



基于双特征嵌套注意力的方面词情感分析算法

肖宇晗, 林慧苹, 汪权彬, 谭莹

引用本文:

肖宇晗, 林慧苹, 汪权彬, 等. 基于双特征嵌套注意力的方面词情感分析算法[J]. *智能系统学报*, 2021, 16(1): 142–151.

XIAO Yuhan, LIN Huiping, WANG Quanbin, et al. An algorithm for aspect–based sentiment analysis based on dual features attention–over–attention[J]. *CAAI Transactions on Intelligent Systems*, 2021, 16(1): 142–151.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202012024>

您可能感兴趣的其他文章

基于注意力机制的显著性目标检测方法

Salient object detection method based on the attention mechanism

智能系统学报. 2020, 15(5): 956–963 <https://dx.doi.org/10.11992/tis.201903001>

加入自注意力机制的BERT命名实体识别模型

BERT named entity recognition model with self–attention mechanism

智能系统学报. 2020, 15(4): 772–779 <https://dx.doi.org/10.11992/tis.202003003>

基于注意力融合的图像描述生成方法

An image caption generation method based on attention fusion

智能系统学报. 2020, 15(4): 740–749 <https://dx.doi.org/10.11992/tis.201910039>

层次化双注意力神经网络模型的情感分析研究

Hierarchical double–attention neural networks for sentiment classification

智能系统学报. 2020, 15(3): 460–467 <https://dx.doi.org/10.11992/tis.201812017>

深度强化学习中状态注意力机制的研究

State attention in deep reinforcement learning

智能系统学报. 2020, 15(2): 317–322 <https://dx.doi.org/10.11992/tis.201809033>

半监督自训练的方面提取

Aspects extraction based on semi–supervised self–training

智能系统学报. 2019, 14(4): 635–641 <https://dx.doi.org/10.11992/tis.201806006>

 微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.202012024

基于双特征嵌套注意力的方面词情感分析算法

肖宇晗¹, 林慧苹¹, 汪权彬², 谭莹²

(1. 北京大学软件与微电子学院, 北京 102600; 2. 北京大学信息科学技术学院, 北京 100871)

摘要: 针对目前方面词情感分析方法忽视了以方面词为核心的局部特征的重要性, 并难以有效减小情感干扰项的负面噪声的问题, 本文提出了一种带有基于变换器的双向编码器表示技术 (bi-directional encoder representations from transformers, BERT) 加持的双特征嵌套注意力模型 (dual features attention-over-attention with BERT, DFAOA-BERT), 首次将 AOA (attention-over-attention) 与 BERT 预训练模型结合, 并设计了全局与局部特征提取器, 能够充分捕捉方面词和语境的有效语义关联。实验结果表明: DFAOA-BERT 在 SemEval 2014 任务 4 中的餐馆评论、笔记本评论和 ACL-14 Twitter 社交评论这 3 个公开数据集上均表现优异, 而子模块的有效性实验, 也充分证明了 DFAOA-BERT 各个部分的设计合理性。

关键词: 情感分析; 方面词; 嵌套注意力; BERT 预训练模型; 全局特征; 局部特征; 深度学习; 机器学习
中图分类号: TP391 **文献标志码:** A **文章编号:** 1673-4785(2021)01-0142-10

中文引用格式: 肖宇晗, 林慧苹, 汪权彬, 等. 基于双特征嵌套注意力的方面词情感分析算法 [J]. 智能系统学报, 2021, 16(1): 142-151.

英文引用格式: XIAO Yuhan, LIN Huiping, WANG Quanbin, et al. An algorithm for aspect-based sentiment analysis based on dual features attention-over-attention [J]. CAAI transactions on intelligent systems, 2021, 16(1): 142-151.

