



(12) 发明专利

(10) 授权公告号 CN 114969318 B

(45) 授权公告日 2023. 04. 07

(21) 申请号 202210069686.4

(22) 申请日 2022.01.21

(65) 同一申请的已公布的文献号
申请公布号 CN 114969318 A

(43) 申请公布日 2022.08.30

(73) 专利权人 哈尔滨工业大学(深圳)(哈尔滨
工业大学深圳科技创新研究院)
地址 518055 广东省深圳市南山区桃源街
道深圳大学城哈尔滨工业大学校区
专利权人 东莞理工学院

(72) 发明人 廖清 柴合言 丁烨 方滨兴
高翠芸 王晔 王轩

(74) 专利代理机构 广州三环专利商标代理有限
公司 44202
专利代理师 郭浩辉

(51) Int. Cl.

G06F 16/35 (2019.01)

G06F 16/36 (2019.01)

G06F 40/284 (2020.01)

G06F 40/211 (2020.01)

G06N 3/0464 (2023.01)

G06N 3/08 (2023.01)

审查员 孙韬敏

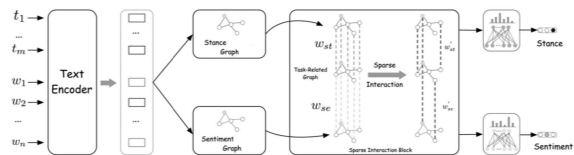
权利要求书3页 说明书11页 附图1页

(54) 发明名称

一种基于多图稀疏交互网络的多任务立场
检测方法

(57) 摘要

本发明公开了一种基于多图稀疏交互网络的多任务立场检测方法。该方法通过将输入文本输入至多图稀疏交互网络模型,得到所述输入文本的立场检测极性和情感分类极性;所述多图稀疏交互网络模型包括文本编码模块、多图构建模块、多图稀疏交互模块和任务相关注意力模块;所述多图构建模块用于构建所述多图稀疏交互网络模型的立场任务图、情感任务图和任务关系图;多图稀疏交互模块用于对立场任务图、情感任务图和任务关系图的图内节点特征进行更新,和对节点特征在图间的稀疏交互进行更新;所述任务相关注意力模块用于计算输入文本的检测立场的极性和分类情感的极性。本发明技术方案提高了针对推文文本进行立场检测的准确性。



1. 一种基于多图稀疏交互网络的多任务立场检测方法,其特征在于,

将输入文本输入至多图稀疏交互网络模型,得到所述输入文本的立场检测极性和情感分类极性;所述多图稀疏交互网络模型为图卷积神经网络模型,所述多图稀疏交互网络模型包括文本编码模块、多图构建模块、多图稀疏交互模块和任务相关注意力模块;所述输入文本包括推文文本和目标文本;

所述文本编码模块用于将输入文本处理成多个词向量,以供所述多图构建模块、多图稀疏交互模块和任务相关注意力模块使用;

所述多图构建模块用于构建所述多图稀疏交互网络模型的立场任务图、情感任务图和任务关系图;所述立场任务图为根据所述输入文本的句法依赖树构建的图和进行立场检测任务时输入文本的单词的语用权重构建的图;所述情感任务图为根据所述输入文本的句法依赖树和进行情感分类任务时输入文本的单词的语用权重构建的图;所述任务关系图为根据输入文本的单词的词类别重要性权重和所述输入文本的句法依赖树构建的图;所述构建所述多图稀疏交互网络模型的立场任务图和情感任务图,包括:根据所述推文文本的句法结构构建第一句法依赖树,并获取所述第一句法依赖树的根词集合;根据所述目标文本和所述根词集合之间的单词连接关系,将所述目标文本中的单词添加至所述第一句法依赖树中,得到输入文本的第二句法依赖树;计算所述推文文本的第一语用权重和第二语用权重,以及所述目标文本的第三语用权重,所述第一语用权重为所述推文文本的各个单词在立场检测任务中的语用权重,所述第二语用权重为所述推文文本的各个单词在情感分类任务中的语用权重;根据所述第一语用权重、第三语用权重和第二句法依赖树构建所述立场任务图,根据所述第二语用权重和第二句法依赖树构建所述情感任务图;构建所述多图稀疏交互网络模型的任务关系图,包括:构造词类别重要性权重,再根据所述词类别重要性权重构造立场标签重要性关系向量和情感标签重要性关系向量;所述词类别重要性权重为单词跟不同任务的标签类别之间的关系;根据所述立场标签重要性关系向量和情感标签重要性关系向量计算任务交互关系,并根据所述任务交互关系构造所述任务关系图;

所述多图稀疏交互模块用于对立场任务图、情感任务图和任务关系图的图内节点特征进行更新,和对节点特征在图间的稀疏交互进行更新;所述图间的稀疏交互包括立场任务图和任务关系图之间的稀疏交互,以及情感任务图和任务关系图之间的稀疏交互;

所述任务相关注意力模块用于根据所述立场任务图的立场特征表示和情感任务图的情感特征表示计算输入文本的检测立场的极性和分类情感的极性。

2. 根据权利要求1所述的基于多图稀疏交互网络的多任务立场检测方法,其特征在于,对立场任务图、情感任务图和任务关系图的图内节点特征进行更新,具体为:

根据公式 $\mathbf{g}_l^{task} = \sigma(\hat{E}_{task}^{\frac{1}{2}}(\mathcal{A}^{task} + \mathbf{I})\hat{E}_{task}^{-\frac{1}{2}}\mathbf{g}_{l-1}^{task}\mathbf{W}_{task}^l)$, 分别对所述立场任务图、情感任务图和任务关系图进行图内节点特征更新,式中 $task \in \{st, se, re\}$ 分别表示立场任务图、情感任务图和任务关系图, \mathbf{I} 是单位矩阵, σ 是非线性激活函数, l 表示当前迭代的网络层数, \mathbf{W}_{task}^l 表示第 l 层网络的参数, \mathcal{A}^{task} 表示立场任务图、情感任务图或任务关系图的邻接矩阵; $\hat{E}_{task} = \sum_j (\mathcal{A}_j^{task} + \mathbf{I})$, j 表示邻接矩阵对角线填充1。

