab|c}^5$ 表示由第a个单词和第b个单词所组成词对所对应的相对距离向量在第五邻接张量的c关系类型对应的通道所对应的值。

[0068] 在本发明的一些实施方式中,基于差异损失函数计算总损失函数,根据如下公式:

$$[0069] \quad \mathcal{L} = \alpha \mathcal{L}_p + \beta \mathcal{L}_1 + \gamma (\mathcal{L}_2 + \mathcal{L}_3 + \mathcal{L}_4 + \mathcal{L}_5)$$

[0070]  $\mathcal{L}$ 表示总损失函数的值, $\mathcal{L}_p$ 表示差异损失函数值的值, $\mathcal{L}_1$ 表示第一损失函数的值, $\mathcal{L}_2$ 表示第二损失函数的值, $\mathcal{L}_3$ 表示第三损失函数的值, $\mathcal{L}_4$ 表示第四损失函数的值, $\mathcal{L}_5$ 表示第五损失函数的值, $\alpha, \beta, \gamma$ 均表示用于调节对应损失影响的超参数。

[0071] 在本发明的一些实施方式中,根据总损失函数对所述第一模型进行训练包括对第一模型中双仿射注意力机制的参数值、对每个所述单词进行编码得到第一词向量步骤中的编码模型的参数值、多层感知机的参数值和对所述词对中分别对应两个单词的两个第二词向量进行注意力计算的步骤中注意力计算的权重参数和偏置、图卷积的权重参数和偏置值、第一模型中用于计算预测词向量的权重参数和偏置进行训练,更新各个参数的值。

[0072] 在本发明的一些实施方式中,在对第一模型进行训练时,不断修改所述第一模型中的参数值,得到多次训练中总损失函数的最小值对应的参数值,即为对第一模型训练完成时第一模型所使用的参数值。

[0073] 本发明的一个方面提供了一种属性情感三元组抽取方法,该方法包括以下步骤:

[0074] 将第二语句输入根据上述方法所训练的第一模型,得到所述第二语句所对应的概率分布张量;

[0075] 对所述第二语句所对应的概率分布张量进行三元组解码,得到第二语句中的属性情感三元组。

[0076] 本发明的附加优点、目的,以及特征将在下面的描述中将部分地加以阐述,且将对于本领域普通技术人员在研究下文后部分地变得明显,或者可以根据本发明的实践而获知。本发明的目的和其它优点可以通过在说明书以及附图中具体指出并获得。

[0077] 本领域技术人员将会理解的是,能够用本发明实现的目的和优点不限于以上具体所述,并且根据以下详细说明将更清楚地理解本发明能够实现的上述和其他目的。

## 附图说明

[0078] 此处所说明的附图用来提供对本发明的进一步理解,构成本申请的一部分,并不构成对本发明的限定。

[0079] 图1为本发明多特征多通道图卷积网络模型训练方法一种实施方式的示意图;

[0080] 图2为本发明属性情感三元组抽取方法一种实施方式的示意图;

[0081] 图3为本发明属性情感三元组抽取方法一种细化步骤的示意图;

- [0082] 图4为本发明属性情感三元组抽取方法另一种细化步骤的示意图；
- [0083] 图5为得到第二语句中的属性情感三元组的一种示例图；
- [0084] 图6为第一语句的每个词对均预设有关联关系的一种示例图；
- [0085] 图7为对所述第一语句中的词对进行标记的示意图。

### 具体实施方式

[0086] 为使本发明的目的、技术方案和优点更加清楚明白，下面结合实施方式和附图，对本发明做进一步详细说明。在此，本发明的示意性实施方式及其说明用于解释本发明，但并不作为对本发明的限定。

[0087] 在此，还需要说明的是，为了避免因不必要的细节而模糊了本发明，在附图中仅仅示出了与根据本发明的方案密切相关的结构和/或处理步骤，而省略了与本发明关系不大的其他细节。

[0088] 应该强调，术语“包括/包含”在本文使用时指特征、要素、步骤或组件的存在，但并不排除一个或多个其它特征、要素、步骤或组件的存在或附加。

[0089] 在此，还需要说明的是，如果没有特殊说明，术语“连接”在本文不仅可以指直接连接，也可以表示存在中间物的间接连接。

[0090] 在下文中，将参考附图描述本发明的实施例。在附图中，相同的附图标记代表相同或类似的部件，或者相同或类似的步骤。

[0091] 如图1所示，本发明的一个方面提供了一种多特征多通道图卷积网络模型训练方法，该方法包括以下步骤：

[0092] 步骤S100，将第一语句输入预设的第一模型中，将第一语句划分为多个单词，对每个所述单词进行编码得到第一词向量，组合多个所述第一词向量得到隐藏状态序列，基于双仿射注意力机制生成对应所述隐藏状态序列的第一邻接张量；

[0093] 在本发明的一些实施方式中，所述第一模型为多特征多通道图卷积网络模型，将第一语句划分为多个单词可以通过分词实现，分词方式可以为BERT的tokenizer分词、Jieba分词、SnowNlp分词或nltk分词等，优选为BERT的tokenizer分词，对每个所述单词进行编码得到第一词向量的编码器可以为BERT编码器，利用BERT作为句子编码器抽取每个词的隐藏状态表示。给定一个输入的第一语句 $X$ ， $w_1, w_2, \dots, w_z$ 均为 $X$ 中的单词， $X = \{w_1, w_2, \dots, w_z\}$ ，经过编码层输出隐藏状态序列 $\{A_1, A_2, \dots, A_z\}$ ，其中 $A_1$ 为对应单词 $w_1$ 的第一词向量，其中 $A_2$ 为对应单词 $w_2$ 的第一词向量， $\dots$ ，其中 $A_z$ 为对应单词 $w_z$ 的第一词向量。

[0094] 在本发明的一些实施方式中，所述第一邻接张量的通道与预设的词对关系表中的关系一一对应，预设的词对关系表可以如下表所示：

[0095]

#	关系	含义
1	B-A	属性词的开头
2	I-A	属性词的其他部分
3	A	词对 $(w_i, w_j)$ 属于同一个属性词
4	B-O	观点词的开头
5	I-O	观点词的其他部分
6	O	词对 $(w_i, w_j)$ 属于同一个观点词
7	POS	词对 $(w_i, w_j)$ 中的 $w_i$ 和 $w_j$ 分别属于属性词和观点词中的词, 并且该属性词与该观点词形成正向情感的属性-观点对
8	NEU	词对 $(w_i, w_j)$ 中的 $w_i$ 和 $w_j$ 分别属于属性词和观点词中的词, 并且该属性词与该观点词形成中性情感的属性-观点对
9	NEG	词对 $(w_i, w_j)$ 中的 $w_i$ 和 $w_j$ 分别属于属性词和观点词中的词, 并且该属性词与该观点词形成负向情感的属性-观点对
10	$\perp$	词对 $(w_i, w_j)$ 没有以上关系