An algorithm for aspect-based sentiment analysis based on dual features attention-over-attention

XIAO Yuhan¹, LIN Huiping¹, WANG Quanbin², TAN Ying²

(1. School of Software and Microelectronics, Peking University, Beijing 102600, China; 2. School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China)

Abstract: Aspect-based sentiment analysis is of great significance to making full use of product reviews to analyze potential user needs. The current research work still has deficiencies. Many studies ignore the importance of local features centered on aspects and fail to handle emotional disturbances effectively. To address these problems, this article proposes a dual features attention-over-attention model with BERT (DFAOA-BERT). For the first time, an AOA (attention-over-attention) mechanism is combined with the BERT pretrained model. DFAOA-BERT also designs global and local feature extractors to fully capture an effective semantic association between aspects and context. According to the experimental results, DFAOA-BERT performs well on three public datasets: restaurant and laptop review datasets from SemEval 2014 Task 4 and the ACL-14 Twitter social review dataset. The effectiveness experiment of submodules also fully proves that each part of DFAOA-BERT makes a significant contribution to the excellent performance.

Keywords: sentiment analysis; aspect; attention-over-attention; BERT pretrained model; global feature; local feature; deep learning; machine learning

随着互联网的迅速发展, 人们越来越倾向于

在网络上发表自己对商品或服务的看法。提取评论所蕴含的情感态度与需求, 有利于商家深入分析用户的潜在需求, 判断市场走向, 改进产品或服务设计。

值得注意的是, 人们在进行评论时, 一般不会

收稿日期: 2020-12-15.

基金项目: 国家重点研发计划项目 (2018AAA0102301, 2018AAA0100302, 2018YFB1702900); 国家自然科学基金项目 (62076010).

通信作者: 谭莹. E-mail: ytan@pku.edu.cn.

发表长篇大论, 而是倾向于用一两句话概括某个事物的某个方面的好坏, 甚至会在同一句话中对多个方面表达不同的情感态度。因此, 方面词情感分析任务的研究近年来颇受关注。该任务的要求是, 在给定文本以及方面词的情况下, 分别对这些方面词所承载的情感态度进行分类^[1-2]。例如, 给定“这家店的环境不错, 服务周到, 但食材不新鲜”这段文本, 并给出“环境”、“服务”、“食材”这 3 个方面词, 算法应当判断出用户对“环境”和“服务”的情感态度是正向, 而对“食材”持负面态度。考虑到方面词往往蕴含了用户对产品或服务的需求, 方面词情感分析算法的研究将对细粒度分析用户潜在需求有着重要意义。

与文档级和句子级情感分析任务相比, 方面词情感分析任务粒度更细、更具挑战。其中一个棘手的问题是, 如何在保留情感有效项信息的同时, 尽可能减小干扰项对情感判断的负面影响。例如, “我买过这家的鼠标、电脑和平板, 鼠标很快就没用了, 至于这个电脑, 不好也不坏, 平板倒是挺结实”, 如果要求对方面词“电脑”进行情感分析, 答案理论上应是中立, 但一些算法很容易将其误判为正向或负向, 原因就在于未能有效减小“鼠标很快就没用了”和“平板倒是挺结实”对情感分析的干扰作用。除此以外, 语言的丰富性与灵活性也显著提高了方面词情感分析任务的难度。一些习语以及复杂句式也会对情感分析的结果产生不利影响。例如, “I have never had a bad meal”, 就方面词“meal”而言, 情感态度是正向的, 但部分模型会根据设计的规则或以往的训练经验, 因“bad”的存在而给出错误的分类结果^[3]。

目前, 处理方面词情感分析任务的主流方式是通过深度学习方法来赋予模型细粒度的情感分析能力。为了避免在神经网络层数较多时出现梯度消失问题, 长短时记忆网络^[4](long short-term memory, LSTM) 成为大多数深度学习模型的基本网络结构选择。Tang 等^[5]在此基础上提出了基于目标的长短时记忆网络模型(target dependent long short-term memory model, TD-LSTM), 分别通过 2 个独立的双向长短时记忆网络来提取方面词左边语境和右边语境的语义特征, 获得了比基础长短时记忆网络更好的实验结果。而注意力机制的引入, 则使得深度学习模型能够更好地利用有效项的语义特征, 并弱化情感干扰项的负面影响。Ma 等^[6]在互动注意力网络模型(interactive attention networks, IAN) 中首次提出了语境特征和方面词特征之间的互动式学习, 他们通过池化整