3. 根据权利要求2所述的基于多图稀疏交互网络的多任务立场检测方法,其特征在于,所述对节点特征在图间的稀疏交互进行更新,具体为

根据以下公式:

$$\begin{aligned} g_i^{st} &= \overline{g_i^{st}} + w_i^{st} \odot z_i^{st} \cdot g_{i-1}^{re}; \\ g_i^{se} &= \overline{g_i^{se}} + w_i^{se} \odot z_i^{se} \cdot g_{i-1}^{re}; \\ g_i^{re} &= \alpha \overline{g_{i-1}^{re}} + \frac{1-\alpha}{2} g_i^{st} + \frac{1-\alpha}{2} g_i^{se}; \end{aligned}$$

对节点特征在图间的稀疏交互进行更新,所述节点特征包括立场任务图、情感任务图和任务关系图,式中 $\overline{g_i^{st}}$, $\overline{g_i^{se}}$ 和 $\overline{g_i^{re}}$ 分别表示立场任务图 g^{st} 、情感任务图 g^{se} 和任务关系图 g^{re} 使用GCN图神经网络进行图内的节点特征更新后的图的节点特征的矩阵表示, α 为超参数,用来控制任务关系图融合的立场任务图的节点信息的数量和情感任务图的节点信息的数量,1表示当前迭代的网络层数。

4. 根据权利要求3所述的基于多图稀疏交互网络的多任务立场检测方法,其特征在于,根据第一损失函数鼓励所述立场任务图和任务关系图之间、情感任务图和任务关系图之间进行稀疏交互;根据第二损失函数鼓励立场任务图和任务关系图之间、情感任务图和任务关系图之间的交互。

5. 根据权利要求4所述的基于多图稀疏交互网络的多任务立场检测方法,其特征在于,所述第一损失函数具体为:

$$\mathcal{L}_{\text{sparsity}} = \frac{\sum_{l \leq L} \log z_l^{st} + \sum_{l \leq L} \log z_l^{se}}{2};$$

所述第二损失函数具体为:

$$\mathcal{L}_{\text{sharing}} = \frac{\sum_{l \leq L} \frac{L-l}{L} |z_l^{st} - z_l^{se}| + \sum_{l \leq L} \frac{L-l}{L} |z_l^{se} - z_l^{st}|}{2};$$

式中 z_l^{st} 为立场任务的稀疏掩码矩阵, z_l^{se} 情感任务的稀疏掩码矩阵,1表示当前迭代的网络层数,L表示预设的迭代网络的总层数。

6. 根据权利要求5所述的基于多图稀疏交互网络的多任务立场检测方法,其特征在于,根据以下公式计算所述立场任务图的立场特征表示 r^{st} :

$$r^{st} = \sum_{i=1}^{m+n} \alpha_i h_i;$$

式中 α_i 为立场特征表示的注意力权重, h_i 为第i个词经过BERT模型编码后的特征向量, $m+n$ 表示所述输入文本的长度。

7. 根据权利要求6所述的基于多图稀疏交互网络的多任务立场检测方法,其特征在于,根据以下公式计算所述情感任务图的情感特征表示 r^{se} :

$$r^{se} = \sum_{i=1}^{m+n} \alpha'_i g_i^{se};$$

式中 α'_i 为情感特征表示的注意权重, g_i^{se} 为情感相关图的输出 g^{se} 中第i个节点的特征

表示, $m+n$ 表示所述输入文本的长度。

8. 根据权利要求1至7任一项所述的基于多图稀疏交互网络的多任务立场检测方法, 其特征在于, 根据以下公式计算输入文本的检测立场和分类情感的极性:

$$y^{\text{task}} = \text{softmax}(W^{\text{task}}r^{\text{task}} + b^{\text{task}});$$

式中 y^{task} 为所述多图稀疏交互网络模型预测的立场检测极性和情感分类极性, W^{task} 为全连接层的权重, b^{task} 为 W^{task} 对应的偏置, r^{task} 为立场特征表示和情感特征表示, softmax 为激活函数。

一种基于多图稀疏交互网络的多任务立场检测方法

技术领域

[0001] 本发明涉及立场检测技术领域,尤其涉及一种基于多图稀疏交互网络的多任务立场检测方法。

背景技术

[0002] 现有的立场检测方法主要使用机器学习方法和深度学习的方法。机器学习的方法需要做大量的特征工程的工作,手动提取特征,然后设计机器学习模型,来训练提取的特征,例如支持向量机模型(Support Vector Machine)、决策树模型、随机森林等。这类方法主要的缺点是需要耗费大量的时间来进行特征工程,人工挑选的特征含有的信息有限,在一定程度上降低了模型的性能;同时机器学习方法大都包含大量的超参数,需要手动挑选出其最优值,耗时耗力,而且无法大规模应用。早期的基于深度学习的立场检测方法主要使用循环神经网络(Recurrent Neural Network,RNN)、卷积神经网络(Convolutional Neural Network,CNN)和注意力机制(Attention Mechanism)来自动的捕捉新闻文本的特征,提高立场检测的表现。受到Transformers流行的影响,最近一些工作聚焦于使用Bert模型来提高立场检测任务的表现,主要是利用Bert强大的词嵌入的能力。与本发明相关的基于深度学习的立场检测方法是使用辅助信息,构造辅助任务来帮助提升立场检测任务的表现,例如情感信息,表情信息,文本的主观或者客观本质等。主要设计了辅助LSTM网络来提取情感特征,主LSTM网络来提取主任务的特征,然后将情感特征和立场特征简单地拼接在一起,预测新闻文本的立场。

[0003] 现有技术中采用基于深度学习的立场检测方法无法捕捉目标表达和立场表达的相关性,例如RNN和CNN。基于辅助任务的立场检测方法中,均简单地拼接情感特征和立场特征,忽略了任务之间关系的复杂性,将两个任务视为重要性相同,简单地将其拼接起来,从而产生较大的负迁移,造成模型性能的下降。且采用辅助任务的形式时,仅仅关注主任务的表现,对于辅助任务的性能提升没有任何考虑。因此,造成现有技术对推文文本进行立场检测时的准确率较低。

发明内容

[0004] 本发明提供一种基于多图稀疏交互网络的多任务立场检测方法,提高了针对推文文本进行立场检测的准确性。

[0005] 本发明一实施例提供一种基于多图稀疏交互网络的多任务立场检测方法,包括以下步骤:

[0006] 将输入文本输入至多图稀疏交互网络模型,得到所述输入文本的立场检测极性和情感分类极性;所述多图稀疏交互网络模型为图卷积神经网络模型,所述多图稀疏交互网络模型包括文本编码模块、多图构建模块、多图稀疏交互模块和任务相关注意力模块;所述输入文本包括推文文本和目标文本;

[0007] 所述文本编码模块用于将输入文本处理成多个词向量,以供所述多图构建模块、

多图稀疏交互模块和任务相关注意力模块使用；

[0008] 所述多图构建模块用于构建所述多图稀疏交互网络模型的立场任务图、情感任务图和任务关系图；所述立场任务图为根据所述输入文本的句法依赖树构建的图和进行立场检测任务时输入文本的单词的语用权重构建的图；所述情感任务图为根据所述输入文本的句法依赖树和进行情感分类任务时输入文本的单词的语用权重构建的图；所述任务关系图为根据输入文本的单词的词类别重要性权重和所述输入文本的句法依赖树构建的图；

[0009] 所述多图稀疏交互模块用于对立场任务图、情感任务图和任务关系图的图内节点特征进行更新，和对节点特征在图间的稀疏交互进行更新；所述图间的稀疏交互包括立场任务图和任务关系图之间的稀疏交互，以及情感任务图和任务关系图之间的稀疏交互；

[0010] 所述任务相关注意力模块用于根据所述立场任务图的立场特征表示和情感任务图的情感特征表示计算输入文本的检测立场的极性和分类情感的极性。

[0011] 进一步的，构建所述多图稀疏交互网络模型的立场任务图和情感任务图，包括以下步骤：

[0012] 根据所述推文文本的句法结构构建第一句法依赖树，并获取所述第一句法依赖树的根词集合；

[0013] 根据所述目标文本和所述根词集合之间的单词连接关系，将所述目标文本中的单词添加至所述第一句法依赖树中，得到输入文本的第二句法依赖树；

[0014] 计算所述推文文本的第一语用权重和第二语用权重，以及所述目标文本的第三语用权重，所述第一语用权重为所述推文文本的各个单词在立场检测任务中的语用权重，所述第二语用权重为所述推文文本的各个单词在情感分类任务中的语用权重；

[0015] 根据所述第一语用权重、第三语用权重和第二句法依赖树构建所述立场任务图，根据所述第二语用权重和第二句法依赖树构建所述情感任务图。

[0016] 进一步的，构建所述多图稀疏交互网络模型的任务关系图，具体为：

[0017] 构造词类别重要性权重，再根据所述词类别重要性权重构造立场标签重要性关系向量和情感标签重要性关系向量；所述词类别重要性权重为单词跟不同任务的标签类别之间的关系；

[0018] 根据所述立场标签重要性关系向量和情感标签重要性关系向量计算任务交互关系，并根据所述任务交互关系构造所述任务关系图。

[0019] 进一步的，对立场任务图、情感任务图和任务关系图的图内节点特征进行更新，具体为：

[0020] 根据公式 $\hat{g}_l^{task} = \sigma(\hat{E}_{task}^{\frac{1}{2}}(\mathcal{A}^{task} + I)\hat{E}_{task}^{-\frac{1}{2}}g_{l-1}^{task}W_{task}^l)$ ，分别对所述立场任务图、情感任务图和任务关系图进行图内节点特征更新，式中 $task \in \{st, se, re\}$ 分别表示立场任务图、情感任务图和任务关系图， I 是单位矩阵， σ 是非线性激活函数， l 表示当前迭代的网络层数， W_{task}^l 表示第 l 层网络的参数， \mathcal{A}^{task} 表示立场任务图、情感任务图或任务关系图的邻接矩阵； $\hat{E}_{task} = \sum_j (\mathcal{A}_j^{task} + I)$ ， j 表示邻接矩阵对角线填充 1。

[0021] 进一步的，所述对节点特征在图间的稀疏交互进行更新，具体为根据以下公式：

$$[0022] \quad g_i^{st} = \overline{g_i^{st}} + w_i^{st} \odot z_i^{st} \cdot g_{i-1}^{re};$$

$$[0023] \quad g_i^{se} = \overline{g_i^{se}} + w_i^{se} \odot z_i^{se} \cdot g_{i-1}^{re};$$

$$[0024] \quad g_i^{re} = \alpha \overline{g_{i-1}^{re}} + \frac{1-\alpha}{2} g_i^{st} + \frac{1-\alpha}{2} g_i^{se};$$

[0025] 对节点特征在图间的稀疏交互进行更新,所述节点特征包括立场任务图、情感任务图和任务关系图的,式中 $\overline{g_i^{st}}$, $\overline{g_i^{se}}$ 和 $\overline{g_i^{re}}$ 分别表示立场任务图 g^{st} 、情感任务图 g^{se} 和任务关系图 g^{re} 使用GCN图神经网络进行图内的节点特征更新后的图的节点特征的矩阵表示, α 为超参数用来控制任务关系图融合的立场任务图的节点信息的数量和情感任务图的节点信息的数量, l 表示当前迭代的网络层数。

[0026] 进一步的,根据第一损失函数鼓励所述立场任务图和任务关系图之间、情感任务图和任务关系图之间进行稀疏交互;根据第二损失函数鼓励立场任务图和任务关系图之间、情感任务图和任务关系图之间多交互。

[0027] 进一步的,所述第一损失函数具体为:

$$[0028] \quad \mathcal{L}_{\text{sparsity}} = \frac{\sum_{l \leq L} \log z_i^{st} + \sum_{l \leq L} \log z_i^{se}}{2};$$

[0029] 所述第二损失函数具体为:

$$[0030] \quad \mathcal{L}_{\text{sharing}} = \frac{\sum_{l \leq L} \frac{L-l}{L} |z_i^{st} - z_i^{st}| + \sum_{l \leq L} \frac{L-l}{L} |z_i^{se} - z_i^{se}|}{2};$$

[0031] 式中 z_i^{st} 为立场任务的稀疏掩码矩阵, z_i^{se} 情感任务的稀疏掩码矩阵, l 表示当前迭代的网络层数, L 表示预设的迭代网络的总层数。

