(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 109408823 B
(45)授权公告日 2019.08.06

(21)申请号 201811291881.1

(22)申请日 2018.10.31

(65)同一申请的已公布的文献号
申请公布号 CN 109408823 A

(43)申请公布日 2019.03.01

(73)专利权人 华南师范大学
地址 510000 广东省广州市天河区中山大道西55号华南师范大学物理与电信工程学院

(72)发明人 袁婷 黎海辉 薛云 胡晓晖

(74)专利代理机构 广州市科丰知识产权代理事务所(普通合伙) 44467
代理人 王海曼

(51)Int.Cl.
G06F 17/27(2006.01)
G06F 16/35(2019.01)
G06Q 30/02(2012.01)

(56)对比文件
CN 107092596 A,2017.08.25,
CN 108446275 A,2018.08.24,
CN 108460089 A,2018.08.28,
CN 108399158 A,2018.08.14,
郭宝震等.采用词向量注意力机制的双路卷积神经网络句子分类模型.《浙江大学学报(工学版)》.2018,第52卷(第9期),第1729-1937页.

审查员 黄长霞

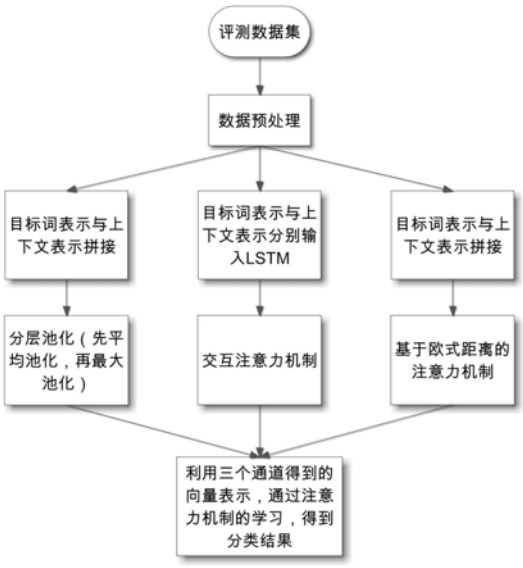
权利要求书2页 说明书13页 附图6页

(54)发明名称

一种基于多通道模型的特定目标情感分析方法

(57)摘要

本发明公开了一种基于多通道模型的特定目标情感分析方法,旨在提供一种特定目标情感分析方法,将目标词与上下文充分利用起来,本方法设置了三个通道,分别利用分层池化、交互注意力机制、基于欧式距离的注意力机制获取目标词和上下文的表示。通过三个通道,目标词和上下文可以学习到有助于情感分类的表示;其技术方案:(1)输入SemEval2014数据集,对数据集进行预处理并划分成训练集、测试集;(2)预处理后的数据分别输入到三个通道,进行特征提取以获得向量r_1、r_2、r_3、r_4和r_5;(3)利用向量r_1、r_2、r_3、r_4和r_5,通过注意力机制的学习,得到分类结果;(4)用训练好的模型对测试集中每个评论文本的特定目标进行情感分类,并与测试集本身的标签对比,以计算分类准确率;属于自然语言处理技术与情感计算领域。



CN 109408823 B

1.一种基于多通道模型的特定目标情感分析方法,其特征在于,包括以下步骤:

(1)输入评测数据集,对评测数据集进行预处理,并划分为训练集、测试集;

(2)分别输入到三个通道进行特征提取,以获得向量r_1、r_2、r_3、r_4和r_5;

(3)利用向量r_1、r_2、r_3、r_4和r_5,通过注意力机制的学习,得到分类结果;

(4)用训练好的模型对测试集中每个评论文本的特定目标进行情感分类,并与测试集本身的标签对比,计算分类准确率;

步骤(2)所述的三个通道进行特征提取获得向量r_1、r_2、r_3、r_4、r_5具体为:

(2-1)第一个通道是将上下文表示W_c与目标词表示W_t进行直接拼接,得到矩阵$W_{1,tc}$,其中$W_c \in \mathbb{R}^{n\times d_c}$、$W_t \in \mathbb{R}^{m\times d_c}$、$W_{1,tc} \in \mathbb{R}^{(m+n)\times d_c}$,m,n分别是目标词和上下文中词的个数,$d_c$是词向量维度,将$W_{1,tc}$通过LSTM得到隐含状态$H_{1,tc}$,其中$H_{1,tc} \in \mathbb{R}^{(m+n)\times d}$,d为LSTM隐藏层的维度,然后对$H_{1,tc}$进行分层池化操作得到向量$r_1$,$r_1 \in \mathbb{R}^{d}$;

(2-2)第二个通道是将W_t与W_c分别输入到LSTM,得到目标词和上下文的隐含状态$H_{2,t}$和$H_{2,c}$,$H_{2,t} \in \mathbb{R}^{m\times d}$,$H_{2,c} \in \mathbb{R}^{n\times d}$,将$H_{2,t}$和$H_{2,c}$分别进行平均池化,得到目标词和上下文的平均池化向量$t_{2,avg}$和$c_{2,avg}$,其中$t_{2,avg} \in \mathbb{R}^{d}$,$c_{2,avg} \in \mathbb{R}^{d}$,然后引入交互注意力机制,使目标信息与上下文信息充分交互,得到向量r_3,$r_3 \in \mathbb{R}^{d}$;

(2-3)第三个通道将W_t与W_c分别输入到LSTM,得到目标词和上下文对应的隐含状态$H_{3,t}^1$与$H_{3,c}^1$,其中$H_{3,t}^1 \in \mathbb{R}^{m\times d}$,$H_{3,c}^1 \in \mathbb{R}^{n\times d}$,引入基于欧式距离的注意力机制,充分利用语义信息,得到注意力机制权重矩阵$H_{3,tc}$,其中$H_{3,tc} \in \mathbb{R}^{m\times n}$,通过$H_{3,tc}$的转置与$H_{3,t}^1$相乘得到$H_{3,t}^2$,$H_{3,tc}$与$H_{3,c}^1$相乘得到$H_{3,c}^2$,其中$H_{3,t}^2 \in \mathbb{R}^{n\times d}$,为上下文对目标词的基于欧式距离注意力机制后的表示,$H_{3,c}^2 \in \mathbb{R}^{m\times d}$,为目标词对上下文的基于欧式距离注意力机制后的表示,将$H_{3,t}^2$和$H_{3,c}^2$输入到LSTM,得到输出$H_{3,t}^3$和$H_{3,c}^3$,其中$H_{3,t}^3 \in \mathbb{R}^{n\times d}$,$H_{3,c}^3 \in \mathbb{R}^{m\times d}$,将$H_{3,t}^3$、$H_{3,c}^3$进行平均池化得到$t_{3,avg}$和$c_{3,avg}$,其中$t_{3,avg} \in \mathbb{R}^{d}$,$c_{3,avg} \in \mathbb{R}^{d}$,引入交互注意力机制后,得到向量$r_4$与$r_5$,其中$r_4 \in \mathbb{R}^{d}$,$r_5 \in \mathbb{R}^{d}$。

2.根据权利要求1所述的一种基于多通道模型的特定目标情感分析方法,其特征在于,步骤(1)具体是:对特定目标情感分析的测评数据进行预处理,包括获得评论文本、特定目标及其情感极性,然后按3:1的比例将测评数据随机划分成训练集和测试集,且保证两者中积极和消极的评论数基本平衡。

3.根据权利要求1所述的一种基于多通道模型的特定目标情感分析方法,其特征在于,步骤(3)具体为:利用向量r_1、r_2、r_3、r_4和r_5,通过注意力机制的学习,得到最终表示r,$r \in \mathbb{R}^{d}$,最后将r输入到全连接神经网络当中,激活函数为softmax,得到最终的输出y,其中$y \in \mathbb{R}^{C}$,C为分类的类数;d为LSTM隐藏层的维度。

4.根据权利要求1所述的一种基于多通道模型的特定目标情感分析方法,其特征在于,

步骤(4)具体为:利用训练集训练得到模型结构参数后,便对测试集进行预测,并求出分类准确率。

5.根据权利要求3所述的一种基于多通道模型的特定目标情感分析方法,其特征在于,步骤(2-1)所述的分层池化操作是先进行平均池化,再进行最大池化。

一种基于多通道模型的特定目标情感分析方法

技术领域

[0001] 本发明属于自然语言处理技术与情感计算领域,具体地说是一种基于多通道深度学习模型的英文特定目标情感分析方法。

背景技术

[0002] 随着电子商务行业的发展,网络购物越来越受到人们的认可,由此也产生了大量的网络评论文本数据。面对这些海量的网络评论,一方面消费者需要快速了解评论的情感倾向,从其他消费者的经验中得到对该物品的评价信息,优化自己的购买决策;另一方面商家也需要从消费者的网络评论情感倾向中总结得到商品的市场反馈信息,对商品进行改善。因此,如何对评论文本进行情感分类已经成为自然语言处理领域的一个重要研究课题。

[0003] 传统的情感分类主要是为了得到句子的整体的情感倾向,当句子中含有多个目标词的时候,传统的情感分类的方法就忽略了具体每一个目标词的情感。因此对特定目标词的情感分析任务正逐渐被学术界所重视。对特定目标词的情感分类指的是当一个句子中含有多个目标词的时候,需要对不同的目标词进行情感倾向的判断。例如:The appetizers is ok,but the service is slow and the environment is bad.在这段评价中,通过分析知道,这段评论主要对三个方面进行了评价,分别是开胃菜,服务和环境。这三个目标词的情感倾向分别是积极,消极和消极。因此传统的情感分类面对这种多目标词的情况就显得不太适用。