[0096] B和I分别表示抽取对象的开头和除开头以外的其他部分,而A和O则能够决定抽取对象的角色是属性词还是观点词。表中的A和O关系分别用于检测词对中两个不同的词是否属于一个相同的属性词和观点词。三种情感关系用于去匹配词对是否属于属性词或观点词,并判断其情感极性。

[0097] 步骤S200,根据所述第一语句中单词的词性生成第二邻接张量,根据所述第一语句中单词之间的句法依存类型生成第三邻接张量,根据所述第一语句中单词之间基于树的词对距离生成第四邻接张量,根据所述第一语句中单词的相对距离生成第五邻接张量;

[0098] 在本发明的一些实施方式中,所述单词的词性包括形容词、否定词和情态动词等;所述句法依存类型可以依照Stanford依存句法关系,包括名词短语、动词短语和形容词短语等;所述第一语句基于树的词对距离需要预先根据Stanford依存句法关系为第一语句建立建立关系树,计算两个单词在关系树上的边的数量,当计算边的数量时,我们将关系树中的依存弧视为无向边,所述边的数量即为两个单词的词对距离;所述第一语句中单词的相对距离为单词之间的距离,若第一语句为“The gourmet food is delicious”,则The和gourmet的相对距离为1,The和food的相对距离为2,gourmet和is的相对距离为3。

[0099] 采用上述方案,通过词性、句法依存类型、句法依存类型、基于树的词对和词对相对位置多特征增强多特征多通道图卷积网络模型,提高模型对属性情感三元组提取的准确度。

[0100] 步骤S300,将所述隐藏状态序列分别与第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量进行图卷积,生成第一特征序列、第二特征序列、第三特征序列、第四特征序列和第五特征序列,对所述第一特征序列、第二特征序列、第三特征序列、第四特征序列和第五特征序列进行平均池化,得到联合特征序列;

[0101] 步骤S400,将所述第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量进行拼接得到联合张量;

[0102] 在本发明的一些实施方式中,所述第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量均存在多个通道,对所述第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量进行拼接为将上述第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量的多个通道进行叠放,若预设的词对关系表中的存在10中关系类型,第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量均设置有10个通道,则联合张量存在50个通道。

[0103] 步骤S500,将第一语句中每两个单词组合为词对,根据联合张量和联合特征序列为每个词对生成第一词对向量,基于分类函数对多个第一词对向量进行处理,得到概率分布张量;

[0104] 在本发明的一些实施方式中,本申请组合词对的方式为将第一语句中的每个单词均与第一语句中的所有单词中的每个单词进行组合,若第一语句包括单词 $w_i, w_j, w_z$ ,则与单词 $w_i$ 相关的词对包括 $(w_i, w_j), (w_i, w_i)$ 和 $(w_i, w_z)$ 。

[0105] 在本发明的一些实施方式中,所述概率分布张量的通道数量与第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量相同。

[0106] 步骤S600,根据所述概率分布张量计算差异损失函数,基于差异损失函数计算总损失函数,根据总损失函数对所述第一模型进行训练。

[0107] 本发明的多特征多通道图卷积网络模型训练方法,我们首先为属性情感三元组任务定义多种关系类型,然后采用双仿射注意力机制(Biaffine)构建第一邻接张量表示词之间的关系,多特征多通道图卷积网络(Multi-feature Multi-channel Graph Convolutional Network, MMGCN)可以将句子转换成一个多通道图结构,如本申请中的各个邻接张量,所述关系类型的数量与各个邻接张量的通道数相等,因此,MMGCN可以学习到关系意识的节点表示。此外,我们引入了词性、句法、基于树的词对距离和单词的相对距离,作为多种特征以增强MMGCN模型,充分考虑了句子中词与词之间的关系,提高三元组提取的准确度。

[0108] 现有的属性情感三元组抽取(Aspect Sentiment Triplet Extraction, ASTE)任务,为通过一种流水线型(pipeline)方法抽取属性词、观点词和情感极性所形成的三元组。流水线型方法独立地抽取三元组中的元素,这种方法会忽略掉元素之间的潜在的关联和交互,从而很可能导致错误传播和额外的计算开销。

[0109] 以往的大多数研究都是单独地处理属性词和观点词的子任务,而忽略了它们之间的依赖性;本申请从多个层面建立邻接张量,充分考虑了词与词之间的关联性,提高属性情感三元组抽取准确率。

[0110] 在本发明的一些实施方式中,基于双仿射注意力机制生成对应所述隐藏状态序列的第一邻接张量的步骤包括:

[0111] 步骤S110,利用多层感知机求得对应每个第一词向量的第二词向量;

[0112] 在本发明的一些实施方式中,多层感知器(MLP, Multilayer Perceptron)是一种前馈人工神经网络模型。

[0113] 步骤S120,对所述词对中分别对应两个单词的两个第二词向量进行注意力计算,

得到对应词对的词组向量中每个维度的值,组合所述词组向量中每个维度的值,得到词组向量;

[0114] 步骤S130,组合所述第一语句中所有词对对应的词组向量得到第一邻接张量。

[0115] 在本发明的一些实施方式中,所述词组向量的维度数与所述第一邻接张量的通道数相等,所述词组向量在第一个维度的值处于所述第一邻接张量的第一个通道上,所述第一语句的所有词组向量在第一个维度的值构成所述第一邻接张量的第一个通道。

[0116] 在本发明的一些实施方式中,利用多层感知机求得对应每个第一词向量的第二词向量,根据如下公式:

$$[0117] \quad h_i^a = \text{MLP}_a(h_i)$$

$$[0118] \quad h_j^o = \text{MLP}_o(h_j)$$

[0119]  $h_i$ 表示第i个单词对应的第一词向量, $h_j$ 表示第j个单词对应的第一词向量, $h_i^a$ 表示第i个单词对应的第二词向量, $h_j^o$ 表示第j个单词对应的第二词向量, $\text{MLP}_a$ 表示对第i个单词进行多层感知机处理, $\text{MLP}_o$ 表示对第j个单词进行多层感知机处理。