[0032] 进一步的,根据以下公式计算所述立场任务图的立场特征表示 r^{st} :

$$[0033] \quad r^{st} = \sum_{i=1}^{m+n} \alpha_i h_i;$$

[0034] 式中 α 为立场特征表示的注意力权重, h_i 为第 i 个词经过BERT模型编码后的特征向量, $m+n$ 表示所述输入文本的长度。

[0035] 进一步的,根据以下公式计算所述情感任务图的情感特征表示 r^{se} :

$$[0036] \quad r^{se} = \sum_{i=1}^{m+n} \alpha'_i g_i^{se};$$

[0037] 式中 α' 为情感特征表示的注意力权重, g_i^{se} 为情感相关图的输出 g^{se} 中第 i 个节点的特征表示, $m+n$ 表示所述输入文本的长度。

[0038] 进一步的,根据以下公式计算输入文本的检测立场和分类情感的极性:

$$[0039] \quad y^{\text{task}} = \text{softmax}(W^{\text{task}} r^{\text{task}} + b^{\text{task}});$$

[0040] 式中 $y^{\text{task}}|_{\text{task}=\{\text{st}, \text{se}\}}$ 为所述多图稀疏交互网络模型预测的立场检测极性和情感分类极性, $W^{\text{task}}|_{\text{task}=\{\text{st}, \text{se}\}}$ 为全连接层的权重, b^{task} 为 W^{task} 对应的偏置, $r^{\text{task}}|_{\text{task}=\{\text{st}, \text{se}\}}$ 为立场特征表示和情感特征表示,softmax为激活函数。

[0041] 本发明的实施例,具有如下有益效果:

[0042] 本发明提供了一种基于多图稀疏交互网络的多任务立场检测方法,该方法通过采用多任务立场检测方法,以及联合新闻的情感分析任务和立场检测任务进行模型训练,利用任务之间的关系设计多图稀疏交互机制帮助任务之间的进行信息共享,从而提升立场检测的准确性。本发明通过联合训练情感分析任务和立场检测任务,为每个任务构建一个任务特定的图(即立场任务图和情感任务图)和任务相关的图(即任务关系图),并构建图之间的稀疏交互模块,实现立场任务图和情感任务图之间的稀疏交互,从而帮助任务之间进行信息共享,提升所述多图稀疏交互网络模型在各个任务上的表现力。

附图说明

[0043] 图1是本发明一实施例提供的多图稀疏交互网络模型的结构示意图。

具体实施方式

[0044] 下面将结合本发明中的附图,对本发明中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0045] 如图1所示,本发明一实施例提供的一种基于多图稀疏交互网络的多任务立场检测方法,将输入文本输入至多图稀疏交互网络模型,得到所述输入文本的立场检测极性和情感分类极性;所述多图稀疏交互网络模型为图卷积神经网络模型,所述多图稀疏交互网络模型包括文本编码模块、多图构建模块、多图稀疏交互模块和任务相关注意力模块;所述输入文本包括推文文本和目标文本;

[0046] 所述文本编码模块用于将输入文本处理成多个词向量,以供所述多图构建模块、多图稀疏交互模块和任务相关注意力模块使用;

[0047] 所述多图构建模块用于构建所述多图稀疏交互网络模型的立场任务图、情感任务图和任务关系图;所述立场任务图为根据所述输入文本的句法依赖树构建的图和进行立场检测任务时输入文本的单词的语用权重构建的图;所述情感任务图为根据所述输入文本的句法依赖树和进行情感分类任务时输入文本的单词的语用权重构建的图;所述任务关系图为根据输入文本的单词的词类别重要性权重和所述输入文本的句法依赖树构建的图;

[0048] 所述多图稀疏交互模块用于对立场任务图、情感任务图和任务关系图的图内节点特征进行更新,和对节点特征在图间的稀疏交互进行更新;所述图间的稀疏交互包括立场任务图和任务关系图之间的稀疏交互,以及情感任务图和任务关系图之间的稀疏交互;

[0049] 所述任务相关注意力模块用于根据所述立场任务图的立场特征表示和情感任务图的情感特征表示计算输入文本的检测立场的极性和分类情感的极性。

[0050] 作为其中一种实施例,构建模型的任务关系图,包括以下步骤:

[0051] 构造词类别重要性权重,再根据所述词类别重要性权重构造立场标签重要性关系向量和情感标签重要性关系向量;

[0052] 根据所述立场标签重要性关系向量和情感标签重要性关系向量计算任务交互关系,并根据所述任务交互关系构造所述任务关系图。

[0053] 作为其中一种实施例,所述文本编码模块使用BERT来编码推文文本内容C和目标

文本中的目标词T。首先将推文文本内容C和目标词T拼接起来形成输入文本S,其中 $S = \{t_1, \dots, t_m, w_1, w_2, \dots, w_n\}$ 。然后将处理成BERT模型的输入格式: $[\text{CLS}] t_1 \dots t_m w_1 w_2 \dots w_n [\text{SEP}]$ 。再将其输入到BERT网络模型以捕捉输入文本的上下文特征。其过程可由如下公式定义:

[0054] $H = \text{BERT}(S)$

[0055] 其中,H是BERT网络模型的输出, $H = \{h_1, h_2, \dots, h_{m+n}\}$ 。其中H中的每一个元素为输入文本中一个词的特征表示, $h_t \in \mathbb{R}^{2d}$ 为第t个词包含上下文信息的特征表示。

[0056] 作为其中一种实施例,构建所述立场任务图和情感任务图,包括以下步骤:

[0057] 根据所述推文文本的句法结构构建第一句法依赖树,并获取所述第一句法依赖树的根词集合;

[0058] 根据所述目标文本和所述根词集合之间的单词连接关系,将所述目标文本中的单词添加至所述第一句法依赖树中,得到输入文本的第二句法依赖树;计算所述推文文本的第一语用权重和第二语用权重,以及所述目标文本的第三语用权重,所述第一语用权重为所述推文文本的各个单词在立场检测任务中的语用权重(即利用单词与立场标签(“支持”和“反对”)在整个语料中的相对共现频率和词频计算得到词关于立场任务的语用权重);所述第二语用权重为所述推文文本的各个单词在情感分类任务中的语用权重(即利用单词与情感标签在整个语料中的相对共现频率和词频计算得到词关于立场任务的语用权重)。所述语用权重是指推文中的词对于特定的目标的依赖关系(或影响)。