[0004] 情感分类方法主要经历了三个阶段,第一阶段是基于词典和语言学规则的情感分类方法,这种分类方法本质上依赖于情感词典和判断规则的质量,需要人工设计,因此这类方法的优劣很大程度上都取决于人工设计和先验知识。在网络迅速发展的今天,出现了很多新词语,如:给力、坑爹,这使得词典需要实时的更新,即便如此也很难跟上时代的发展,因此基于词典的方法变得不再实用。第二阶段是基于机器学习的情感分类方法,其核心是特征提取和分类器设计,常用的特征工程包括:句法特征,(term frequency-inverse document frequency,TF-IDF)特征等;常见的分类器有决策树、贝叶斯分类器、支持向量机等。然而这些传统的情感分析方法都高度依赖于所获得的特征质量,并且特征工程的工作量也十分巨大,所以显得不太适用。因此基于深度学习的方法应运而生,深度学习的方法首先通过海量的样本学习到词的表示即词向量,之后将词向量作为神经网络的输入,经过多层网络提取出特征,最后通过全连接得到分类结果。

[0005] 近年来,基于深度学习实现特定目标词情感分析的方法正逐步受到认可。Tang以目标词为中心点,将句子划分为左右两部分,并用两个独立的LSTM网络分别对这两部分的句子进行建模,从而能够充分地利用目标词的语义信息,捕获到了目标词和上下文之间的关系,体现了LSTM在文本任务中能力,但是这种无差别的操作不能体现不同词对情感分类的差异性;Huang则将目标词和上下文中每一个词语分别进行拼接后输入LSTM网络,获得隐藏层输出后,再使用注意力机制确定给定目标对句子的影响,这种方式通过注意力机制将主要的信息加大权重,有助于后面的情感分析;Liu将句子分成了三部分,分别是目标词,目

标词的上文和目标词下文,将三部分分别输入到LSTM之后得到隐藏层的输出,然后再通过注意力机制得到输出向量,最后通过全连接得到情感分析结果;Ma则通过LSTM网络对目标词和上下文分别进行建模,再利用各自隐藏层的输出进行交互注意力机制的操作,Ma提出的交互注意力机制,即将目标词和上下文进行交互,然后通过注意力机制进行权重的选择。实际上,上述工作中都利用了目标词和上下文,而Ma则是将目标词和上下文尽可能的交互,从而取得了更好的分类效果。但是Ma在交互过程中,求目标词和上下文整体向量时采用了平均池化操作,忽略了不同词对于情感分类的影响。Shen将词向量先经过平均池化,再进行最大池化(简称分层池化)操作得到向量,并且通过大量的实验证明该方法在一定程度上比向量直接输入到LSTM中效果更好。Yin提出基于欧式距离的注意力机制,通过实验证明了该模型的有效性。虽然上述方法都取得了一定效果,但是这些方法获得的信息都不够全面,从而导致分析的结果不够可靠。

发明内容

[0006] 在针对评论文本的特定目标情感分类任务上,为了克服现有技术存在的上述不足,本发明提出一种基于多通道模型的特定目标情感分析方法;本申请提供的技术方案为了尽可能利用目标词与上下文,设置了三个通道,分别利用了分层池化,交互注意力机制以及基于欧式距离的注意力机制。

[0007] 为此,本发明通过的技术方案如下:

[0008] 一种基于多通道模型的特定目标情感分析方法,包括以下步骤:

[0009] (1)获取SemEval 2014评测数据集,对评测数据集进行预处理,并将其划分成训练集、测试集;

[0010] (2)分别通过三个通道进行特征提取,获得向量r_1、r_2、r_3、r_4和r_5;

[0011] (3)利用向量r_1、r_2、r_3、r_4和r_5,通过注意力机制的学习,得到分类结果;

[0012] (4)用训练好的模型对测试集中每个评论文本的特定目标进行情感分类,得到分类结果,并与测试集本身的标签对比,计算分类准确率。

[0013] 上述的一种基于多通道模型的特定目标情感分析方法,步骤(1)具体是:对特定目标情感分析的测评数据进行预处理,包括获得评论文本、特定目标及其情感极性,然后按3:1的比例将测评数据随机划分成训练集和测试集,并保证两者中积极和消极的评论数基本平衡;

[0014] 上述的一种基于多通道模型的特定目标情感分析方法,步骤(2)具体是:使用斯坦福大学公开的300维glove词向量与输入文本中的词进行匹配,使得文本中的每个词都能对应得到的300维向量,对于没有匹配到的词,则在[-0.1,0.1]中随机取值后得到词向量,作为模型的输入并分别进入到三个通道;

[0015] 其中三个通道分别为:

[0016] (2-1)第一个通道是将上下文表示W_c与目标表示W_t进行直接拼接,得到矩阵$W_{1,tc}$,其中$W_c \in \mathbb{R}^{n\times d_e}$、$W_t \in \mathbb{R}^{m\times d_e}$、$W_{1,tc} \in \mathbb{R}^{(m+n)\times d_e}$,m,n分别是目标词和上下文中词的个数,$d_e$是词向量维度,将$W_{1,tc}$通过LSTM得到隐含状态$H_{1,tc}$,其中$H_{1,tc} \in \mathbb{R}^{(m+n)\times d}$,d为LSTM隐藏层的维

度，然后对$H_{1,tc}$进行分层池化操作，得到向量r_1，$r_1 \in \mathbb{R}^d$；

[0017] 更进一步的，第一个通道进行特征提取获得向量r_1的方法，包括以下步骤：

[0018] 1)将W_t与W_c进行直接拼接，得到矩阵$W_{1,tc}$；

[0019] 2)LSTM网络中分别设计了输入门、遗忘门、输出门来控制信息的输入，保留和输出。其前向传播公式如下所示：(下面公式有乱码)

[0020] $i^t=\sigma W_i \cdot [h^{t-1};w^t]+b_i$

[0021] $f^t=\sigma W_f \cdot [h^{t-1};w^t]+b_f$

[0022] $o^t=\sigma W_o \cdot [h^{t-1};w^t]+b_o$

[0023] $g^t=\tanh\ W_r \cdot [h^{t-1};w^t]+b_r$

[0024] $c^t=i^t \odot g^t+f^t \odot c^{t-1}$

[0025] $h^t=o^t \odot \tanh\ c^t$

[0026] 其中i^t，f^t，o^t分别表示的是输入门，遗忘门，输出门。h^{t-1}为上个细胞单元的输出，g^t为当前细胞输入的状态，c^t和h^t分别为当前细胞状态和隐藏层输出，w^t为t时刻的输入向量，W_i，W_f，W_o，W_r为参数矩阵，b_i，b_f，b_o，b_r为偏置，⊙为点乘，σ为softmax激活函数；

[0027] 将拼接后的矩阵作为LSTM的输入，得到隐含状态$H_{1,tc} \in \mathbb{R}^{(m+n)\times d}$；

[0028] 3)进行分层池化，即先进行平均池化，然后再进行最大池化，在实验数据集SemEval2014Task4中，使用的平均池化窗口是8×d，然后得到平均池化后的矩阵，对整个矩阵采用最大池化得到向量r_1；

[0029] (2-2)第二个通道是将W_t与W_c分别输入到LSTM，得到目标词和上下文的隐含状态$H_{2,t}$和$H_{2,c}$，其中$H_{2,t} \in \mathbb{R}^{m\times d}$，$H_{2,c} \in \mathbb{R}^{n\times d}$，将$H_{2,t}$和$H_{2,c}$分别进行平均池化操作，得到目标词和上下文的平均池化向量$t_{2,avg}$和$c_{2,avg}$，$t_{2,avg} \in \mathbb{R}^d$，$c_{2,avg} \in \mathbb{R}^d$，然后引入交互注意力机制，使目标信息与上下文信息充分交互，得到向量r_3，$r_3 \in \mathbb{R}^d$；

[0030] 更进一步的，第二个通道进行特征提取获得向量r_2和向量r_3的方法，包括以下步骤：

[0031] 1)将W_t与W_c分别输入到LSTM，得到目标词和上下文的隐含状态$H_{2,t}$和$H_{2,c}$；

[0032] 2)将$H_{2,c}$进行平均池化，如下公式所示，得到上下文的平均池化向量$c_{2,avg}$；

[0033] $$c_{2,avg}=\sum_{i=1}^{n} h_{2,c}^{i} / n$$

[0034] 其中$h_{2,c}^{i}$是$H_{2,c}$中的行向量，i∈[1,n]，因此$c_{2,avg}$包含了$H_{2,c}$的信息；

[0035] 3)将$H_{2,t}$进行平均池化，如下公式所示，得到目标词的平均池化向量$t_{2,avg}$；

[0036] $$t_{2,avg}=\sum_{j=1}^{m} h_{2,t}^{j} / \mathrm{m}$$

[0037] 其中$h_{2,t}^{j}$是$H_{2,t}$中的行向量，j∈[1,m]，因此$t_{2,avg}$包含了$H_{2,t}$的信息；

[0038] 4)利用$c_{2,avg}$与$H_{2,t}$的第j个行向量$h_{2,t}^{j}$，通过交互学习，得到$\gamma(h_{2,t}^{j},c_{2,avg})$，公式如下所示：