句话对应的隐藏层状态来获取语境对方面词的注意力权重, 从而使得 IAN 能够重点考虑语境中的情感有效项信息。针对训练数据不足的问题, Xu 等^[7]提出了半监督序列生成模型(target-oriented semi-supervised sequential generative model, TSSGM), 在减小对训练数据的依赖性的同时能够提升情感分类器的判断精度。

在最近 2 年里, 通过预训练语言模型来增强模型的语义表示能力成为方面词情感分析任务的研究热点。取自语言模型的词嵌入^[8](embeddings from language models, ELMo), 以及基于 transformer^[9]的生成式预训练^[10](generative pre-training, GPT), 均是其中的优秀代表。而 BERT^[11](bi-directional encoder representations from transformers) 的提出, 则更是里程碑式地刷新了多达 11 项自然语言处理任务的最高成绩。Karimi 等^[12]在实验中使用基础 BERT 模型来完成处理方面词情感分析任务, 发现其效果已经超过绝大多数普通神经网络模型。Song 等^[13]探索了 BERT 在方面词情感分析任务中的优势, 分别提出了用于句对分类的 BERT 模型(BERT model for sentence pair classification, BERT-SPC) 与带有 BERT 的注意力编码网络(attentional encoder network with BERT, AEN-BERT), 在多个数据集上都取得了前沿性的结果。

尽管现有的模型在方面词情感分析任务上有着不俗的表现, 但仍在 2 个方面存在不足。一方面, 在很多研究工作中, 方面词的重要性未能得到充分的重视。Jiang 等^[14]在 Twitter 数据集上评估了一个情感分类器的效果, 结果发现 40% 的分类错误都缘于没有考虑方面词的指示作用。这说明, 如何充分使用方面词提供的位置特征和语义特征将是影响模型效果的重要因素。另一方面, 方面词的情感倾向往往与邻近的形容词、动词等主观性词语关联更大, 位置较远的词语很有可能成为干扰项。有些算法虽然已经考虑到方面词的重要价值, 但却没有真正把握好如何利用方面词和语境之间的联系, 往往只是提取方面词特征和整个评论的全局语义特征来进行运算, 未能有效地消除干扰项的不利影响。

针对上述不足, 本文首次将阅读理解领域提出的 AOA(attention-over-attention) 与 BERT 预训练模型相结合, 提出了含 BERT 加持的双特征嵌套注意力模型(dual features attention-over-attention with BERT, DFAOA-BERT), 借助 BERT 来获取优质的词向量和语义编码, 在此基础上计算 AOA 数值, 并提取关键语义特征。此外, DFAOA-

BERT 还分别设计了全局和局部特征提取器,使得模型既能获得全面的语义信息,还能有效减小情感干扰项的负面作用。实验证明, DFAOA-BERT 在 3 个公开数据集上均取得了优异的成绩,而在主实验基础上进行的子模块有效性实验,也充分证明了模型主要组成部分的设计合理性。

1 AOA 注意力机制

AOA 机制由 Cui 等^[15]提出,用于处理阅读理解领域中的完形填空问题。它将完形填空任务转换为问答任务,把每个填空的候选项看作对文档的查询(query),并设计了一套算法,计算查询对文档的注意力数值以及文档对查询的注意力数值,将计算获得的双向注意力结果视为最终语义匹配特征,在此基础上为每个候选项进行打分和排序。

Huang 等^[3]则对阅读理解领域的 AOA 进行了调整,将其应用到方面词情感分析任务中。方面词被视为查询,方面词邻近的语境被看作是待查询的文档,通过计算方面词和语境之间的双向注意力数值来获得语义匹配特征,最终将产生的结果输入到情感分类层中。实验结果表明, AOA 在方面词情感分析任务中有着优秀的表现,能够帮助深度学习模型更为充分地利用方面词本身提供的信息以及它和语境之间的语义关联。