[0120] 在本发明的一些实施方式中,多层感知机 $\text{MLP}_a$ 和 $\text{MLP}_o$ 的内部参数值不同。

[0121] 在本发明的一些实施方式中,对所述词对中分别对应两个单词的两个第二词向量进行注意力计算,得到对应词对的词组向量中每个维度的值,根据如下公式:

$$[0122] \quad g_{i,j} = h_i^{aT} U_1 h_j^o + U_2 (h_i^a \oplus h_j^o) + b_1$$

[0123]  $g_{i,j}$ 表示由第i个单词和j个单词所组成词对的注意力计算的结果向量, $h_i^{aT}$ 表示第i个单词对应的第二词向量的转置, $h_j^o$ 表示第j个单词对应的第二词向量, $h_i^a$ 表示第i个单词对应的第二词向量, $U_1$ 和 $U_2$ 均表示注意力计算的权重参数, $b_1$ 表示注意力计算的偏置;

$$[0124] \quad r_{i,j,k} = \frac{\exp(g_{i,j,k})}{\sum_{l=1}^m \exp(g_{i,j,l})}$$

[0125]  $r_{i,j,k}$ 表示由第i个单词和j个单词所组成词对对应的词组向量中第k个维度的值, $g_{i,j,k}$ 表示结果向量第k个维度的值,m表示结果向量的维度数量,l表示结果向量的第1个维度。

[0126] 在本发明的一些实施方式中,根据所述第一语句中单词的词性生成第二邻接张量的步骤为,基于每个单词的词性对词对的词性进行标记,对标记相同的词对生成相同的词性向量,组合所有词对的词性向量得到第二邻接张量;

[0127] 在本发明的一些实施方式中,所述单词的词性包括限定词、常用名词、形容词和动词(现在时态,第三人称单数),将限定词标记为DT,将常用名词标记为NN,形容词标记为JJ,将动词(现在时态,第三人称单数)标记为VBZ,则若第一语句为The food is delicious,则将第一语句标记为如图7,part-of-speech combination部分所示。

[0128] 在本发明的一些实施方式中,为相同的词对随机生成词性向量,所述词性向量的维度与预设的词对关系表中的关系数量相等,组合第一语句中所有词对的词性向量得到第一语句所对应的第二邻接张量。

[0129] 根据所述第一语句中单词之间的句法依存类型生成第三邻接张量的步骤为,基于

每个词对中两个单词在第一语句中的句法依存类型对词对进行标记,对标记相同的词对生成相同的句法向量,组合所有词对的句法向量得到第三邻接张量;

[0130] 在本发明的一些实施方式中,单词之间句法依存类型包括相同词关系、决定词关系、名词主语、(命题主词与谓词间的)连系和无关系,将相同词关系标记为self,将决定词关系标记为det,将名词主语标记为nsubj,将名词主语和(命题主词与谓词间的)连系标记为cop,将无关系标记为-,则若第一语句为The food is delicious,The和food为决定词关系,food和delicious为名词主语关系,is和delicious为(命题主词与谓词间的)连系,则将第一语句标记为如图7,syntactic dependency type部分所示。

[0131] 将无关系以零向量表示。

[0132] 根据所述第一语句中单词之间基于树的词对距离生成第四邻接张量的步骤为,基于每个词对中两个单词基于树的词对距离对词对进行标记,对标记距离相同的词对生成相同的距离向量,组合所有词对的距离向量得到第四邻接张量;

[0133] 在本发明的一些实施方式中,若第一语句为The food is delicious,则建立如图7顶部所示的关系树,可知,The和food存在一条边det,则由The和food所组成词对的词对距离为1,The和delicious之间存在两条边det和nsubj,则由The和delicious所组成词对的词对距离为2,The和is之间存在三条边det、nsubj和cop,则由The和is所组成词对的词对距离为3,则将第一语句标记为如图7,tree-based distance部分所示。

[0134] 相同的词所组成的词对的词对距离为0,以零向量表示。

[0135] 根据所述第一语句中单词的相对距离生成第五邻接张量的步骤为,基于每个词对中两个单词在第一语句中的相对距离对词对进行标记,对标记相对距离相同的词对生成相同的相对距离向量,组合所有词对的相对距离向量得到第五邻接张量。

[0136] 在本发明的一些实施方式中,若第一语句为The food is delicious,则The和food的相对距离为1,The和is的相对距离为2,The和delicious的相对距离为3,则将第一语句标记为如图7,relative position distance部分所示。

[0137] 相同的词所组成的词对的相对词对距离为0,以零向量表示。

[0138] 受卷积神经网络(CNN)的启发,图卷积神经网络(GCN)是一种可以直接基于图结构进行卷积操作的CNN变体。通常的做法是利用句法依存树构建一个邻接矩阵表示该句子的图结构。

[0139] 本申请为了建模词之间的不同关系,本申请的MMGCN在原始GCN基础之上利用了双仿射构造了多通道的多个邻接张量,邻接张量中的每个通道都刻画了词与词之间的某一种关系,提高通过第一模型提取属性情感三元组的准确度。

[0140] 在本发明的一些实施方式中,将所述隐藏状态序列分别与第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量进行图卷积,生成第一特征序列、第二特征序列、第三特征序列、第四特征序列和第五特征序列的步骤包括:

[0141] 分别提取第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量中每个通道的通道切片;

[0142] 在本发明的一些实施方式中,所述通道切片为第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量或第五邻接张量单个通道的数据,所述第一邻接张量由词组向量组成,每个词组向量的第一维度的数据相组合即为第一邻接张量第一通道的数据,每个词



性向量的第一维度的数据相组合即为第二邻接张量第一通道的数据,每个句法向量的第一维度的数据相组合即为第三邻接张量第一通道的数据,每个距离向量的第一维度的数据相组合即为第四邻接张量第一通道的数据,每个相对距离向量的第一维度的数据相组合即为第五邻接张量第一通道的数据。

[0143] 根据图卷积分别求出每个通道切片对应的切片特征序列;

[0144] 收集第一邻接张量每个通道切片对应的切片特征序列,对第一邻接张量的所有切片特征序列进行平均池化,得到第一特征序列;

[0145] 收集第二邻接张量每个通道切片对应的切片特征序列,对第二邻接张量的所有切片特征序列进行平均池化,得到第二特征序列;

[0146] 收集第三邻接张量每个通道切片对应的切片特征序列,对第三邻接张量的所有切片特征序列进行平均池化,得到第三特征序列;

[0147] 收集第四邻接张量每个通道切片对应的切片特征序列,对第四邻接张量的所有切片特征序列进行平均池化,得到第四特征序列;