[0039] $\gamma(h_{2,t}^{j},c_{2,avg})=\tanh(h_{2,t}^{j}\cdot W_{2,b}\cdot c_{2,avg}^{T}+b_{2,b})$

[0040] 其中$W_{2,b}$是交互学习的参数矩阵，维度为$\mathbb{R}^{d\times d}$，$h_{2,t}^{j}$是$H_{2,t}$的一个行向量，$c_{2,avg}^{T}$为$c_{2,avg}$的转置，$b_{2,b}$为偏置；

[0041] 5)对每个$\gamma(h_{2,t}^{j},c_{2,avg})$进行归一化，求得对应$H_{2,t}$第j个行向量$h_{2,t}^{j}$的系数$\beta_j$，公式如下所示：

[0042] $$\beta_j=\frac{\exp(\gamma(h_{2,t}^{j},c_{2,avg}))}{\sum_{k=1}^{m}\exp(\gamma(h_{2,t}^{k},c_{2,avg}))}$$

[0043] 其中k、j∈[1:m]；

[0044] 6)将β_j与$H_{2,t}$的第j个特征向量$h_{2,t}^{j}$相乘，加权求和的结果即为采用注意力机制之后得到的向量r_2，公式如下所示：

[0045] $$r_2=\sum_{j=1}^{m}\beta_j h_{2,t}^{j}$$

[0046] 7)同理，利用$t_{2,avg}$与$H_{2,c}$的第i个行向量$h_{2,c}^{i}$，通过交互学习，得到向量r_3，原理与4)—6)类似，这里不再重复，公式如下所示：

[0047] $$\gamma(h_{2,c}^{i},t_{2,avg})=\tanh(h_{2,c}^{i}\cdot W_{2,a}\cdot t_{2,avg}^{T}+b_{2,a})$$

[0048] $$\alpha_i=\frac{\exp(\gamma(h_{2,c}^{i},t_{2,avg}))}{\sum_{l=1}^{n}\exp(\gamma(h_{2,c}^{l},t_{2,avg}))}$$

[0049] $$r_3=\sum_{i=1}^{n}\alpha_i h_{2,c}^{i}$$

[0050] 其中l、i∈[1:n]；

[0051] (2-3)第三个通道将W_t与W_c分别输入到LSTM，得到目标词和上下文的隐含状态$H_{3,t}^{1}$与$H_{3,c}^{1}$，其中$H_{3,t}^{1}\in\mathbb{R}^{m\times d}$，$H_{3,c}^{1}\in\mathbb{R}^{n\times d}$，引入基于欧式距离的注意力机制，充分利用语义信息，得到注意力机制权重矩阵$H_{3,tc}$，其中$H_{3,tc}\in\mathbb{R}^{m\times n}$，通过$H_{3,tc}$的转置与$H_{3,t}^{1}$相乘得到$H_{3,t}^{2}$，$H_{3,tc}$与$H_{3,c}^{1}$相乘得到$H_{3,c}^{2}$，其中$H_{3,t}^{2}\in\mathbb{R}^{n\times d}$，为上下文对目标词的基于欧式距离注意力机制后的表示，$H_{3,c}^{2}\in\mathbb{R}^{m\times d}$为目标词对上下文的基于欧式距离注意力机制后的表示，将$H_{3,t}^{2}$和$H_{3,c}^{2}$输入到LSTM，得到隐含状态$H_{3,t}^{3}$和$H_{3,c}^{3}$，其中$H_{3,t}^{3}\in\mathbb{R}^{n\times d}$，$H_{3,c}^{3}\in\mathbb{R}^{m\times d}$，将$H_{3,t}^{3}$、$H_{3,c}^{3}$进行平均池化得到$t_{3,avg}$和$c_{3,avg}$，其中$t_{3,avg}\in\mathbb{R}^{d}$，$c_{3,avg}\in\mathbb{R}^{d}$，引入交互注意力机制后，得到向量$r_4$与$r_5$，其中$r_4\in\mathbb{R}^{d}$，$r_5\in\mathbb{R}^{d}$；

[0052] 更进一步的，第三个通道进行特征提取获得r_4和r_5的方法，包括以下步骤：

[0053] 1)将W_t和W_c分别输入到LSTM中，得到$H_{3,t}^{1}$，$H_{3,c}^{1}$；

[0054] 2)引入基于欧式距离的注意力机制，充分利用语义信息，得到注意力机制权重矩阵$H_{3,tc}$，计算公式如下所示：

[0055] $$h_{3,tc}^{i,j}=\frac{1}{1+|h_{3,c}^{1,i}-h_{3,t}^{1,j}|}$$

[0056] 其中$h_{3,tc}^{i,j}$为$H_{3,tc}$中的第i行第j列元素，$h_{3,c}^{1,i}$为上下文特征矩阵$H_{3,c}^{1}$中的第i个行向量，$h_{3,t}^{1,j}$为目标特征矩阵$H_{3,t}^{1}$中的第j个行向量，维度为d，$|h_{3,c}^{1,i}-h_{3,t}^{1,j}|$为两个向量的欧式距离，加1操作是为了防止两个完全一样的向量导致分母为0；其意义为距离较近的两个向量之间相互影响较大，则注意力机制权重较大，反之，距离较远的两个向量之间相互影响较小，则注意力机制权重较小；

[0057] 3)$H_{3,tc}$与$H_{3,t}^{1}$相乘得到基于欧式距离的注意力机制后的表示$H_{3,t}^{2}$，公式如下所示：

[0058] $$H_{3,t}^{2}=H_{3,tc}\times H_{3,t}^{1}$$

[0059] 其中$H_{3,t}^{1}\in\mathbb{R}^{m\times d}$，$H_{3,tc}\in\mathbb{R}^{n\times m}$，$H_{3,t}^{2}\in\mathbb{R}^{n\times d}$

[0060] 4)$H_{3,tc}$与$H_{3,c}^{1}$相乘得到基于欧式距离的注意力机制后的表示$H_{3,c}^{2}$，公式如下所示：

[0061] $$H_{3,c}^{2}=H_{3,tc}^{T}\times H_{3,c}^{1}$$

[0062] 其中$H_{3,c}^{1}\in\mathbb{R}^{n\times d}$，$H_{3,tc}\in\mathbb{R}^{n\times m}$，$H_{3,c}^{2}\in\mathbb{R}^{m\times d}$

[0063] 5)将$H_{3,t}^{2}$，$H_{3,c}^{2}$输入到LSTM进一步提取特征，且得到输出为$H_{3,t}^{3}$、$H_{3,c}^{3}$，$H_{3,t}^{3}$、$H_{3,c}^{3}$的维度分别与$H_{3,t}^{2}$，$H_{3,c}^{2}$的维度一致；

[0064] 6)将$H_{3,t}^{3}$进行平均池化，得到目标信息的平均池化向量$t_{3,avg}$，公式如下所示：

[0065] $$t_{3,avg}=\sum_{i=1}^{n}h_{3,t}^{3,i}/n$$

[0066] 其中$h_{3,t}^{3,i}$为矩阵$H_{3,t}^{3}$的行向量，$i\in[1:n]$，$t_{3,avg}\in\mathbb{R}^{d}$

[0067] 7)利用$t_{3,avg}$与$H_{3,c}^{3}$，通过交互学习，得到$\gamma(h_{3,c}^{3,j},t_{3,avg})$，公式如下所示：

[0068] $$\gamma(h_{3,c}^{3,j},t_{3,avg})=\tanh(h_{3,c}^{3,j}\cdot W_{3,a}\cdot t_{3,avg}^{T}+b_{3,a})$$

[0069] 其中$W_{3,a}$为交互学习的参数矩阵，维度为$R^{d\times d}$，$b_{3,a}$为偏置项；

[0070] 8)对每一个$\gamma(h_{3,c}^{3,j},t_{3,avg})$进行归一化，得到$H_{3,c}^{3}$第j个行向量$h_{3,c}^{3,j}$的注意力机制权重系数$\alpha_j$，公式如下所示：

[0071] $$\alpha_j=\frac{\exp(\gamma(h_{3,c}^{3,j},t_{3,avg}))}{\sum_{k=1}^{m}\exp(\gamma(h_{3,c}^{3,k},t_{3,avg}))}$$

[0072] 其中$j,k\in[1:m]$；

[0073] 9)将α_j与$H_{3,t}^{3}$第j个行向量$h_{3,c}^{3,j}$相乘，加权求和得到向量r_4，公式如下所示：

[0074] $$r_4=\sum_{i=1}^{m}\alpha_j h_{3,c}^{3,j}$$

[0075] 10)将$H_{3,c}^{3}$进行平均池化，得到上下文信息的平均池化向量$c_{3,avg}$，具体如下公式所

示；

[0076] $$c_{3,avg}=\sum_{j=1}^{m}h_{3,c}^{3,j}/\mathrm{m}$$

[0077] 其中$h_{3,c}^{3,j}$为矩阵$H_{3,c}^{3}$的第j个行向量，j∈[1:m]。

[0078] 11)同理，利用$c_{3,avg}$与$H_{3,t}^{3}$第i个行向量$h_{3,t}^{3,i}$，通过交互学习，得到向量r_5，原理与8)—10)类似，这里不再重复，公式如下所示：

[0079] $$\gamma(h_{3,t}^{3,i},c_{3,avg})=\tanh(h_{3,t}^{3,i}\cdot W_{3,b}\cdot {c_{3,avg}}^{T}+b_{3,b})$$