设 AOA 的输入为方面词的语义编码 h_a 和语境的语义编码 h_c , 其中, $h_a \in \mathbf{R}^{md}$, $h_c \in \mathbf{R}^{nd}$, m 是方面词序列的长度(单词数), n 是语境序列的长度, d 是语义编码的维数。计算 AOA 数值的第 1 步是获取两者的语义联系矩阵 M :

$$M = h_c h_a^T \quad (1)$$

式中 $M \in \mathbf{R}^{nm}$, M 中第 i 行 j 列的元素表示语境序列中第 i 个词和方面词序列中第 j 个词的语义匹配分数。

接着,对 M 的列进行 softmax 操作,获得方面词对语境的注意力数值 α ;对 M 的行进行 softmax 操作,获得语境对方面词的注意力数值 β 。

$$\alpha_{ij} = \frac{\exp(M_{ij})}{\sum_i \exp(M_{ij})} \quad (2)$$

$$\beta_{ij} = \frac{\exp(M_{ij})}{\sum_j \exp(M_{ij})} \quad (3)$$

对 β 按列计算平均值获得 $\bar{\beta} \in \mathbf{R}^m$, 最后的 AOA 注意力数值 $\gamma \in \mathbf{R}^n$ 则由 α 与 $\bar{\beta}$ 点乘得到:

$$\bar{\beta}_j = \frac{1}{n} \sum_i \beta_{ij} \quad (4)$$

$$\gamma = \alpha \bar{\beta}^T \quad (5)$$

2 本文模型

给定包含方面词的语境序列 $\mathbf{W} = \{w_1, w_2, \dots, w_n\}$ 和方面词序列 $\mathbf{W}' = \{w'_1, w'_2, \dots, w'_m\}$, DFAOA-BERT 模型将输出方面词对应的情感分类结果,且分类结果是负面、中立、正面中的一个。整体模型结构如图 1 所示。宏观上可以分为 4 个部分:输入层、局部特征提取器、全局特征提取器以及分类层。而局部特征提取器和全局特征提取器可以细分为 3 个子模块:词嵌入层、编码层、注意力机制。这 2 个提取器的设计要点都是以 AOA 为核心的注意力机制部分,最大的不同在于局部特征提取器中额外添加了语境动态加权模块,词嵌入层和编码层的计算方式则基本相同。

模型的介绍将围绕输入层、词嵌入层、编码层、注意力机制模块以及分类层这 5 个部分展开。

2.1 输入层

对于局部特征提取器而言,输入序列就是包含方面词的语境序列以及方面词序列本身。全局特征提取器的输入序列则有所不同,其语境序列需要经过特殊处理。受 BERT-SPC 模型^[13]启发,为了充分发挥 BERT 在本任务中的作用,全局特征提取器将原语境序列处理成“[CLS] + 原语境序列 + [SEP] + 方面词序列 + [SEP]”的形式。其中,“[CLS]”和“[SEP]”是 BERT 使用的 2 种标记符号,前者是特殊分类向量标记符,聚集了分类相关的信息,后者是分隔符,在输入多个序列时用于分隔不同序列。将原语境序列处理成这种特殊形式其实是将原格式转化成了 BERT 在处理文本分类任务时所要求的输入格式,从而充分发挥 BERT 的效果。对于方面词序列,全局特征提取器与局部特征提取器的处理方式相同。

2.2 词嵌入层

DFAOA-BERT 模型在该部分将单词一一映射到低维向量空间中,映射方式选用 BERT 词嵌入。与 GloVe^[16]不同的是, BERT 提供的是一个语义理解模型,而 GloVe 提供的只是一个词嵌入矩阵。

设 GloVe 提供的矩阵为 M , 则 $M \in \mathbf{R}^{d_e \times |V|}$, 其中, d_e 是词嵌入的维数, $|V|$ 表示单词的总数,通过矩阵 M , 可将任意单词 w_i 映射成维数为 d_e 的向量。而 BERT 词嵌入层则是一个基于序列到序列技术^[4]的预训练语义理解模型,其参数并不是像 GloVe 这样固定不变的。因此,使用一个 BERT 词嵌入层来获取局部特征提取器输入序列