[0148] 收集第五邻接张量每个通道切片对应的切片特征序列,对第五邻接张量的所有切片特征序列进行平均池化,得到第五特征序列。

[0149] 在本发明的一些实施方式中,根据图卷积分别求出每个通道切片对应的切片特征序列,根据如下公式:

$$[0150] \quad \widetilde{H}_k = \sigma(R_{:, :, k} H W_k + b_k)$$

[0151] R表示第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量的任一个, $R_{:, :, k}$ 表示第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量的任一个在第k个通道的通道切片, $\sigma$ 表示激活函数ReLU,H表示隐藏状态序列, $W_k$ 表示图卷积的权重, $b_k$ 表示图卷积的偏置, $\widetilde{H}_k$ 表示第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量的任一个在第k个通道的通道切片所对应的切片特征序列。

[0152] 在本发明的一些实施方式中,对第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量、第五邻接张量的所有切片特征序列进行平均池化,根据如下公式:

$$[0153] \quad \widehat{H} = f(\widetilde{H}_1, \widetilde{H}_2, \dots, \widetilde{H}_m)$$

[0154]  $\widetilde{H}_1, \widetilde{H}_2, \dots, \widetilde{H}_m$ ;分别表示第1,2...m通道的通道切片所对应的切片特征序列, $\widehat{H}$ 表示第一特征序列、第二特征序列、第三特征序列、第四特征序列和第五特征序列中的任一个,f表示平均池化。

[0155] 在本发明的一些实施方式中,对所述第一特征序列、第二特征序列、第三特征序列、第四特征序列和第五特征序列进行平均池化,得到联合特征序列,根据如下公式:

$$[0156] \quad H a = f(\widehat{H}^1, \widehat{H}^2, \widehat{H}^3, \widehat{H}^4, \widehat{H}^5)$$

[0157] Ha表示联合特征序列, $\widehat{H}^1$ 表示第一特征序列, $\widehat{H}^2$ 表示第二特征序列, $\widehat{H}^3$ 表示第三特征序列, $\widehat{H}^4$ 表示第四特征序列, $\widehat{H}^5$ 表示第五特征序列,f表示平均池化。

[0158] 在本发明的一些实施方式中,将所述第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量进行拼接得到联合张量,根据如下公式:

$$[0159] \quad Ra = R^1 \oplus R^2 \oplus R^3 \oplus R^4 \oplus R^5$$

[0160]  $R_a$ 表示联合张量,  $R^1$ 表示第一邻接张量,  $R^2$ 表示第二邻接张量,  $R^3$ 表示第三邻接张量,  $R^4$ 表示第四邻接张量,  $R^5$ 表示第五邻接张量,  $\oplus$ 表示拼接。

[0161] 在本发明的一些实施方式中,所述联合特征序列包括对应第一语句中每个单词的第三词向量,所述联合张量包括对应每个词对的第二词对向量,根据联合张量和联合特征序列为每个词对生成第一词对向量,根据如下公式:

$$[0162] \quad s_{ab} = h_a \oplus h_b \oplus r_{ab} \oplus r_{aa} \oplus r_{bb}$$

[0163]  $h_a$ 表示第一语句中第a个词对应的第三词向量,  $h_b$ 表示第一语句中第b个词对应的第三词向量,  $r_{ab}$ 表示所述联合张量中第a个词和第b个词所组成词对所对应的第二词对向量,  $r_{aa}$ 表示所述联合张量中两个第a个词所组成词对所对应的第二词对向量;  $r_{bb}$ 表示所述联合张量中两个第b个词所组成词对所对应的第二词对向量,  $s_{ab}$ 表示由第a个词和第b个词所组成词对所对应的第一词对向量。

[0164] 在本发明的一些实施方式中,基于分类函数对多个第一词对向量进行处理,得到概率分布张量的步骤还包括:

[0165] 基于所述分类函数得出第一词对向量所对应的预测词向量,组合所有预测词向量得到概率分布张量。

[0166] 在本发明的一些实施方式中,若分类函数为softmax函数,则基于所述分类函数得出第一词对向量所对应的预测词向量,根据如下公式:

$$[0167] \quad p_{ab} = \text{softmax}(W_p s_{ab} + b_p)$$

[0168]  $p_{ab}$ 表示第一语句中第a个词和第b个词所组成词对所对应预测词向量,  $W_p$ 表示权重参数,  $b_p$ 表示偏置, softmax表示将softmax函数作为分类函数,  $s_{ab}$ 表示由第a个词和第b个词所组成词对所对应的第一词对向量。

[0169] 在本发明的一些实施方式中,  $W_p$ 表示第一模型中用于计算预测词向量的权重参数,  $b_p$ 表示第一模型中用于计算预测词向量的偏置。

[0170] 在本发明的一些实施方式中,所述第一语句预设有实际标签张量,所述实际标签张量与概率分布张量的通道数相等,所述实际标签张量和概率分布张量的通道均与预设的词对关系表中的关系一一对应,所述第一语句中的词对在实际标签张量或概率分布张量的每个通道的值对应应该词对在该种关系的得分,如图6所示,所述第一语句的每个词对均预设有实际关系,所述词对在实际关系对应通道的值大于其他通道的值,根据所述概率分布张量计算差异损失函数,根据如下公式:

$$[0171] \quad \mathcal{L}_p = - \sum_a^n \sum_b^n \sum_{c \in \mathcal{C}} I(y_{ab} = c) \log(p_{ab|c})$$

[0172]  $\mathcal{L}_p$ 表示差异损失函数的值, a表示第一语句中的第a个单词, b表示第一语句中的第b个单词, n表示第一语句中的总的单词数,  $\mathcal{C}$ 表示预设的关系表中所有的关系类型, c表示关系表中的任一关系类型,  $y_{ab}$ 表示由第a个单词和第b个单词所组成词对的实际关系的关系类型,  $p_{ab|c}$ 表示由第a个单词和第b个单词所组成词对所对应的预测词向量在概率分布张量的c关系类型对应的通道所对应的值,  $I(\cdot)$ 表示指示函数,若满足 $(\cdot)$ 的条件则函数值为1,否则为0,即c关系类型若与由第a个单词和第b个单词所组成词对的实际关系的关系类型相

同,则函数值为1,否则为0。

[0173] 采用上述方案,通过计算损失函数不断修正第一模型,提高第一模型准确度。

[0174] 在本发明的一些实施方式中,若词对  $(w_1, w_2)$  的实际关系为词对关系表中的B-0关系,若B-0关系对应实际标签张量的第四个通道,则  $(w_1, w_2)$  在实际标签张量的各个通道的值可以为  $(0, 0, 0, 1, 0, 0, 0, 0, 0, 0)$ 。