[0080] $$\beta_i=\frac{\exp(\gamma(h_{3,t}^{3,i},c_{3,avg}))}{\sum_{l=1}^{n}\exp(\gamma(h_{3,t}^{3,l},c_{3,avg}))}$$

[0081] $$r_5=\sum_{i=1}^{n}\beta_i h_{3,t}^{3,i}$$

[0082] 其中$W_{3,b}$为交互学习的参数矩阵，维度是$\mathbb{R}^{d\times d}$，$b_{3,b}$为偏置项，i,l∈[1:n]；

[0083] 上述的一种基于多通道模型的特定目标情感分析方法，步骤(3)利用向量r_1、r_2、r_3、r_4和r_5，通过注意力机制的学习，得到r，公式如下所示：

[0084] $$c_p=\frac{\exp(r_p w_p^T)}{\sum_{z=1}^{5}\exp(r_z w_z^T))}$$

[0085] $$r=\sum_{p=1}^{5}c_p r_p$$

[0086] 其中r_p∈[r_1,r_2,r_3,r_4,r_5]，w_p^T和w_z^T为参数向量，$w_p^T\in\mathbb{R}^{d\times 1}$，$w_z^T\in\mathbb{R}^{d\times 1}$，$c_p$为注意力机制系数；

[0087] 最后将r经过全连接层，并使用softmax作为激活函数，得到最终的分类结果。

[0088] 上述的一种基于多通道模型的特定目标情感分析方法，步骤(4)具体为：用训练好的模型对测试集中每个评论文本的特定目标进行情感分类，得到分类结果，并与测试集本身的标签对比，计算分类准确率。

[0089] 与现有技术相比，本发明采用三个通道分别进行提取特征，可以获取到更全面的信息，第一个通道是利用分层池化，获取目标词和上下文的表示，即先进行平均池化，再进行最大池化，其中平均池化利用了所有词的特征信息，而最大池化则利用了最突出的特征信息；第二个是交互注意力机制，使得目标信息与上下文信息进可以交互学习，从而获得交互信息；第三个通道是基于欧式距离的注意力机制，充分利用语义信息，即语义较近的词语彼此相互影响较大，分配的权重较大，反之，语义较远的词语彼此相互影响较小，分配的权重较小。本发明能够准确地挖掘评论中特定目标的情感极性，适用于各种领域评论的特定目标情感分类，一方面可以使潜在消费者在购买商品前了解商品的评价信息，另一方面可以使商家更加充分地了解消费者的意见，从而提高服务质量，具有极大的实用价值。

附图说明

[0090] 图1是实例中的流程示意图。

[0091] 图2是第一个通道模型图。

[0092] 图3是第二个通道模型图。

[0093] 图4是第三个通道模型图

[0094] 图5是整体模型图。

[0095] 图6是分层池化图。

具体实施方式

[0096] 以下结合附图和实例对本发明的实施方式作进一步说明，但本发明的实施和保护不限于此。

[0097] 本实例的一种基于多通道的特定目标情感分析方法，以SemEval 2014评测数据集为例，主要包括以下部分：(1)获取SemEval 2014评测数据，对评测数据集进行预处理，并将其划分成训练集、测试集；(2)预处理后的数据分别输入到三个通道进行特征提取后获得向量r_1、r_2、r_3、r_4和r_5；(3)利用向量r_1、r_2、r_3、r_4和r_5，通过注意力机制的学习，得到分类结果；(4)用训练好的模型对测试集中每个评论文本的特定目标进行情感分类，并与测试集本身的标签对比，计算分类准确率。流程示意图如图1所示，整体模型图如图5所示。下面将进行详细的介绍。

[0098] 其包括以下步骤：

[0099] (1)对特定目标情感分析的测评数据进行预处理，包括获得评论文本、特定目标及其情感极性，处理后的数据集格式为第一行原始文本，特定目标由“aspect_term”代替，例如“aspect_term is super fast,around anywhere from 35seconds to 1minute.”特定目标Boot time被替代；第二行为特定目标；第三行为特定目标的情感极性；然后按3:1的比例将测评数据随机划分成训练集和测试集，并保证两者中积极和消极的评论数基本平衡；

[0100] (2)使用斯坦福大学公开的300维glove词向量与输入文本中的词进行匹配，使得文本中的每个词都能对应得到300维向量，对于没有匹配到的词，则在[-0.1,0.1]中随机取值后得到词向量，作为模型的输入并分别进入到三个通道；

[0101] 其中三个通道分别为：

[0102] (2-1)第一个通道是将上下文表示W_c与目标表示W_t进行直接拼接，得到矩阵$W_{1,tc}$，其中$W_c \in \mathbb{R}^{n\times d_c}$、$W_t \in \mathbb{R}^{m\times d_c}$、$W_{1,tc} \in \mathbb{R}^{(m+n)\times d_c}$，m,n分别是目标词和上下文中词的个数，$d_c$是词向量维度，将$W_{1,tc}$通过LSTM得到隐含状态$H_{1,tc}$，其中$H_{1,tc} \in \mathbb{R}^{(m+n)\times d}$，d为LSTM隐藏层的维度，然后对$H_{1,tc}$进行分层池化操作，得到向量$r_1$，$r_1 \in \mathbb{R}^d$；

[0103] 更进一步的，第一个通道进行特征提取获得向量r_1的方法，包括以下步骤：

[0104] 1)将W_t与W_c进行直接拼接，得到矩阵$W_{1,tc}$；

[0105] 2)LSTM网络中分别设计了输入门、遗忘门、输出门来控制信息的输入，保留和输出。其前向传播公式如下所示：

[0106] $i^t=\sigma W_i \cdot [h^{t-1};w^t]+b_i$

[0107] $f^t=\sigma W_f \cdot [h^{t-1};w^t]+b_f$

[0108] $o^t=\sigma W_o \cdot [h^{t-1};w^t]+b_o$

[0109] $g^t = \tanh\ W_r \cdot [h^{t-1}; w^t] + b_r$

[0110] $c^t = i^t \odot g^t + f^t \odot c^{t-1}$

[0111] $h^t = o^t \odot \tan \odot h\ c^t$

[0112] 其中i^t,f^t,o^t分别表示的是输入门,遗忘门,输出门。h^{t-1}为上个细胞单元的输出,g^t为当前细胞输入的状态,c^t和h^t分别为当前细胞状态和隐藏层输出,w^t为t时刻的输入向量,W_i,W_f,W_o,W_r为参数矩阵,b_i,b_f,b_o,b_r为偏置,⊙为点乘,σ为softmax激活函数;

[0113] 将拼接后的矩阵作为LSTM的输入,得到隐含状态$H_{1,tc} \in \mathbb{R}^{(m+n)\times d}$;

[0114] 3)进行分层池化,即先进行平均池化,然后再进行最大池化,在实验数据集SemEval2014Task4中,使用的平均池化窗口为8×d,然后得到平均池化后的矩阵,对整个矩阵采用最大池化得到向量r_1;

[0115] (2-2)第二个通道是将W_t与W_c分别输入到LSTM,得到目标词和上下文的隐含状态$H_{2,t}$和$H_{2,c}$,其中$H_{2,t} \in \mathbb{R}^{m\times d}$, $H_{2,c} \in \mathbb{R}^{n\times d}$,将$H_{2,t}$和$H_{2,c}$分别进行平均池化操作,得到目标词和上下文的平均池化向量$t_{2,avg}$和$c_{2,avg}$,$t_{2,avg} \in \mathbb{R}^d$, $c_{2,avg} \in \mathbb{R}^d$,然后引入交互注意力机制,使目标信息与上下文信息充分交互,得到向量r_3,$r_3 \in \mathbb{R}^d$;

[0116] 更进一步的,第二个通道进行特征提取获得向量r_2和向量r_3的方法,包括以下步骤:

[0117] 1)将W_t与W_c分别输入到LSTM,得到目标词和上下文的隐含状态$H_{2,t}$和$H_{2,c}$;

[0118] 2)将$H_{2,c}$进行平均池化,如下公式所示,得到上下文的平均池化向量$c_{2,avg}$;

[0119] $$c_{2,avg} = \sum_{i=1}^{n} h_{2,c}^i / n$$

[0120] 其中$h_{2,c}^i$是$H_{2,c}$中的行向量,i∈[1,n],因此$c_{2,avg}$包含了$H_{2,c}$的信息;

[0121] 3)将$H_{2,t}$进行平均池化,如下公式所示,得到目标词的平均池化向量$t_{2,avg}$;

[0122] $$t_{2,avg} = \sum_{j=1}^{m} h_{2,t}^j / \mathrm{m}$$

[0123] 其中$h_{2,t}^j$是$H_{2,t}$中的行向量,j∈[1,m],因此$t_{2,avg}$包含了$H_{2,t}$的信息;

[0124] 4)利用$c_{2,avg}$与$H_{2,t}$的第j个行向量$h_{2,t}^j$,通过交互学习,得到$\gamma(h_{2,t}^j, c_{2,avg})$,公式如下所示:

[0125] $$\gamma(h_{2,t}^j, c_{2,avg}) = \tanh(h_{2,t}^j \cdot W_{2,b} \cdot c_{2,avg}^T + b_{2,b})$$

[0126] 其中$W_{2,b}$是交互学习的参数矩阵,维度为$\mathbb{R}^{d\times d}$,$h_{2,t}^j$是$H_{2,t}$的一个行向量,$c_{2,avg}^T$为$c_{2,avg}$的转置,$b_{2,b}$为偏置;