与全局特征提取器序列的词嵌入, 和使用 2 个互相独立的 BERT 词嵌入层所产生的结果不完全相

同。本模型中采用 2 个独立的 BERT 词嵌入层, 原因在于该选择会小幅度提升算法的训练效果。

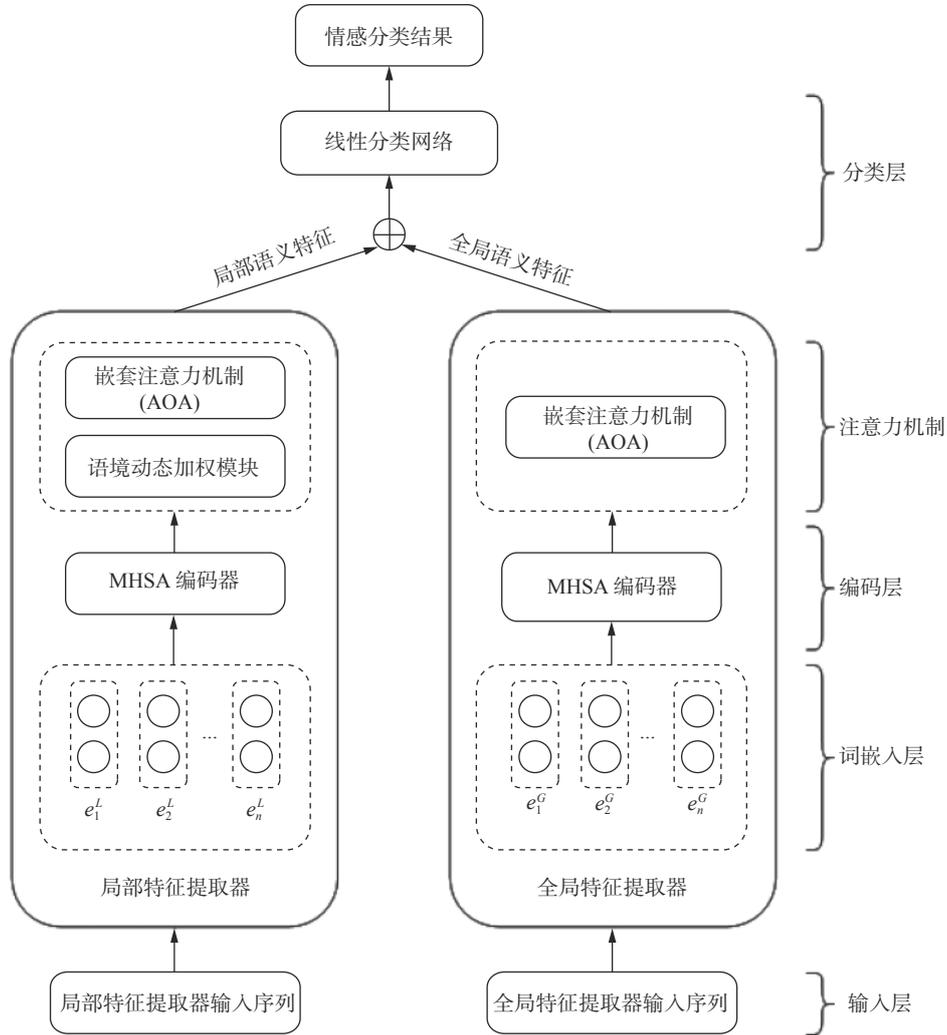


图 1 DFAOA-BERT 模型结构

Fig. 1 Structure of DFAOA-BERT model

2.3 编码层

本层的任务是将输入序列对应的词嵌入编码成含有丰富语义信息的向量。词嵌入本身包含的信息较为基础, 基于循环神经网络的模型一般使用长短时记忆网络来对词嵌入进行编码, DFAOA-BERT 则使用 MHSA^[9] (multi-head self-attention) 提取语义特征。MHSA 的主要思想是通过多个注意力函数来对每个单词计算注意力数值, 将这多个注意力数值拼接在一起后进行激活函数运算, 其结果即为输入序列的语义表示。MHSA 中的自注意力机制函数可以选用不同的设计, DFAOA-BERT 使用的是 SDA (scaled dot product attention), 相较于其他自注意力机制而言, 其优点在于既可以保证编码效果, 又能够有效提高计算的效率。