[0175] 在本发明的一些实施方式中,若词对  $(w_1, w_2)$  的实际关系为词对关系表中的B-0关系,若B-0关系对应概率分布张量的第四个通道,词对  $(w_1, w_2)$  所对应的预测词向量在概率分布张量的各个通道的值可以为  $(0.11, 0.12, 0.09, 0.95, 0.11, 0.11, 0.15, 0.12, 0.13, 0.16)$ 。

[0176] 在本发明的一些实施方式中,基于差异损失函数计算总损失函数的步骤包括:

[0177] 分别计算所述第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量与实际标签张量的第一损失函数、第二损失函数、第三损失函数、第四损失函数和第五损失函数;

[0178] 联合差异损失函数、第一损失函数、第二损失函数、第三损失函数、第四损失函数和第五损失函数进行计算,得到总损失函数。

[0179] 采用上述方案,本申请联合差异损失函数、第一损失函数、第二损失函数、第三损失函数、第四损失函数和第五损失函数进行计算,得到总损失函数,通过总损失函数修正第一模型,综合考虑了各个邻接张量的损失,进一步保证模型精准度。

[0180] 在本发明的一些实施方式中,分别计算所述第一邻接张量、第二邻接张量、第三邻接张量、第四邻接张量和第五邻接张量与实际标签张量的第一损失函数、第二损失函数、第三损失函数、第四损失函数和第五损失函数,根据如下公式:

$$[0181] \quad \mathcal{L}_q = - \sum_a^n \sum_b^n \sum_{c \in \mathcal{C}} I(y_{ab} = c) \log(p_{ab|c}^q)$$

[0182]  $\mathcal{L}_q$ 表示第一损失函数、第二损失函数、第三损失函数、第四损失函数和第五损失函数的任一个的值,当 $q=1$ 时, $\mathcal{L}_1$ 表示第一损失函数的值,当 $q=2$ 时, $\mathcal{L}_2$ 表示第二损失函数的值,当 $q=3$ 时, $\mathcal{L}_3$ 表示第三损失函数的值,当 $q=4$ 时, $\mathcal{L}_4$ 表示第四损失函数的值,当 $q=5$ 时, $\mathcal{L}_5$ 表示第五损失函数的值, $a$ 表示第一语句中的第 $a$ 个单词, $b$ 表示第一语句中的第 $b$ 个单词, $n$ 表示第一语句中的总的单词数, $\mathcal{C}$ 表示预设的关系表中所有的关系类型, $c$ 表示关系表中的任意一关系类型, $y_{ab}$ 表示由第 $a$ 个单词和第 $b$ 个单词所组成词对的实际关系的关系类型,当 $q=1$ 时, $p_{ab|c}^q$ 为 $p_{ab|c}^1$ , $p_{ab|c}^1$ 表示由第 $a$ 个单词和第 $b$ 个单词所组成词对所对应的词组向量在第一邻接张量的 $c$ 关系类型对应的通道所对应的值,当 $q=2$ 时, $p_{ab|c}^q$ 为 $p_{ab|c}^2$ , $p_{ab|c}^2$ 表示由第 $a$ 个单词和第 $b$ 个单词所组成词对所对应的词性向量在第二邻接张量的 $c$ 关系类型对应的通道所对应的值,当 $q=3$ 时, $p_{ab|c}^q$ 为 $p_{ab|c}^3$ , $p_{ab|c}^3$ 表示由第 $a$ 个单词和第 $b$ 个单词所组成词对所对应的句法向量在第三邻接张量的 $c$ 关系类型对应的通道所对应的值,当 $q=4$ 时, $p_{ab|c}^q$ 为 $p_{ab|c}^4$ , $p_{ab|c}^4$ 表示由第 $a$ 个单词和第 $b$ 个单词所组成词对所对应的距离向量在第四邻接张量

的c关系类型对应的通道所对应的值,当 $q=5$ 时, $p_{ab|c}^q$ 为 $p_{ab|c}^5$ , $p_{ab|c}^5$ 表示由第a个单词和第b个单词所组成词对所对应的相对距离向量在第五邻接张量的c关系类型对应的通道所对应的值。

[0183] 在本发明的一些实施方式中,基于差异损失函数计算总损失函数,根据如下公式:

$$[0184] \quad \mathcal{L} = \alpha \mathcal{L}_p + \beta \mathcal{L}_1 + \gamma (\mathcal{L}_2 + \mathcal{L}_3 + \mathcal{L}_4 + \mathcal{L}_5)$$

[0185]  $\mathcal{L}$ 表示总损失函数的值, $\mathcal{L}_p$ 表示差异损失函数值的值, $\mathcal{L}_1$ 表示第一损失函数的值, $\mathcal{L}_2$ 表示第二损失函数的值, $\mathcal{L}_3$ 表示第三损失函数的值, $\mathcal{L}_4$ 表示第四损失函数的值, $\mathcal{L}_5$ 表示第五损失函数的值, $\alpha, \beta, \gamma$ 均表示用于调节对应损失影响的超参数。

[0186] 在本发明的一些实施方式中,根据总损失函数对所述第一模型进行训练包括对第一模型中双仿射注意力机制的参数值、对每个所述单词进行编码得到第一词向量步骤中的编码模型的参数值、多层感知机的参数值和对所述词对中分别对应两个单词的两个第二词向量进行注意力计算的步骤中注意力计算的参数值。

[0187] 在本发明的一些实施方式中,在对第一模型进行训练时,不断修改所述第一模型中的参数值,得到多次训练中总损失函数的最小值对应的参数值,即为对第一模型训练完成时第一模型所使用的参数值。

[0188] 在本发明的一些实施方式中,对第一模型进行训练时,不断修改所述第一模型中的参数值,所述参数值包括基于双仿射注意力机制生成对应所述隐藏状态序列的第一邻接张量的步骤中多层感知机的参数值、注意力计算中的权重参数和偏置、根据图卷积分别求出每个通道切片对应的切片特征序列步骤中图卷积的权重参数和偏置值、以及分类函数中的参数值等。

[0189] 如图2所示,本发明的一个方面提供了一种属性情感三元组抽取方法,该方法包括以下步骤:

[0190] 步骤S001,将第二语句输入根据上述方法所训练的第一模型,得到所述第二语句所对应的概率分布张量;

[0191] 步骤S002,对所述第二语句所对应的概率分布张量进行三元组解码,得到第二语句中的属性情感三元组。

[0192] 采用上述方案,将第二语句输入已经训练好的第一模型,能够通过第一模型直接得到概率分布张量,通过对概率分布张量进行三元组解码,可直接得到属性情感三元组,在保证属性情感三元组的获取准确率的前提下,提高获取效率。