[0059] 根据所述第一语用权重、第三语用权重和第二句法依赖树构建所述立场任务图,根据所述第二语用权重和第二句法依赖树构建所述情感任务图。

[0060] 具体地,使用句法分析工具对每条推文文本进行句法分析以构建句法依赖树;并通过将单词表示为图的节点(即图的节点为推文文本中的每个词的词向量),将句法依赖树中单词之间的关系表示为节点之间的连接线,以构建目标文本T和推文文本C的基图 \mathcal{G} ,包括以下步骤:

[0061] 根据所述推文文本C的句法结构构建第一句法依赖树 $\mathcal{T}^c \in \mathbb{R}^{n \times n}$,并通过句法解析器获取所述第一句法依赖树的根词集合 w^r 。

[0062] 由于目标文本T不是一个完整的句子,而是一个短语或者单词,无法被建模成一棵句法依赖树,因此,本发明实施例根据所述目标文本和所述根词集合之间的单词连接关系,将所述目标文本中的单词添加至所述第一句法依赖树中,得到所述输入文本S的第二句法依赖树 \mathcal{T} ,所述第二句法依赖树 \mathcal{T} 的计算公式如下:

$$[0063] \quad \mathcal{T}(w_i, w_j) = \begin{cases} 1 & \text{if } w_i \in T \text{ and } w_j \in w^r \\ \mathcal{T}^c(w_i, w_j) & \text{otherwise} \end{cases};$$

[0064] 式中, w^r 表示第一句法依赖树 \mathcal{T}^c 的根词, $\mathcal{T} \in \mathbb{R}^{(m+n) \times (m+n)}$ 表示输入文本S的第二句法依赖树 \mathcal{T} , w_i 和 w_j 表示推文文本和目标文本构成的输入文本S中的任意两个不相同的词。

[0065] 为了捕捉输入文本中单词的重要性和单词之间的交互特征,需要计算不同任务的单词的语用权重和单词的词频。计算所述输入文本中每个单词在整个语料库中出现的频率

$p(w_i) = \frac{N(w_i)}{N}$, 式中 $N(w_i)$ 为单词 w_i 出现在语料库中的次数, N 为语料库中所有的单词的数量。本发明实施例针对不同的任务计算单词在不同任务中的语用权重, 包括第一语用权重 $\phi^{\text{task}}(w_i) \big|_{\text{task=stance}}$ 和第二语用权重 $\phi^{\text{task}}(w_i) \big|_{\text{task=sentiment}}$, 在计算第一语用权重和第二语用权重时只考虑了有实际意义的类别来计算, 忽略了中立的立场和中性的情感这两个不包含有用信息的标签类别, 其具体计算过程如下式所示:

$$[0066] \quad \rho^{\text{task}}(w_i) = \left| \frac{N(w_i, \text{label}_+)}{N(\text{label}_+)} - \frac{N(w_i, \text{label}_-)}{N(\text{label}_-)} \right|;$$

$$[0067] \quad \phi^{\text{task}}(w_i) = \frac{\rho^{\text{task}}(w_i) - \mu(\rho^{\text{task}})}{\delta(\rho^{\text{task}})}, w_i \in C;$$

[0068] 式中 $N(w_i, \text{label}_+)$ 和 $N(w_i, \text{label}_-)$ 分别表示词 w_i 出现在立场任务标签“支持”和“反对”中的数量, 或者分别表示词 w_i 出现在情感任务标签“积极”和“消极”中的数量; $N(\text{label}_+)$ 和 $N(\text{label}_-)$ 分别表示立场任务标签“支持”和“反对”的总数量, 或者分别表示情感任务标签“积极”和“消极”的总数量, μ 为均值, δ 为标准差。

[0069] 根据公式 $\phi^{\text{task}}(s_j) = 1, s_j \in T \text{ and } s_j \notin C$, 计算所述目标文本的第三语用权重。所述目标文本的第三语用权重用于建立目标文本和推文文本之间的联系, 即构造图中目标文本和推文文本之间的边。

[0070] 根据所述第一语用权重、第三语用权重和第二句法依赖树计算所述立场任务图的第一邻接矩阵, 进而根据所述第一邻接矩阵得到所述立场任务图。

[0071] 根据所述第二语用权重和第二句法依赖树计算所述情感任务图的第二邻接矩阵, 进而根据所述第二邻接矩阵得到所情感任务图。

[0072] 具体地, 根据以下公式计算所述第一邻接矩阵和第二邻接矩阵:

$$[0073] \quad \mathcal{A}_{i,j}^{\text{task}} = \begin{cases} \sum_{t \in \{i,j\}} \phi^{\text{task}}(s_t) p(s_t) & \text{if } \mathcal{T}(s_i, s_j) = 1, s_t \in S; \\ 0 & \text{otherwise} \end{cases}$$

[0074] 式中 s_j 和 s_i 为输入文本中的单词。

[0075] 作为其中一种实施例, 构建所述多图稀疏交互网络模型的任务关系图, 包括以下步骤:

[0076] 构造词类别重要性权重, 再根据所述词类别重要性权重构造立场标签重要性关系向量和情感标签重要性关系向量;

[0077] 根据所述立场标签重要性关系向量和情感标签重要性关系向量计算任务交互关系, 并根据所述任务交互关系构造所述任务关系图。

[0078] 具体地, 考虑到不同任务之间细粒度的交互关系, 即任务在不同类别上的关系是不同的, 本发明实施例用词类别重要性权重 $\varphi^{\text{task}}(w_i, c_i)$ 来建立单词跟不同任务的标签类别之间的关系, 从而获取不同单词对于不同的任务的标签重要性关系特征, 从而帮助不同任务之间的信息交互, 以提升所述多图稀疏交互网络模型的任务表现力。根据公式

$$\varphi^{\text{task}}(w_i, c_i) = \frac{|w_i^{(c_i)}|}{|N_{c_i}|} \times \frac{|w_i^{(c_i)}|}{|w_i|}$$