令词嵌入层的输出为 X_e , SDA 的计算方式为

$$Q = X_e W^q \quad (6)$$

$$K = X_e W^k \quad (7)$$

$$V = X_e W^v \quad (8)$$

$$SDA(X_e) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (9)$$

式中: W^q 、 W^k 、 W^v 均为权重矩阵; $W^q \in \mathbf{R}^{d_h \times d_q}$; $W^k \in \mathbf{R}^{d_h \times d_k}$; $W^v \in \mathbf{R}^{d_h \times d_v}$ 。这 3 个矩阵中的权重属于模型的可训练参数。 d_q 、 d_k 、 d_v 是矩阵维度, 三者的值都为 d_h/h , 其中, d_h 是隐藏层维数, h 是多头自注意力机制中的头数。根据每个注意力头的 SDA 计算结果, 可得到整个 MHSA 的输出 O^e :

$$SDA = SDA_0 \oplus SDA_1 \oplus \dots \oplus SDA_{h-1} \quad (10)$$

$$O^e = \text{Tanh}(SDA W^{\text{MHSA}}) \quad (11)$$

式中: \oplus 表示向量的拼接操作; W^{MHSA} 是一个向量矩阵; $W^{\text{MHSA}} \in \mathbf{R}^{hd_h \times d_h}$; Tanh 为激活函数。

通过以上步骤, MHSA 将基础的词嵌入编码

为包含丰富信息的语义向量表示。为方便下文阐释,记语境序列对应的语义编码为 O_c^e , 方面词序列对应的语义编码为 O_a^e 。

2.4 注意力机制

DFAOA-BERT 模型的注意力机制设计以 AOA 为核心, 相对于全局特征提取器而言, 局部特征提取器额外设计了语境动态加权模块来增强其获取局部关键信息的能力。

1) 语境动态加权模块。

语境动态加权模块的设计思想在于, 一个方面词的情感态度往往只与邻近的词有关, 远距离的词反而可能提供噪声信息而干扰模型的判断。因此, 从本质上讲, 语境动态加权模块就是一个位置相关的注意力机制, 保留邻近词的积极作用, 根据位置关系弱化远距离词的干扰。

受 Zeng 等^[17] 提出的局部情境焦点 (local context focus, LCF) 模型启发, DFAOA-BERT 的语境动态加权模块定义了语义相关距离 D , 根据 D 来判断如何给语境所对应的语义向量设定计算权重。设输入的原语境序列为 $W = \{w_1, w_2, \dots, w_i, w_{i+1}, \dots, w_{i+m-1}, \dots, w_n\}$, 其中 $\{w_i, w_{i+1}, \dots, w_{i+m-1}\}$ 为方面词序列, 对语境序列中的任意单词 $w_j (1 \leq j \leq n)$, D 的计算方法为

$$D_j = \left| j - \left(i + \left\lfloor \frac{m}{2} \right\rfloor \right) \right| - \left\lfloor \frac{m}{2} \right\rfloor \quad (12)$$

获得每个单词对应的语义相关距离之后, 需要设置语义相关距离的阈值 T_D , 之后即可进行语境动态加权计算。对于语境序列中的任意单词 w_j , 其语义编码向量所对应的权重 s_j 计算方式为

$$s_j = \begin{cases} 1, & D_j \leq T_D \\ 1 - \frac{D_j - T_D}{n}, & D_j > T_D \end{cases} \quad (13)$$

对语境序列的每个单词按顺序计算权重, 得到权重矩阵 $S = (s_1, s_2, \dots, s_n)$ 。语境动态加权的最终输出 O^w , 即为语境序列的语义编码 O_c^e 与 S 相乘的结果:

$$O^w = O_c^e S^T \quad (14)$$

2) AOA 机制。

AOA 机制是 DFAOA-BERT 注意力机制的主体部分, 局部特征提取器和全局特征提取器都以此为核心。其主要设计思想与计算方法在第 1 节已作详细阐述。由于 2 个特征提取器在该部分的运算方式基本一致, 下面以局部特征提取器为例进行具体说明。