[0193] 如图3、4所示,图4中的 $\mathcal{L}_r$ 的即为 $\mathcal{L}_1$ ,图4中的 $\mathcal{L}_{pos}$ 的即为 $\mathcal{L}_2$ ,图4中的 $\mathcal{L}_{dep}$ 的即为 $\mathcal{L}_2$ ,图4中的 $\mathcal{L}_{tba}$ 的即为 $\mathcal{L}_4$ ,图4中的 $\mathcal{L}_{rpd}$ 的即为 $\mathcal{L}_5$ 。

[0194] 在本发明的一些实施方式中,将第二语句输入根据上述方法所训练的第一模型,执行上述步骤S100-S500,得到第二语句所对应的概率分布张量。

[0195] 在本发明的一些实施方式中,对所述第二语句所对应的概率分布张量进行三元组解码,可以为遍历所述第二语句的每个词对在各个通道的值,提取最大值所对应的通道,对照预设的关系表中该通道所对应的关系类型,若关系类型属于POS、NEU和NEG中的任一个,则输出关于该词对的三元组。

[0196] 在本发明的一些实施方式中,对所述第二语句所对应的概率分布张量进行三元组解码的步骤还包括:

[0197] 分别抽取所述第二语句中的属性词、观点词和属性情感分类;

[0198] POS表示正向情感、NEU表示中性情感和NEG表示负向情感,若 $w_1$ 为属性词, $w_2$ 为观点词, $w_1$ 对应的属性情感分类为positive,则词对 $(w_1, w_2)$ 输出的关系类型为POS,则输出属性情感三元组为 $(w_1, w_2, \text{positive})$ ;若 $w_1$ 为属性词, $w_2$ 为观点词, $w_1$ 对应的属性情感分类为neutral,若词对 $(w_1, w_2)$ 输出的关系类型为NEU,则输出属性情感三元组为 $(w_1, w_2, \text{neutral})$ ;若 $w_1$ 为属性词, $w_2$ 为观点词, $w_1$ 对应的属性情感分类为negative,若词对 $(w_1, w_2)$ 输出的关系类型为NEG,则输出属性情感三元组为 $(w_1, w_2, \text{negative})$ 。

[0199] 根据上述方法输出第二语句中所有的属性情感三元组。

[0200] 在本发明的一些实施方式中,输出属性情感三元组的集合为 $\mathcal{T} = \{(a', o', s')_{\delta}\}_{\delta=1}^{|\mathcal{T}|}$ ;  $\mathcal{T}$ 表示第二语句总共有 $|\mathcal{T}|$ 个属性情感三元组, $\delta$ 表示属性情感三元组的标号第二语句总共有3个属性情感三元组,则包括 $\{(a', o', s')_1, (a', o', s')_2, (a', o', s')_3\}$ ,其中1、2、3均为 $\delta$ 的标号,若 $a'$ 表示属性词, $o'$ 表示观点词, $s'$ 表示属性情感分类。

[0201] 如图5所示,若第二语句为The gourmet food is delicious but the service is poor,则第二语句中的属性词(ATE)包括gourmet food和service,观点词包括delicious和poor,gourmet food对应的属性情感分类为positive,service对应的属性情感分类为negative;则输出的属性情感三元组为(gourmet food, delicious, positive)和(service, poor, negative)。

[0202] 对于单词(“gourmet”, “food”), “gourmet”和“food”属于同一个属性词“gourmet food”。同时,对于词对(“food”, “delicious”), “food”是“delicious”的观点目标,并被赋予了正向的情感极性。因此,为了有效地抽取属性词“gourmet food”,使“gourmet”获得“food”的信息,“food”也获得“gourmet”的信息。为了判断属性词的情感极性,观点词“delicious”被传递到属性词“gourmet food”,能够词之间的关系学习任务依赖的词表示。

[0203] 属性词“gourmet food”和“service”都是名词,而观点词“delicious”和“poor”都是形容词,名词和形容词组成的词对能够形成属性-观点词对。其次,从图5中的关系树可以看出不同的词对之间具有不同的依存类型。这些依存类型不仅能够帮助属性词和观点词的抽取,也能帮助词对匹配的预测,词对匹配是指词和词分别属于属性词和观点词中的某个词,如果该属性词和观点词源于同一个三元组,则词对匹配。

[0204] 采用上述方案,本申请提供一种有效的细化策略,可以在判断词对是否匹配时考虑到属性词和观点词抽取的隐式结果。

[0205] 我们为ASTE任务提出了一种多特征多通道图卷积网络(MMGCN)。为了利用词之间的关系,我们首先设计了一个多通道图结构用于建模不同的关系类型。然后,我们利用图卷积操作在所有通道上进行关系意识的节点表示学习。进一步,我们考虑到词法和句法信息作为多种特征对MMGCN模型进行增强。最后,我们设计了一种细化策略用于抽取属性情感三元组,它可以在判断词对是否匹配时利用属性词和观点词抽取的隐式结果作为先验知识。

[0206] 本领域普通技术人员应该可以明白,结合本文中所公开的实施方式描述的各示例性的组成部分、系统和方法,能够以硬件、软件或者二者的结合来实现。具体究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每

个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本发明的范围。当以硬件方式实现时,其可以例如是电子电路、专用集成电路(ASIC)、适当的固件、插件、功能卡等等。当以软件方式实现时,本发明的元素是被用于执行所需任务的程序或者代码段。程序或者代码段可以存储在机器可读介质中,或者通过载波中携带的数据信号在传输介质或者通信链路上传送。

[0207] 需要明确的是,本发明并不局限于上文所描述并在图中示出的特定配置和处理。为了简明起见,这里省略了对已知方法的详细描述。在上述实施例中,描述和示出了若干具体的步骤作为示例。但是,本发明的方法过程并不限于所描述和示出的具体步骤,本领域的技术人员可以在领会本发明的精神后,作出各种改变、修改和添加,或者改变步骤之间的顺序。

[0208] 本发明中,针对一个实施方式描述和/或例示的特征,可以在一个或更多个其它实施方式中以相同方式或以类似方式使用,和/或与其他实施方式的特征相结合或代替其他实施方式的特征。

[0209] 以上所述仅为本发明的优选实施例,并不用于限制本发明,对于本领域的技术人员来说,本发明实施例可以有各种更改和变化。凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