是对应任务的一个标签类别 (例如, task 为立场检测任务, $c_i \in \{\text{支持}, \text{反对}\}$), $|w_i^{(c_i)}|$ 表示词 w_i 出现在任务 task 的标签类别为 c_i 的次数; $|N_{c_i}|$ 表示在任务 task 中的标签类别 c_i 对应单词出现的总数量; $|w_i|$ 表示单词 w_i 出现在语料库中的数量; task 表示立场任务或者情感任务。所述词类别重要性权重是指词对于不同任务的不同标签类别之间的重要性关系, 即表示词在两个任务之间的相似性关系, 用于捕捉两个任务中基于任务标签层面的相似性关系。先计算词对于立场任务各个标签类别之间的关系, 组成一个向量即立场标签重要性关系向量, 然后计算词对于情感任务的各个标签类别之间的关系, 组成一个向量即情感标签重要性关系向量。最后计算这两个向量的相似性, 将其作为单词对于各个任务的标签类别的重要性权重, 用于衡量单词在两个任务之间的相似性关系。

[0079] 具体地, 根据计算出来的词类别重要性权重, 为每个任务构建一个标签重要性关系向量, 包括立场标签重要性关系向量和情感标签重要性关系向量, 即 $\Phi^{\text{stance}}(w_i)$ 和 $\Phi^{\text{sentiment}}(w_i)$ 。

[0080] 所述重要性关系向量表示每个单词对于不同标签类别的重要性关系, 记作

$\hat{\Phi}^{\text{task}}(w_i): [\hat{\varphi}_{(w_i, c_1)}^{\text{task}}, \dots, \hat{\varphi}_{(w_i, c_i)}^{\text{task}}]$, 其中, task 表示立场任务或者情感任务, c_i 表示 task 的某

一个标签, $\hat{\varphi}_{(w_i, c_i)}^{\text{task}}$ 为词类别重要性权重的归一化后的形式, $\hat{\varphi}_{(w_i, c_i)}^{\text{task}} = \frac{\varphi_{(w_i, c_i)}^{\text{task}} - \mu(\varphi^{\text{task}})}{\sigma(\varphi^{\text{task}})}$,

其中 $\mu(\varphi^{\text{task}})$ 和 $\sigma(\varphi^{\text{task}})$ 是 φ^{task} 的均值和标准差。经过上述计算, 可以得到关于词在不同任务下, 对于不同的标签类别的重要性关系向量, 即立场标签重要性关系向量 $\Phi^{\text{stance}}(w_i)$ 和情感标签重要性关系向量 $\Phi^{\text{sentiment}}(w_i)$ 。

[0081] 根据所述立场标签重要性关系向量和情感标签重要性关系向量计算基于词级别的任务交互关系 $\xi(w_i)$:

$$[0082] \quad \xi(w_i) = \frac{\Phi^{\text{stance}}(w_i) \times \Phi^{\text{se}}(w_i)}{\|\Phi^{\text{stance}}(w_i)\|_2^2 + \|\Phi^{\text{se}}(w_i)\|_2^2 - \Phi^{\text{stance}}(w_i) \times \Phi^{\text{se}}(w_i)},$$

[0083] 其中, stance 表示立场检测任务, se 表示情感分类任务。根据所述输入文本 S 的第二句法依赖树 \mathcal{T} , 将任务交互关系 $\xi(w_i)$ 进行标准化处理后用于构造任务关系图的邻接矩阵 $\mathcal{A}^{\text{related}} \in \mathbb{R}^{(m+n) \times (m+n)}$, 进而根据所述邻接矩阵 $\mathcal{A}^{\text{related}}$ 得到所述任务关系图, 所述邻接矩阵 $\mathcal{A}^{\text{related}}$ 的计算公式为:

$$[0084] \quad \mathcal{A}_{i,j}^{\text{related}} = \begin{cases} \sum_{t \in \{i,j\}} \xi(s_t) & \text{if } \mathcal{T}(s_i, s_j) = 1, s_t \in S \\ 0 & \text{otherwise} \end{cases}.$$

[0085] 作为其中一种实施例, 对立场任务图、情感任务图和任务关系图的图内节点特征进行更新, 即进行横向的图内更新。具体地, 本发明实施例采用图卷积神经网络 (Graph Convolutional Neural Network, GCN) 进行图内迭代更新, 该更新过程在广义上 (只考虑图的横向更新) 对各个图来说是独立的, 即每个图独立更新节点特征。根据如下公式分别对所述立场任务图、情感任务图和任务关系图进行各自的图内节点特征更新:

$$[0086] \quad \mathbf{g}_l^{task} = \sigma(\hat{E}_{task}^{\frac{1}{2}} (\mathcal{A}^{task} + \mathbf{I}) \hat{E}_{task}^{-\frac{1}{2}} \mathbf{g}_{l-1}^{task} \mathbf{W}_{task}^l),$$

[0087] 式中 $task \in \{st, se, re\}$ 分别表示立场任务图、情感任务图和任务关系图, \mathbf{g}_{l-1}^{task} 表示第 $l-1$ 层网络输出的节点特征, \mathbf{W}_{task}^l 表示第 l 层网络的参数, \mathbf{I} 是单位矩阵, σ 是非线性

激活函数; $\hat{E}_{task} = \sum_j (\mathcal{A}_j^{task} + \mathbf{I})$, j 表示邻接矩阵对角线填充 1, \hat{E}_{task} 为图卷积网络卷积核的特征向量和常规图卷积神经网络更新的参数; l 表示当前迭代的网络层数。初始的节点特征表示为 $\mathbf{g}_0^{task} = \hat{\mathbf{H}} = [\hat{\mathbf{h}}_1, \dots, \hat{\mathbf{h}}_{(m+n)}]$ 。