考虑到 BERT 词嵌入的维数较高, 进入 AOA 模块后运算速度很慢, 所以, DFAOA-BERT 首先使用双向长短时记忆网络 (bi-directional long

short-term memory, Bi-LSTM) 压缩维数、提取语义信息。令 h_c 和 h_a 分别为 Bi-LSTM 最后输出的语境隐藏状态和方面词隐藏状态, 其计算方法为

$$\begin{aligned} \vec{h}_c &= \text{LSTM}(O^w) \\ \overleftarrow{h}_c &= \text{LSTM}(O^w) \\ h_c &= \left[\vec{h}_c, \overleftarrow{h}_c \right] \end{aligned} \quad (15)$$

$$\begin{aligned} \vec{h}_a &= \text{LSTM}(O_a^e) \\ \overleftarrow{h}_a &= \text{LSTM}(O_a^e) \\ h_a &= \left[\vec{h}_a, \overleftarrow{h}_a \right] \end{aligned} \quad (16)$$

式中: $h_c \in \mathbf{R}^{n \times 2d_h}$; $h_a \in \mathbf{R}^{m \times 2d_h}$; d_h 是 Bi-LSTM 的隐藏状态维数。

获得以上结果之后, 根据第 1 节阐述的 AOA 原理, 计算得出最终的 AOA 注意力数值 γ 。

2.5 分类层

设 h_c^L 、 h_c^G 分别为局部特征提取器和全局特征提取器中语境序列的 Bi-LSTM 隐藏状态输出, γ^L 、 γ^G 分别是局部特征提取器和全局特征提取器中的 AOA 数值, DFAOA-BERT 最终的语义表示为

$$r^L = (h_c^L)^T \gamma^L \quad (17)$$

$$r^G = (h_c^G)^T \gamma^G \quad (18)$$

$$r = r^L \oplus r^G \quad (19)$$

本模型将 r 视为情感分类所依据的最终分类特征, 通过一个线性网络将其映射到任务所要求的结果空间, 并使用 softmax 来计算方面词所对应的情感态度为 z 的可能性:

$$X = W_l r + b_l \quad (20)$$

$$P(x = z) = \frac{\exp(X_z)}{\sum_{i=1}^Z \exp(X_i)} \quad (21)$$

式中: W_l 、 b_l 分别是线性网络的权重矩阵和偏移。Z 是情感分类结果的种类数, 在本任务中, 情感态度共有负面、中立、正面 3 种, 故 $Z = 3$ 。

在模型的训练方面, DFAOA-BERT 采用交叉熵损失函数来计算训练损失, 同时应用 L_2 正则化来抑制过拟合现象。损失函数为

$$\mathcal{L} = - \sum_i \sum_{z \in Z} I(y_i = z) \log(P(y_i = z)) + \lambda \|\theta\|^2 \quad (22)$$

式中: $I(\cdot)$ 是指示函数; λ 是 L_2 正则项; θ 是 DFAOA-BERT 模型中所有可训参数的集合。

3 实验结果与分析

3.1 数据集信息

本文在 3 个公开的方面词情感分析数据集上进行了实验: SemEval 2014 任务 4 中的餐馆评论数据集和笔记本评论数据集^[18], 以及 ACL-14

Twitter 社交评论数据集^[19](为方便下文说明, 将这 3 个数据集分别简称为餐馆数据集、笔记本数据

集和 Twitter 数据集)。表 1 给出了各数据集的训练评论数与测试评论数。

表 1 数据集组成信息

Table 1 Composition information of the datasets

数据集	正面		中立		负面	
	训练	测试	训练	测试	训练	测试
餐馆	2164	728	637	196	807	196
笔记本	994	341	464	169	870	128
Twitter	1561	173	3127	346	1560	173

3.2 基线

为了全方位评价 DFAOA-BERT 的表现效果, 本文选用了多个基线模型进行结果对比。这些基线模型基于不同的设计思想与技术原理, 包括普通神经网络模型与含 BERT 加持的深度学习模型。

1) 支持向量机^[20](support vector machine, SVM) 是应用传统机器学习解决方面词情感分析任务的代表模型。它使用支持向量机技术来结合特征工程进行情感分析。

2) 基于目标的长短时记忆网络模型 (TD-LSTM)^[5] 分别通过 2 个 LSTM 来提取方面词左右两边的语义特征, 将两部分语义信息拼接后计算情感数值。

3) 互动注意力网络 (IAN)^[6] 使用 2 个 LSTM 来分别获取语境和方面词的语义表示。它根据语境的语义表示的隐藏层状态来构造对方面词的注意力权重, 反之即可构造方面词对语境的注意力权重, 取得了不错的效果。