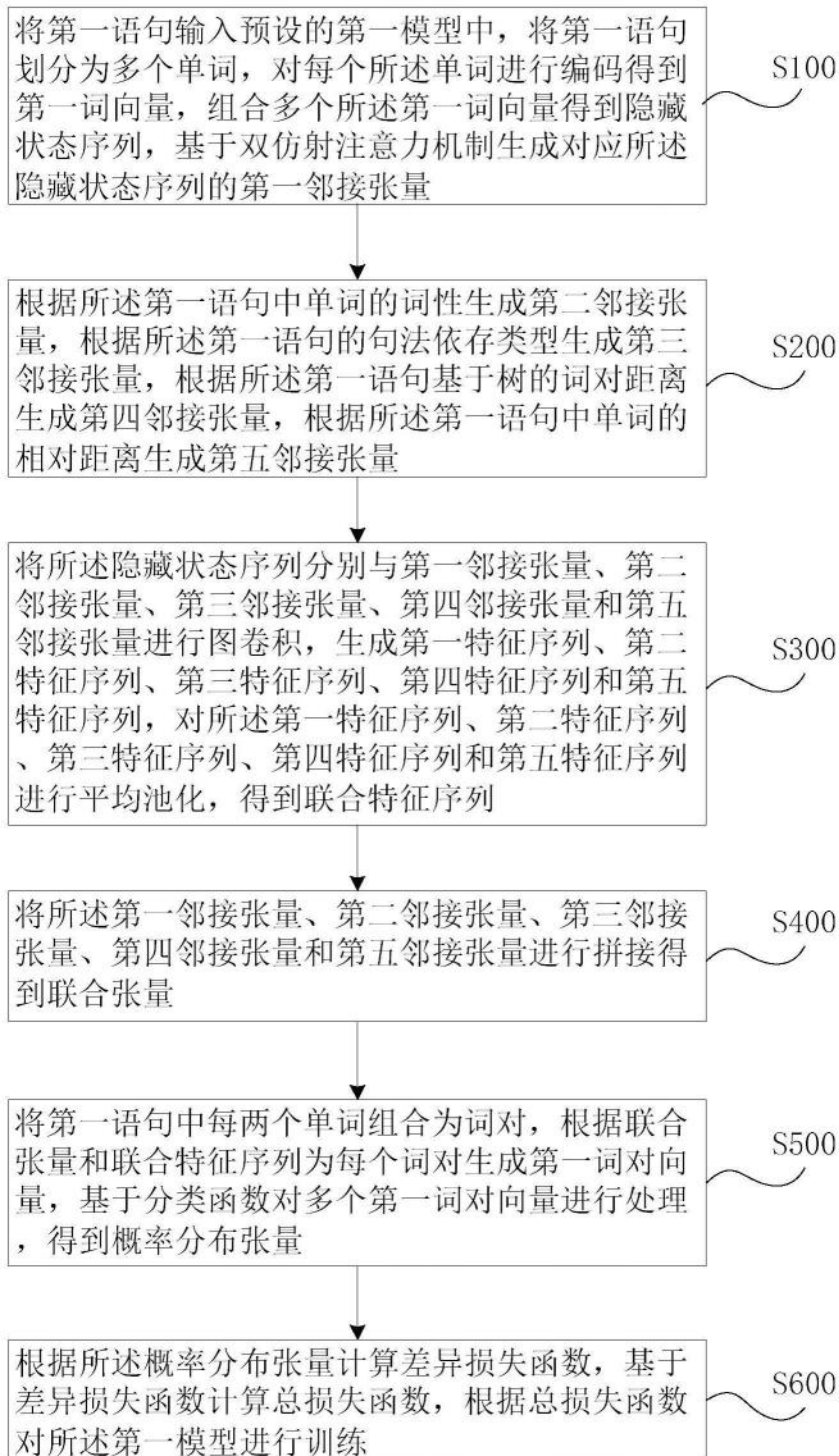


图1

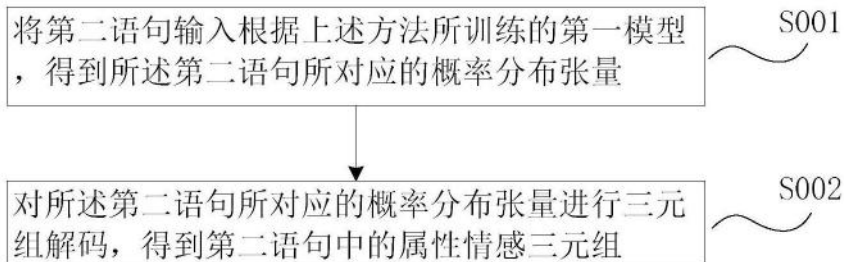


图2

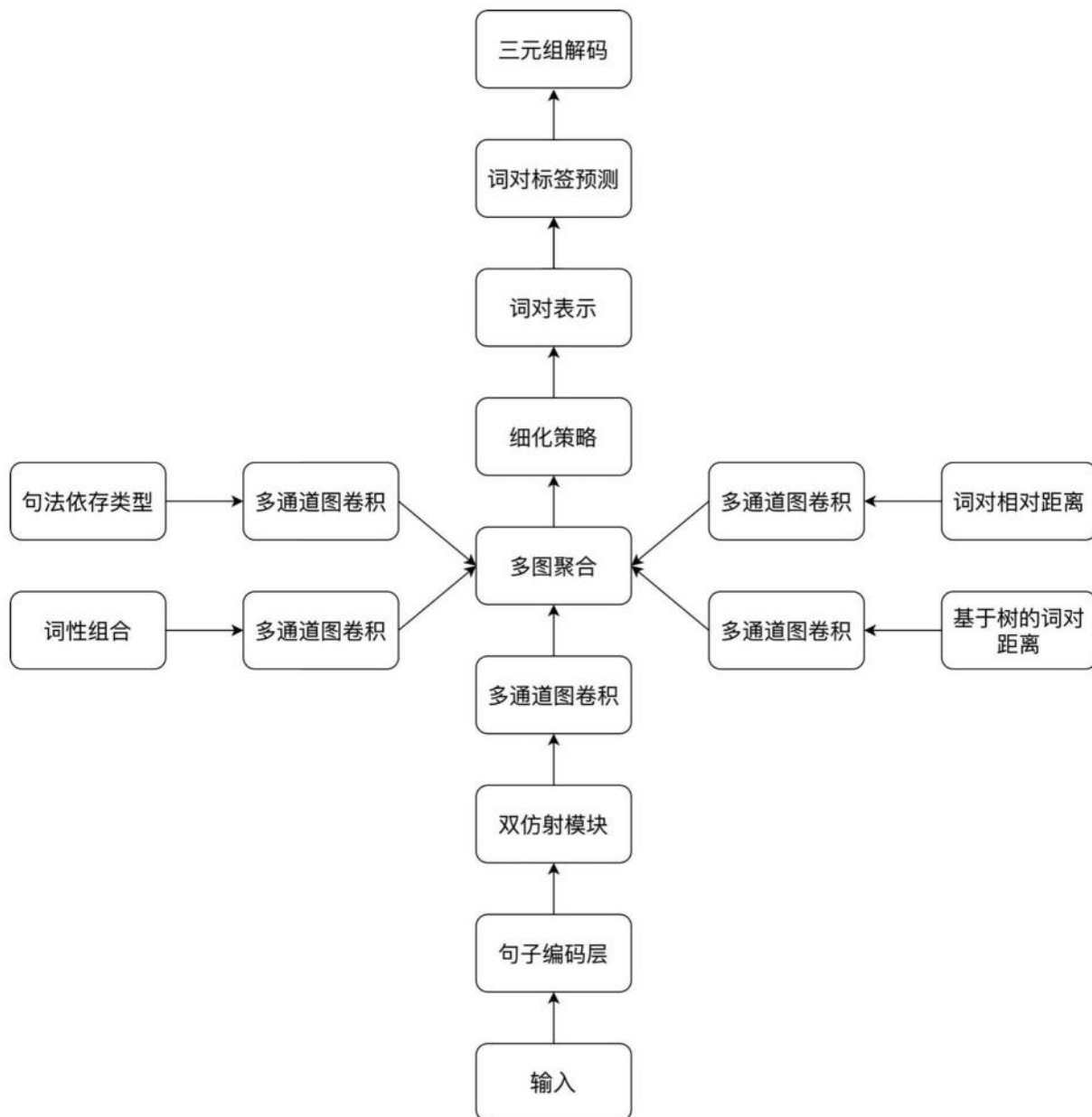


图3



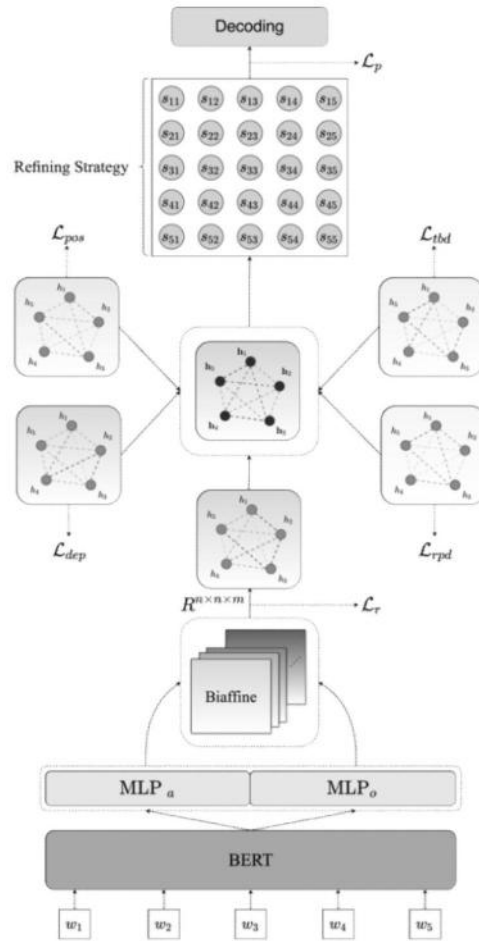


图4

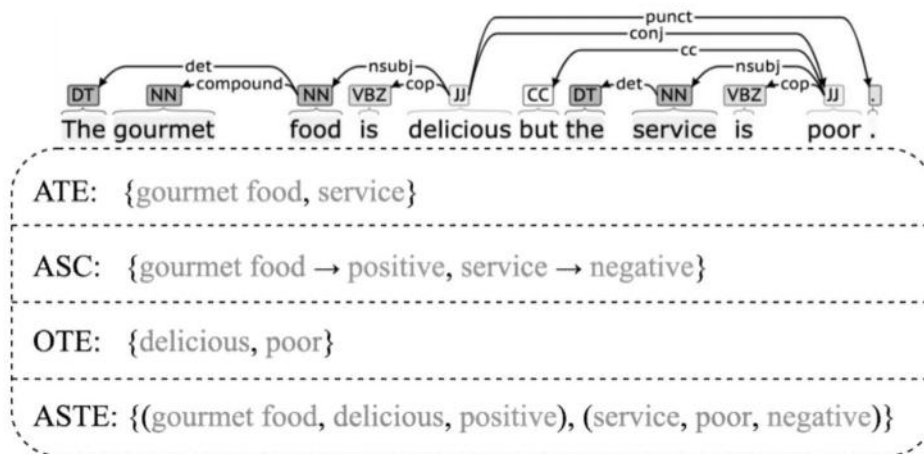


图5

	The	gourmet	food	is	delicious	but	the	service	is	poor
The	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥
gourmet	⊥	B-A	A	⊥	POS	⊥	⊥	⊥	⊥	⊥
food	⊥	A	I-A	⊥	POS	⊥	⊥	⊥	⊥	⊥
is	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥
delicious	⊥	POS	POS	⊥	B-O	⊥	⊥	⊥	⊥	⊥
but	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥
the	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥
service	⊥	⊥	⊥	⊥	⊥	⊥	⊥	B-A	⊥	NEG
is	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥
poor	⊥	⊥	⊥	⊥	⊥	⊥	⊥	NEG	⊥	B-O

图6

The food is delicious					The food is delicious				
The	(DT, DT)	(DT, NN)	(DT, VBZ)	(DT, JJ)	The	self	det	-	-
food	(DT, NN)	(NN, NN)	(NN, VBZ)	(NN, JJ)	food	det	self	-	nsubj
is	(DT, VBZ)	(NN, VBZ)	(VBZ, VBZ)	(VBZ, JJ)	is	-	-	self	cop
delicious	(DT, JJ)	(NN, JJ)	(VBZ, JJ)	(JJ, JJ)	delicious	-	nsubj	cop	self
<b>part-of-speech combination</b>					<b>syntactic dependency type</b>				
The food is delicious					The food is delicious				
The	0	1	3	2	The	0	1	2	3
food	1	0	2	1	food	1	0	1	2
is	3	2	0	1	is	2	1	0	1
delicious	2	1	1	0	delicious	3	2	1	0
<b>tree-based distance</b>					<b>relative position distance</b>				

图7