[0088] 立场任务图 g^{st} 和情感任务图 g^{se} 分别与任务关系图 g^{re} 进行稀疏交互, 从而达到 g^{st} 跟 g^{se} 之间进行交互的效果, 以帮助不同任务实现信息共享。初始化 g^{st} 跟 g^{re} 的边为 w_0^{st} , g^{se} 跟 g^{re} 的边为 w_0^{se} 。定义立场任务的稀疏掩码矩阵 \mathbf{z}_l^{st} 和情感任务的稀疏掩码矩阵 \mathbf{z}_l^{se} , 其中掩码矩阵 \mathbf{z} 表示一组随机的二进制变量, 负责控制立场任务图跟任务关系图之间的交互, 以及控制立场任务图跟任务关系图之间的交互。因此, 在进行第 l 层图的迭代交互时, 根据公式 (1) - (3) 对节点特征在图间的稀疏交互进行更新, 即对立场任务图和任务关系图之间的稀疏交互, 以及情感任务图和任务关系图之间的稀疏交互进行更新:

$$[0089] \quad \mathbf{g}_l^{st} = \overline{\mathbf{g}_l^{st}} + w_l^{st} \odot \mathbf{z}_l^{st} \cdot \mathbf{g}_{l-1}^{re} \quad (1);$$

$$[0090] \quad \mathbf{g}_l^{se} = \overline{\mathbf{g}_l^{se}} + w_l^{se} \odot \mathbf{z}_l^{se} \cdot \mathbf{g}_{l-1}^{re} \quad (2);$$

$$[0091] \quad \mathbf{g}_l^{re} = \alpha \overline{\mathbf{g}_l^{re}} + \frac{1-\alpha}{2} \mathbf{g}_l^{st} + \frac{1-\alpha}{2} \mathbf{g}_l^{se} \quad (3);$$

[0092] 其中, w_l^{st} 和 w_l^{se} 分别表示立场任务图和情感任务图跟任务关系图之间连接边的参数, l 表示当前迭代的网络层数, $\overline{\mathbf{g}_l^{st}}$, $\overline{\mathbf{g}_l^{se}}$ 和 $\overline{\mathbf{g}_l^{re}}$ 分别表示立场任务图 g^{st} 、情感任务图 g^{se} 和任务关系图 g^{re} 使用 GCN 图神经网络进行图内的节点特征更新后的图的节点特征的矩阵表示, α 为超参数用来控制任务关系图融合的立场任务图的节点信息的数量和情感任务图的节点信息的数量。

[0093] 由于掩码矩阵 \mathbf{z} 表示的一组随机的二进制变量是离散型变量且不可导的, 不适用于深度学习模型。因此, 本发明实施例采用 Gumbel-Softmax 分布来将离散型随机变量 \mathbf{z} 连续化为连续型变量 \mathbf{v} , 以适应本发明实施例的多图稀疏交互网络模型的更新过程, 其计算公式如下所示:

$$[0094] \quad v_l = \frac{\exp((\log \pi_l(j) + G_l(j))/\tau)}{\sum_{i \in \{0,1\}} \exp((\log \pi_l(i) + G_l(i))/\tau)}$$

[0095] 其中, i 和 j 表示随机选取的 0 和 1, τ 是缩放参数, G_l 是从标准正态分布 Gumbel(0, 1) 中采样出来的独立同分布的样本, $\pi_l = [1 - z_l, z_l]$ 是掩码矩阵 \mathbf{z} 表示为一组随机的二进制变量。

[0096] 为了提高模型的训练效率,根据第一损失函数(即稀疏正则)鼓励立场任务图和任务关系图之间、情感任务图和任务关系图之间进行稀疏交互,所述第一损失函数为:

$$[0097] \quad \mathcal{L}_{\text{sparsity}} = \frac{\sum_{l \leq L} \log z_l^{st} + \sum_{l \leq L} \log z_l^{se}}{2}。$$

[0098] 式中, l 表示当前迭代的网络层数, L 表示设定的迭代的网络总层数为了避免图间交互太过稀疏,导致整个多图稀疏交互网络模型被分割成三个独立的图网络,再根据第二损失函数(即共享正则)鼓励立场任务图和任务关系图之间、情感任务图和任务关系图之间多交互,即鼓励底层多交互,任务之间多共享底层信息,所述第二损失函数为:

$$[0099] \quad \mathcal{L}_{\text{sharing}} = \frac{\sum_{l \leq L} \frac{L-l}{L} |z_l^{st} - z_l^{st}| + \sum_{l \leq L} \frac{L-l}{L} |z_l^{se} - z_l^{se}|}{2}。$$

[0100] 本发明实施例的多图稀疏交互模块用于两个任务之间进行信息共享,以获取对训练有帮助的信息,对于一些影响任务训练的噪声信息予以过滤,提高了任务之间共享的效率和质量,进而提升了所述多图稀疏交互网络模型执行各个任务的性能。

[0101] 作为其中一种实施例,在所述任务相关注意力模块获取立场任务图的立场特征表示 g^{st} 和情感任务图的情感特征表示 g^{se} 。对于立场检测任务来说,为了获得目标相关的立场特征表示,需要采用掩码机制过滤掉非目标词,具体地,通过设计mask矩阵,将目标词对应位置置1,非目标词对应位置置0,从而得到mask矩阵转换后的立场任务图的特征表示 $g_{\text{mask}}^{st} = \text{mask} \odot g^{st}$ 。再使用基于检索的注意力机制来获取与目标词相关的更丰富的立场特征表示,所述立场特征表示的注意力权重 α 根据下式计算得出:

$$[0102] \quad \beta_t = \sum_{i=1}^{m+n} h_t^\top g_{\text{mask}}^{st,i} = \sum_{i=1}^m h_t^\top g_{\text{mask}}^{st,i};$$

$$[0103] \quad \alpha_t = \frac{\exp(\beta_t)}{\sum_{i=1}^n \exp(\beta_i)};$$

[0104] 式中, h 为所述BERT网络模型的输出, h_t^\top 表示第 t 个词经过BERT网络模型编码后的特征表示; $m+n$ 表示输入文本中词的长度(目标文本的长度为 m ,推文文本的长度 n), $g_{\text{mask}}^{st,i}$ 表示立场相关图经过mask矩阵转换后的输出 g_{mask}^{st} 中第 i 个节点(即第 i 个词向量)的特征表示, β_t 表示第 t 个词向量的注意力权重, β_i 为第 i 个词向量的立场特征表示的注意力权重, α_t 表示第 t 个词向量的注意力权重经过归一化的注意力权重。所述立场特征表示的注意力权重表示推文中所有词对立场表征的注意力。