[0127] 5)对每个$\gamma(h_{2,t}^j, c_{2,avg})$进行归一化,求得对应$H_{2,t}$第j个行向量$h_{2,t}^j$的系数$\beta_j$,公式如下所示:

[0128] $$\beta_j = \frac{\exp(\gamma(h_{2,t}^j, c_{2,avg}))}{\sum_{k=1}^{m} \exp(\gamma(h_{2,t}^k, c_{2,avg}))}$$

[0129] 其中k、j∈[1:m]；

[0130] 6)将β_j与$H_{2,t}$的第j个特征向量$h_{2,t}^j$相乘，加权求和的结果即为采用注意力机制之后得到的向量r_2，公式如下所示：

[0131] $$r_2=\sum_{j=1}^{m}\beta_j h_{2,t}^j$$

[0132] 7)同理，利用$t_{2,avg}$与$H_{2,c}$的第i个行向量$h_{2,c}^i$，通过交互学习，得到向量r_3，原理与4)—6)类似，这里不再重复，公式如下所示：

[0133] $$\gamma(h_{2,c}^i,t_{2,avg})=\tanh(h_{2,c}^i\cdot W_{2,a}\cdot t_{2,avg}^T+b_{2,a})$$

[0134] $$\alpha_i=\frac{\exp(\gamma(h_{2,c}^i,t_{2,avg}))}{\sum_{l=1}^{n}\exp(\gamma(h_{2,c}^l,t_{2,avg}))}$$

[0135] $$r_3=\sum_{i=1}^{n}\alpha_i h_{2,c}^i$$

[0136] 其中l、i∈[1:n]；

[0137] (2-3)第三个通道将W_t与W_c分别输入到LSTM，得到目标词和上下文的隐含状态$H_{3,t}^1$与$H_{3,c}^1$，其中$H_{3,t}^1\in\mathbb{R}^{m\times d}$，$H_{3,c}^1\in\mathbb{R}^{n\times d}$，引入基于欧式距离的注意力机制，充分利用语义信息，得到注意力机制权重矩阵$H_{3,tc}$，其中$H_{3,tc}\in\mathbb{R}^{m\times n}$，通过$H_{3,tc}$的转置与$H_{3,t}^1$相乘得到$H_3^2{}_{,t}$，$H_{3,tc}$与$H_{3,c}^1$相乘得到$H_{3,c}^2$，其中$H_{3,t}^2\in\mathbb{R}^{n\times d}$，为上下文对目标词的基于欧式距离注意力机制后的表示，$H_{3,c}^2\in\mathbb{R}^{m\times d}$为目标词对上下文的基于欧式距离注意力机制后的表示，将$H_{3,t}^2$和$H_{3,c}^2$输入到LSTM，得到隐含状态$H_{3,t}^3$和$H_{3,c}^3$，其中$H_{3,t}^3\in\mathbb{R}^{n\times d}$，$H_{3,c}^3\in\mathbb{R}^{m\times d}$，将$H_{3,t}^3$、$H_{3,c}^3$进行平均池化得到$t_{3,avg}$和$c_{3,avg}$，其中$t_{3,avg}\in\mathbb{R}^d$，$c_{3,avg}\in\mathbb{R}^d$，引入交互注意力机制后，得到向量$r_4$与$r_5$，其中$r_4\in\mathbb{R}^d$，$r_5\in\mathbb{R}^d$；

[0138] 更进一步的，第三个通道进行特征提取获得r_4和r_5的方法，包括以下步骤：

[0139] 1)将W_t和W_c分别输入到LSTM中，得到$H_{3,t}^1$，$H_{3,c}^1$；

[0140] 2)引入基于欧式距离的注意力机制，充分利用语义信息，得到注意力机制权重矩阵$H_{3,tc}$，计算公式如下所示：

[0141] $$h_{3,tc}^{i,j}=\frac{1}{1+|h_{3,c}^{1,i}-h_{3,t}^{1,j}|}$$

[0142] 其中$h_{3,tc}^{i,j}$为$H_{3,tc}$中的第i行第j列元素，$h_{3,c}^{1,i}$为上下文特征矩阵$H_{3,c}^1$中的第i个行向量，$h_{3,t}^{1,j}$为目标特征矩阵$H_{3,t}^1$中的第j个行向量，维度为d，$|h_{3,c}^{1,i}-h_{3,t}^{1,j}|$为两个向量的欧式距离，加1操作是为了防止两个完全一样的向量导致分母为0；其意义为距离较近的两个向量之间相互影响较大，则注意力机制权重较大，反之，距离较远的两个向量之间相互影响较小，则注意力机制权重较小；

[0143] 3)$H_{3,tc}$与$H_{3,t}^{1}$相乘得到基于欧式距离的注意力机制后的表示$H_{3,t}^{2}$，公式如下所示：

[0144] $$H_{3,t}^{2}=H_{3,tc}\times H_{3,t}^{1}$$

[0145] 其中$H_{3,t}^{1}\in\mathbb{R}^{m\times d}$，$H_{3,tc}\in\mathbb{R}^{n\times m}$，$H_{3,t}^{2}\in\mathbb{R}^{n\times d}$

[0146] 4)$H_{3,tc}$与$H_{3,c}^{1}$相乘得到基于欧式距离的注意力机制后的表示$H_{3,c}^{2}$，公式如下所示：

[0147] $$H_{3,c}^{2}=H_{3,tc}^{T}\times H_{3,c}^{1}$$

[0148] 其中$H_{3,c}^{1}\in\mathbb{R}^{n\times d}$，$H_{3,tc}\in\mathbb{R}^{n\times m}$，$H_{3,c}^{2}\in\mathbb{R}^{m\times d}$

[0149] 5)将$H_{3,t}^{2}$，$H_{3,c}^{2}$输入到LSTM进一步提取特征，且得到输出为$H_{3,t}^{3}$、$H_{3,c}^{3}$，$H_{3,t}^{3}$、$H_{3,c}^{3}$的维度分别与$H_{3,t}^{2}$，$H_{3,c}^{2}$的维度一致；

[0150] 6)将$H_{3,t}^{3}$进行平均池化，得到目标信息的平均池化向量$t_{3,avg}$，公式如下所示：

[0151] $$t_{3,avg}=\sum_{i=1}^{n}h_{3,t}^{3,i}/\mathrm{n}$$

[0152] 其中$h_{3,t}^{3,i}$为矩阵$H_{3,t}^{3}$的行向量，i∈[1:n]，$t_{3,avg}\in\mathbb{R}^{d}$

[0153] 7)利用$t_{3,avg}$与$H_{3,c}^{3}$，通过交互学习，得到$\gamma(h_{3,c}^{3,j},t_{3,avg})$，公式如下所示：

[0154] $$\gamma(h_{3,c}^{3,j},t_{3,avg})=\tanh(h_{3,c}^{3,j}\cdot W_{3,a}\cdot t_{3,avg}^{T}+b_{3,a})$$

[0155] 其中$W_{3,a}$为交互学习的参数矩阵，维度为$R^{d\times d}$，$b_{3,a}$为偏置项；

[0156] 8)对每一个$\gamma(h_{3,c}^{3,j},t_{3,avg})$进行归一化，得到$H_{3,t}^{3}$第j个行向量$h_{3,c}^{3,j}$的注意力机制权重系数$\alpha_j$，公式如下所示：

[0157] $$\alpha_{j}=\frac{\exp(\gamma(h_{3,c}^{3,j},t_{3,avg}))}{\sum_{k=1}^{m}\exp(\gamma(h_{3,c}^{3,k},t_{3,avg}))}$$

[0158] 其中j,k∈[1:m]；

[0159] 9)将α_j与$H_{3,t}^{3}$第j个行向量$h_{3,c}^{3,j}$相乘，加权求和得到向量r_4，公式如下所示：

[0160] $$r_{4}=\sum_{i=1}^{m}\alpha_{j}h_{3,c}^{3,j}$$

[0161] 10)将$H_{3,c}^{3}$进行平均池化，得到上下文信息的平均池化向量$c_{3,avg}$，具体如下公式所示；

[0162] $$c_{3,avg}=\sum_{j=1}^{m}h_{3,c}^{3,j}/\mathrm{m}$$

[0163] 其中$h_{3,c}^{3,j}$为矩阵$H_{3,c}^{3}$的第j个行向量，j∈[1:m]。

[0164] 11)同理，利用$c_{3,avg}$与$H_{3,t}^{3}$第i个行向量$h_{3,t}^{3,i}$，通过交互学习，得到向量r_5，原理与8)—10)类似，这里不再重复，公式如下所示：

[0165] $$\gamma(h_{3,t}^{3,i}, c_{3,avg}) = \tanh(h_{3,t}^{3,i} \cdot W_{3,b} \cdot c_{3,avg}^{T} + b_{3,b})$$

[0166] $$\beta_i = \frac{\exp(\gamma(h_{3,t}^{3,i}, c_{3,avg}))}{\sum_{l=1}^{n} \exp(\gamma(h_{3,t}^{3,l}, c_{3,avg}))}$$

[0167] $$r_5 = \sum_{i=1}^{n} \beta_i h_{3,t}^{3,i}$$

[0168] 其中$W_{3,b}$为交互学习的参数矩阵，维度是$\mathbb{R}^{d\times d}$，$b_{3,b}$为偏置项，i,l∈[1:n]；