[0105] 再根据下式计算立场任务图的立场特征表示 r^{st} :

$$[0106] \quad r^{st} = \sum_{i=1}^{m+n} \alpha_i h_i,$$

[0107] h_i 为第 i 个词经过所述BERT模型编码后的特征向量, α_i 为第 i 个词向量的立场特征表示的注意力权重。

[0108] 同理,根据公式(4)-(6)计算情感任务图的情感特征表示 r^{se} :

$$[0109] \quad \beta'_t = \sum_{i=1}^{m+n} \mathbf{h}_t^\top \mathbf{g}_i^{se} \quad (4),$$

$$[0110] \quad \alpha'_t = \frac{\exp(\beta'_t)}{\sum_{i=1}^n \exp(\beta'_i)} \quad (5),$$

$$[0111] \quad \mathbf{r}^{se} = \sum_{i=1}^{m+n} \alpha'_i \mathbf{g}_i^{se} \quad (6),$$

[0112] \mathbf{g}_i^{se} 表示情感相关图的输出 \mathbf{g}^{se} 中第 i 个节点(或词向量)的特征表示, α' 为情感特征表示的注意力权重, β'_t 为第 t 个词向量的情感表征的注意力权重, β'_i 为第 i 个词向量的情感表征的注意力权重。

[0113] 得到最终的立场特征表示 \mathbf{r}^{st} 和情感特征表示 \mathbf{r}^{se} 后,使用一层全连接层来融合文本特征和丰富的上下文特征,并得到所述输入文本的检测立场的极性和分类情感的极性:

$$[0114] \quad \mathbf{y}^{\text{task}} = \text{softmax}(\mathbf{W}^{\text{task}} \mathbf{r}^{\text{task}} + \mathbf{b}^{\text{task}})$$

[0115] 式中, $\mathbf{y}^{\text{task}}|_{\text{task}=\{\text{st}, \text{se}\}}$ 为所述多图稀疏交互网络模型预测的立场检测极性和情感分类极性, $\mathbf{W}^{\text{task}}|_{\text{task}=\{\text{st}, \text{se}\}}$ 为全连接层的权重, \mathbf{b}^{task} 为 \mathbf{W}^{task} 对应的偏置,softmax为激活函数。

[0116] 最后,整个多图稀疏交互模块的目标函数为立场检测任务与情感分析任务的损失函数的线性组合:

$$[0117] \quad \mathcal{L}(\Theta) = \lambda_1 \sum_{d \in \mathcal{D}} y_d^{st} \log \hat{y}_d^{st} + \lambda_2 \sum_{d \in \mathcal{D}} y_d^{se} \log \hat{y}_d^{se} + \lambda_3 \|\Theta\|_2 + \mathcal{L}_{\text{sharing}} + \mathcal{L}_{\text{sparity}};$$

[0118] 式中, Θ 为所述多任务图网络模型模型的参数, λ_1 、 λ_2 、 λ_3 是对应损失项的系数; d 表示第 d 条推文, \mathcal{D} 为所有推文的集合, \hat{y}_d^{se} 为模型所预测出来的第 d 条推文的情感任务标签, \hat{y}_d^{st} 为模型所预测出来的第 d 条推文的立场任务的标签, $\mathcal{L}_{\text{sparity}}$ 表示第一损失函数(即稀疏正则), $\mathcal{L}_{\text{sharing}}$ 表示第二损失函数(即共享正则),鼓励任务共享。

[0119] 本发明实施例的任务相关注意力模块用于计算最终的所述立场任务图的立场特征表示和情感任务图的情感特征表示,并根据立场特征表示和情感特征表示来进行立场检测和情感分类。本发明实施例构造了三种图结构来捕捉任务之间的交互关系,其中构造了任务关系图来捕捉立场检测任务和情感分析任务基于词级别的相关性,从而帮助各个任务进行信息共享,减少信息共享过程中噪音的产生,本发明实施例考虑了立场检测任务和情感分析任务之间的关系会随着训练的进行不断变化,提出了一种任务图稀疏交互的方式来帮助任务之间信息共享,从而提升了所述多图稀疏交互网络模型进行立场检测任务和情感分析任务的性能。

[0120] 本发明实施例根据公众在社交平台上发布的短新闻文本预测其立场及态度。由于新闻文本内容较短,信息量少,采用单任务学习的方式很难取得较高的准确率,本发明提出一种基于多图稀疏交互网络的多任务立场检测方法,联合新闻的情感分析任务和立场检测任务进行模型训练,利用任务之间的关系设计多图稀疏交互机制帮助任务之间的进行信息

共享,从而提升立场检测的准确性。

[0121] 以上所述是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也视为本发明的保护范围。

[0122] 本领域普通技术人员可以理解实现上述实施例中的全部或部分流程,是可以通过计算机程序来指令相关的硬件来完成,所述的程序可存储于一计算机可读取存储介质中,该程序在执行时,可包括如上述各实施例的流程。其中,所述的存储介质可为磁碟、光盘、只读存储记忆体(Read-Only Memory,ROM)或随机存储记忆体(Random Access Memory,RAM)等。

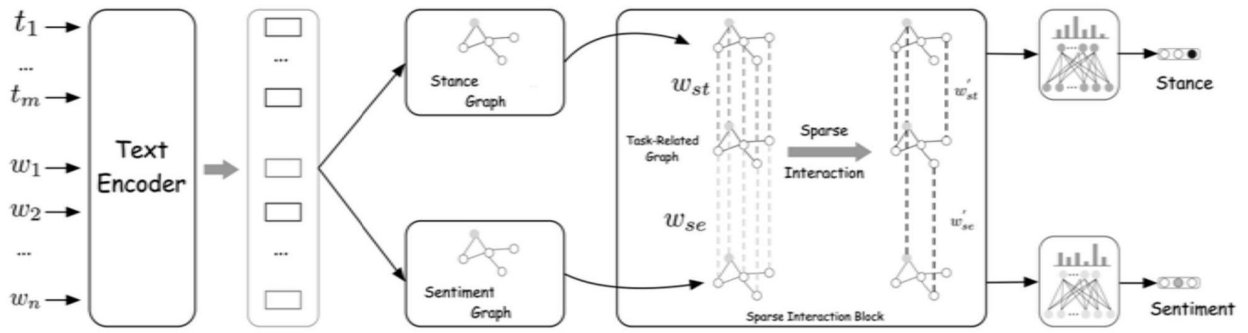


图1