[0169] (3)利用向量r_1、r_2、r_3、r_4和r_5，通过注意力机制的学习，得到r，公式如下所示：

[0170] $$c_p = \frac{\exp(r_p w_p^T)}{\sum_{z=1}^{5} \exp(r_z w_z^T))}$$

[0171] $$r = \sum_{p=1}^{5} c_p r_p$$

[0172] 其中$r_p \in [r_1, r_2, r_3, r_4, r_5]$，$w_p^T$和$w_z^T$为参数向量，$w_p^T \in \mathbb{R}^{d\times 1}$，$w_z^T \in \mathbb{R}^{d\times 1}$，$c_p$为注意力机制系数；

[0173] 最后将r经过全连接层，并使用softmax作为激活函数，得到最终的分类结果。

[0174] (4)用训练好的模型对测试集中每个评论文本的特定目标进行情感分类，得到分类结果，并与测试集本身的标签对比，计算分类准确率。

[0175] 针对本发明方法进行实验论证，具体包括：

[0176] 1.标准数据集：

[0177] 本发明使用SemEval2014Task4中的laptop和restaurant数据集作为标准数据集，验证基于多通道模型的特定目标情感分析方法的有效性，laptop数据集包括3012个句子，其中划分为2358个训练语句和654个测试语句；restaurant包括4722个句子，其中划分为3602个训练语句和1120个测试语句。其中两个数据集都含有积极、消极、中立和矛盾的情感标签。由于本模型进行主要是三分类，没有考虑少量的矛盾句子，因此预处理时删除了标签为矛盾的句子。其中laptop训练集中含有45个标签为conflict的句子，处理后laptop训练集中包括2313个句子，laptop测试集中含有16个标签为conflict的句子，处理后laptop训练集中包括638个句子。restaurant训练集中含有0个标签为conflict的句子，不做处理，仍有3602个训练语句，测试集中含有0个标签为conflict的句子，不做处理，仍有1120个测试语句。表一是两个数据集的说明。

[0178]

Dataset	Neg	Pos	Neu	total
Laptop-train	866	987	460	2313
Laptop-test	128	341	169	638
Restaurant-train	805	2164	633	3602
Restaurant-test	728	196	196	1120

[0179] 表1数据集说明

[0180] 词向量的表示，采用的是斯坦福发布的6B glove词向量，维度为300，文本大小为989M。

[0181] 2.评价指标：

[0182] 本发明使用目前文本情感分类领域常用的评价指标：平均分类准确率(Average Accuracy)作为本实施例的评价指标，其计算公式如下所示：

[0183] $$Average\ Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

[0184] 其中，TP(True Positive)表示分类模型正确预测的积极样本数，TN(True Negative)表示分类模型正确预测的消极样本数，FP(False Positive)表示分类模型错误预测的积极样本数，FN(False Negative)表示分类模型错误预测的消极样本数。

[0185] 3.实验结果

[0186] 为了验证本发明所提方法的有效性，本发明在实验方案中利用了经典的情感分类算法作为对比，最终验证了基于多通道模型的特定目标情感分析方法的有效性。在数据集SemEval2014Task4上本专利模型与几种经典模型三分类的对比情况如表2所示，其中本发明MCM取得了最好的实验结果。

[0187]

Dataset	Restaurant	Laptop
Majority	0.535	0.650
LSTM	0.743	0.665
TD-LSTM	0.756	0.681
AE-LSTM	0.762	0.689
ATAE-LSTM	0.772	0.687

[0188]

MCM	**0.791**	**0.713**

[0189] 表2实验结果

[0190] 4.超参数设置

[0191] 在实验中，所有的词向量都是来自斯坦福大学公布的300维6B的GloVe向量，对于未登录词或者GloVe词典中没有出现的过的词，则是在[-0.1,0.1]中随机取值。所有的参数矩阵数值也是在[-0.1,0.1]中随机取值，所有的偏置的初始值赋为0，LSTM的隐藏层的维度设置为150，学习率设置为0.01，L2正则化设置为0.00001，丢弃率设置为0.5，另外在分层池化的过程中，先进行平均池化，池化的窗口大小是8*1，然后对得到的矩阵取最大池化，池化窗口是所得矩阵的行数*1。

[0192] 5.模型对比

[0193] 为了更全面的评估本专利的模型，我们与一些经典模型进行对比，下面先介绍一下各模型。

[0194] (1)Majority：把句子中最大概率的极性当成特定目标的情感极性。

[0195] (2)LSTM：利用单层的LSTM，对特定目标进行情感分类。

[0196] (3)TD-LSTM：用两个LSTM网络分别对目标词、目标词上文、目标词下文进行建模，从而得到目标词的上下文信息。

[0197] (4)AT-LSTM:首先用LSTM对句子进行建模,将LSTM的隐藏层输出与目标词向量进行拼接,再通过注意力机制得到最终的特征表示。

[0198] (5)ATAE-LSTM:ATAE-LSTM是AT-LSTM的进一步扩展,区别是在LSTM句子进行建模之前,上下文词向量与目标词向量进行拼接。

[0199] 实验结果如表2所示,Majority的结果最差,在res和lap数据集上的准确率分别为53.5%和65.0%;所有其他的模型都采用了LSTM的模型,结果都好于Majority,其中仅仅采用LSTM网络的模型是剩余模型中表现最差的,其在res和lap数据集上的准确分别为74.3%和66.5%,原因是忽略了目标属性信息;TD-LSTM模型,使用了目标属性信息,其在res和lap数据集上的准确率分别高于LSTM1.3%、2.4%,可见效果有了明显的提升,表明了目标信息对分类结果的重要贡献,TD-LSTM虽然考虑了目标信息,但是却同等对待每一个词在最终结果中起到的作用,不能识别出贡献程度大的词语,所以相比引入了注意力机制的AE-LST和ATAE-LSTM模型,其准确率结果分别在res数据集低了0.6%、1.6%,在lap数据集上低了0.8%、0.6%。MCM模型在ATAE-LSTM的基础上,添加了分层池化、交互注意力机制、基于欧式距离的注意力机制,从而获得在res数据集79.1%、lap数据集71.3%的最好结果。

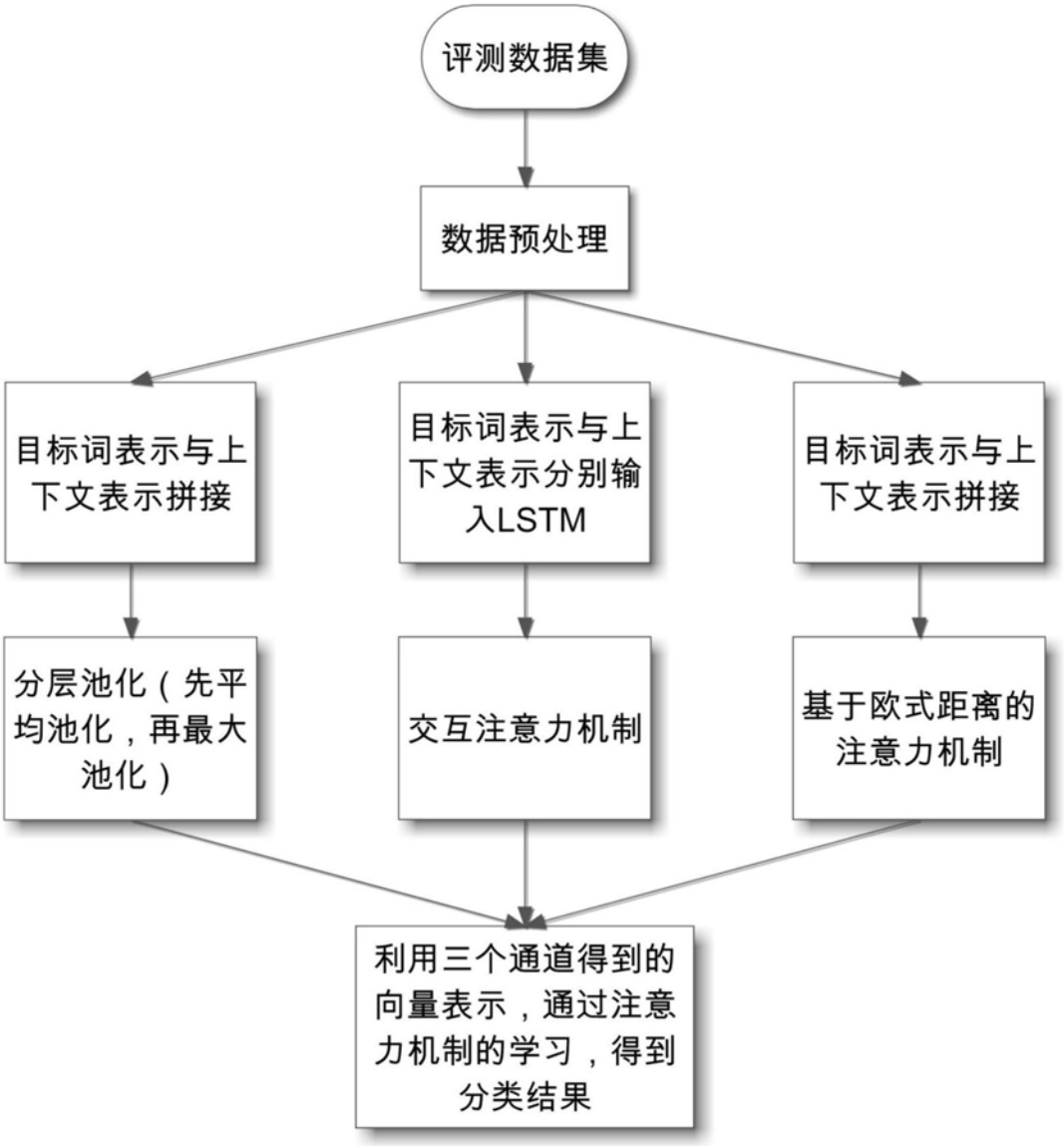

评测数据集
数据预处理
目标词表示与上下文表示拼接
目标词表示与上下文表示分别输入LSTM
目标词表示与上下文表示拼接
分层池化（先平均池化，再最大池化）
交互注意力机制
基于欧式距离的注意力机制
利用三个通道得到的向量表示，通过注意力机制的学习，得到分类结果

图1

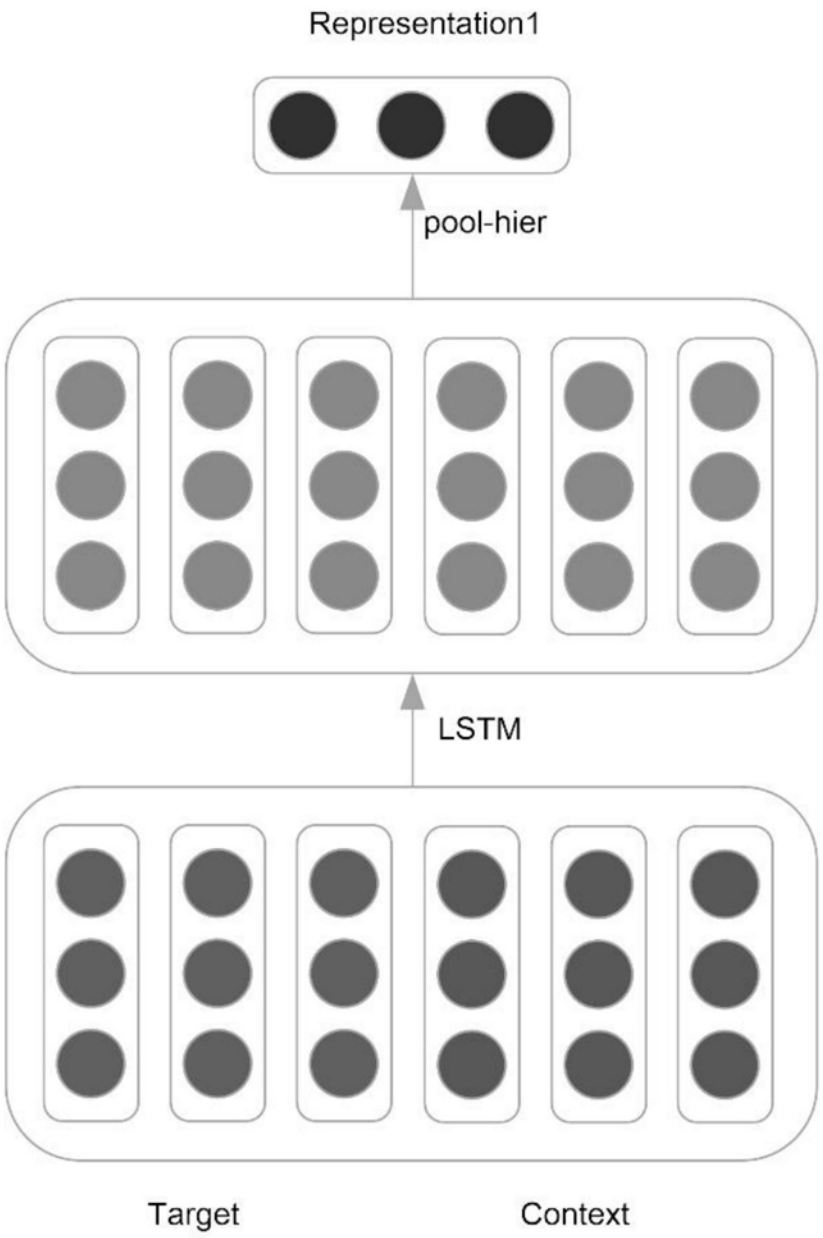


图2

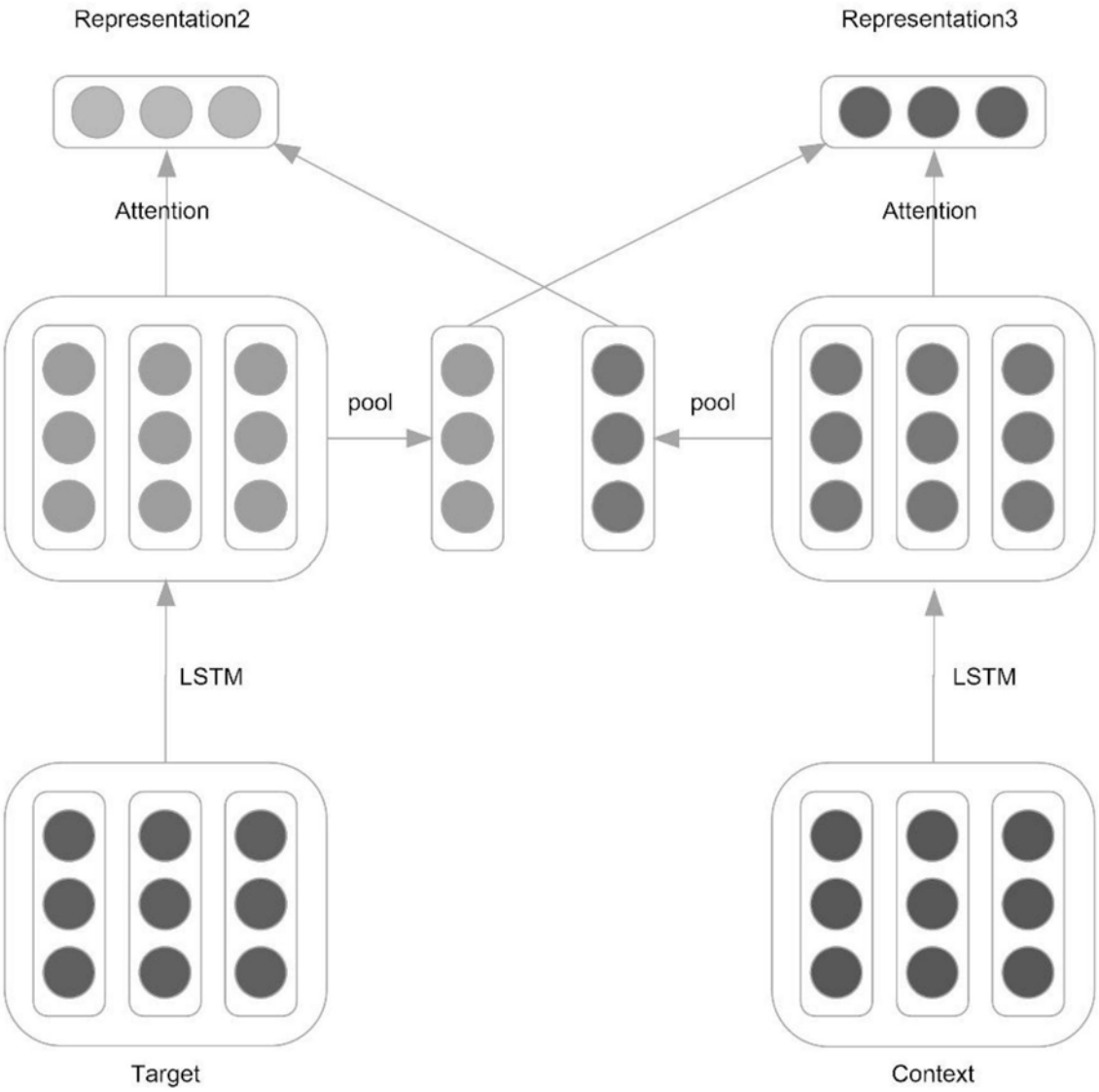

Representation2
Representation3
Attention
Attention
pool
pool
LSTM
LSTM
Target
Context

图3

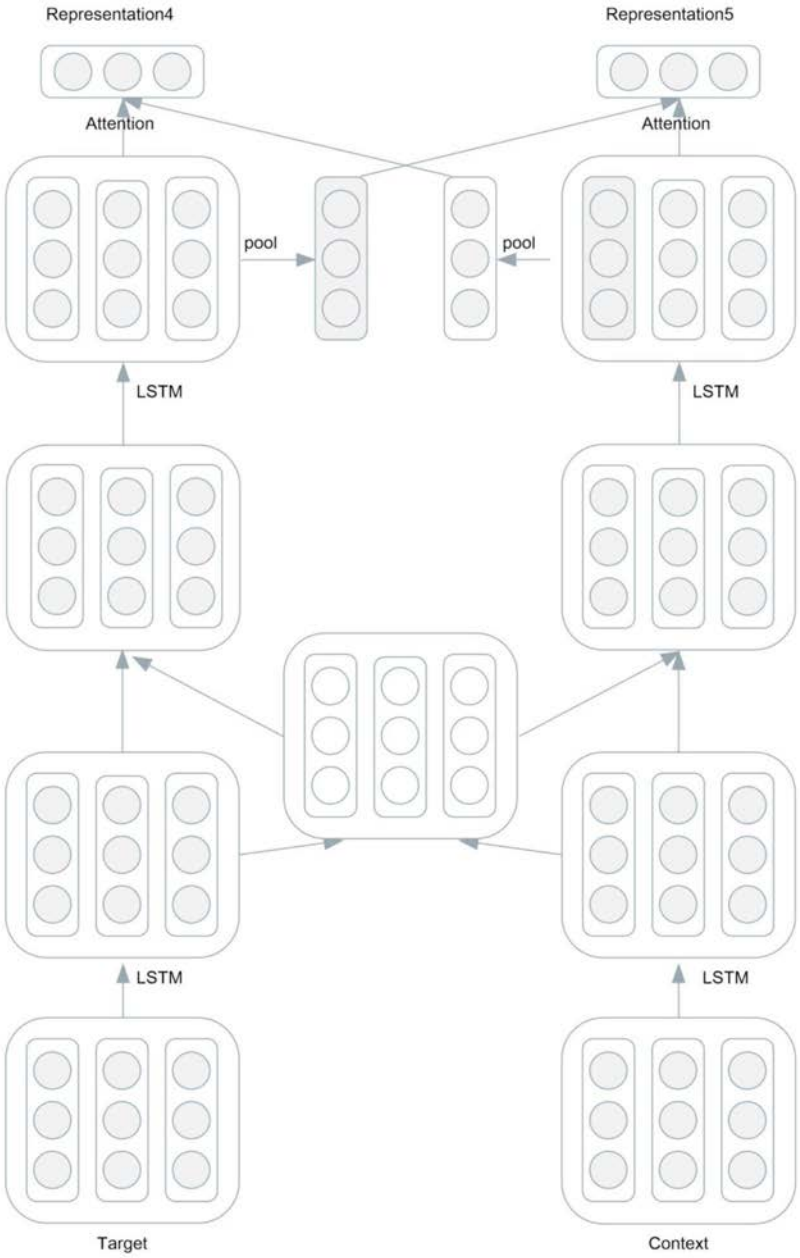

Representation4
Representation5
Attention
Attention
pool
pool
LSTM
LSTM
LSTM
LSTM
Target
Context

图4

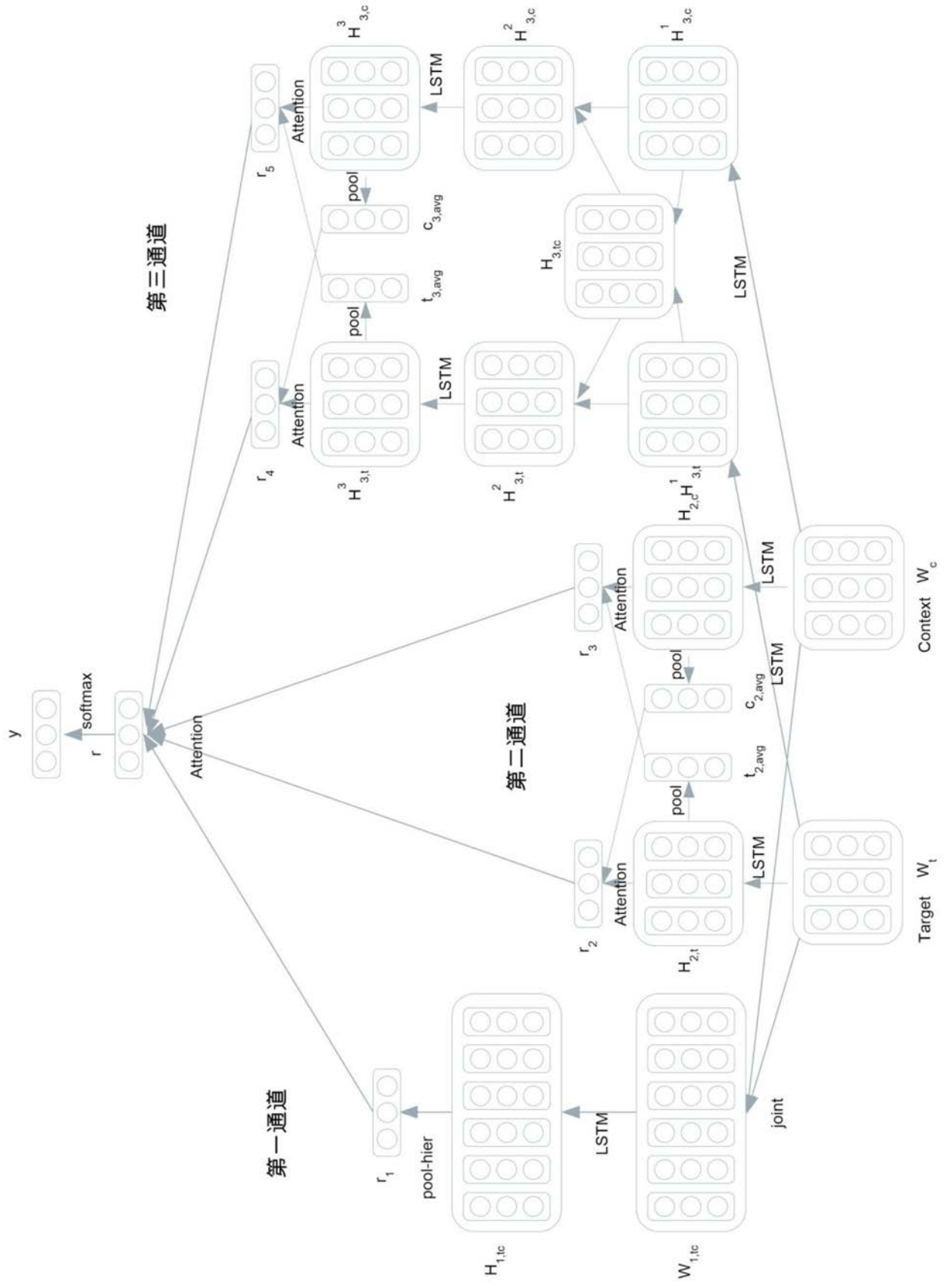

第三通道
第二通道
第一通道
softmax
y
r
Attention
LSTM
pool
pool-hier
joint
Context W_c
Target W_t
r_1
r_2
r_3
r_4
r_5
$H_{1,tc}$
$W_{1,tc}$
$H_{2,t}$
$H_{2,c}$ $H^1_{3,t}$
$H^2_{3,t}$
$H^3_{3,t}$
$H_{3,tc}$
$H^1_{3,c}$
$H^2_{3,c}$
$H^3_{3,c}$
$t_{2,avg}$
$c_{2,avg}$
$t_{3,avg}$
$c_{3,avg}$

图5

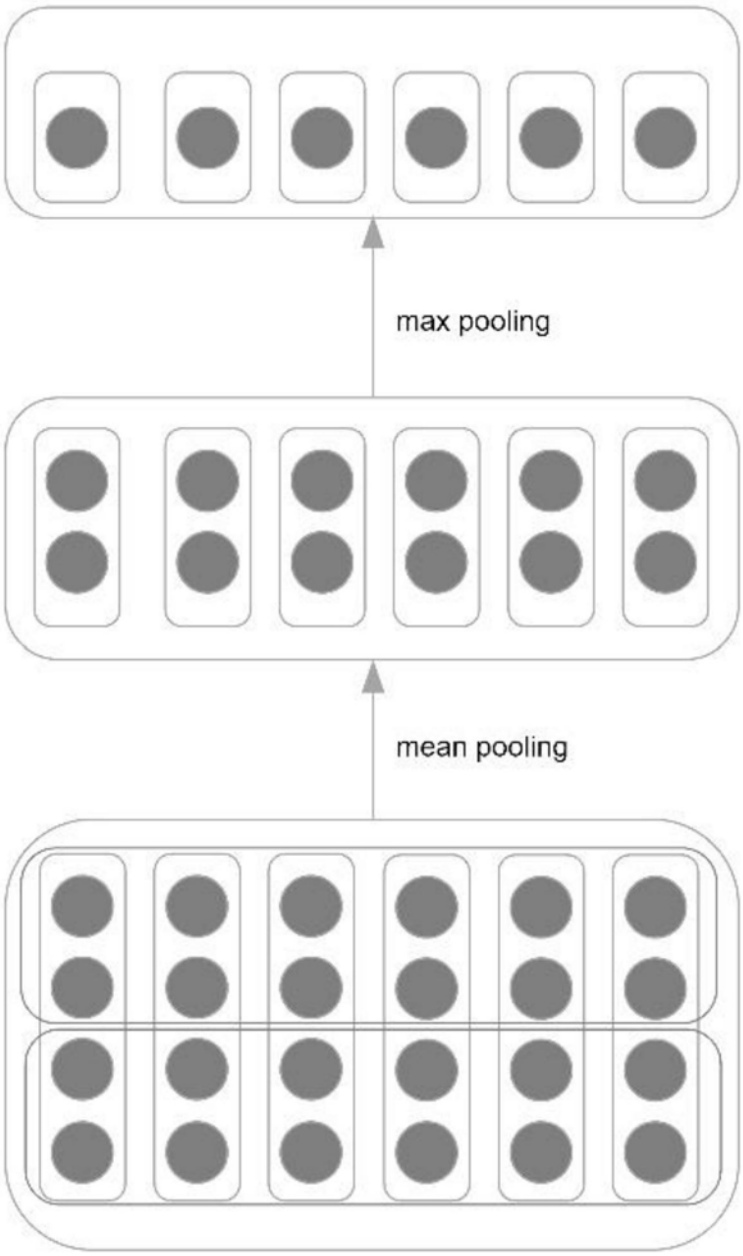


图6