4) 循环注意力记忆网络^[21](recurrent attention network on memory, RAM) 使用 Bi-LSTM 来进行文本记忆表示, 同时创新性地使用了一个循环单元网络来将语境特征的多个注意力输出组合在一起。

5) 嵌套注意力^[3](AOA) 综合考虑到方面词的语义信息和在整个评论中的位置特征, 不仅能生成方面词对整个语境的注意力权重, 还能计算出语境对方面词的注意力权重。

6) 基础 BERT^[11](BERT-base) 是基础的预训练 BERT 模型。即使不添加任何额外模块, 该模型的表现依旧非常突出。

7) 带有 BERT 的注意力编码网络^[13](AEN-BERT) 的设计基于预训练 BERT 模型, 其特点在于完全使用注意力机制来编码语境和方面词的语义表示。

8) 用于句对分类的 BERT 模型^[13](BERT-

SPC) 使用 BERT 模型来处理句对分类任务, 将该模型应用于基于方面词的情感分析任务时, 需要将模型的输入处理成“[CLS] + 整个句子 + [SEP] + 方面词 + [SEP]”的形式。

9) 后训练 BERT^[22](post-trained BERT, BERT-PT) 模型在方面词情感分析任务上探索了 BERT 模型的后训练策略, 成功地提升了 BERT 模型的精调效果。

10) BERT 对抗训练^[12] (BERT adversarial training, BAT) 模型将计算机视觉领域常常涉及的对抗学习思想应用到方面词情感分析任务上, 其训练损失包括主模型的损失与对抗模型的损失这 2 部分。

3.3 实验参数与结果对比

实验中涉及的模型参数设置如下: DFAOA-BERT 在语境动态加权模块定义的语义相关距离阈值 T_D 设为 5。BERT 版本选用 uncased BERT-base, 词嵌入的维数为 768。语境动态加权模块中 Bi-LSTM 的隐藏层维数设为 150。模型中的可训练参数使用 Glorot^[23] 初始化方法来进行初始化, 学习率设为 $2e^{-5}$ (在 Twitter 数据集上实验时设为 $5e^{-6}$), L_2 正则项设为 $1e^{-5}$, dropout 率^[24] 则根据不同数据集进行灵活调整, 餐馆数据集设为 0.1, 笔记本数据集设为 0, 而 Twitter 数据集则设为 0.5。模型采用 Adam 优化器^[25] 来最小化训练损失。由于模型参数数量较大, 为防止显卡内存不足, 每批训练评论数设为 16。模型的评价方法选用准确度与 Macro- F_1 , 两者数值越高则说明模型的情感分析能力越优秀。

实验结果如表 2 所示。表 2 中的 Acc 表示准确度, MF_1 表示 Macro- F_1 , 两者均为模型的评价指标。对比模型的结果取自己发表的论文, “-”指原论文没有发表该结果, 其中, BERT-base 的结果取自 Karimi 等^[12] 的实验结果。DFAOA-BERT 的结果在显卡 NVIDIA Tesla P100 上运行获得。最

优秀的结果加粗表示。可以看到,在 3 个数据集上, DFAOA-BERT 均取得了非常优秀的成绩。

表 2 实验结果对比

Table 2 Comparison of the experimental results

%

类别	模型	餐馆数据集		笔记本数据集		Twitter数据集	
		Acc	MF ₁	Acc	MF ₁	Acc	MF ₁
基线	SVM	80.16	—	70.49	—	63.40	—
	TD-LSTM	75.63	—	68.13	—	70.80	69.00
	IAN	78.60	—	72.10	—	—	—
	RAM	80.23	70.80	74.49	71.35	69.36	67.30
	AOA	81.20	—	74.50	—	—	—
BERT模型	BERT-base	81.54	71.94	75.29	71.91	—	—
	AEN-BERT	83.12	73.76	79.93	76.31	74.71	73.13
	BERT-SPC	84.46	76.98	78.99	75.03	73.55	72.14
	BERT-PT	84.95	76.96	78.07	75.08	—	—
	BAT	86.03	79.24	79.35	76.50	—	—
本文模型	DFAOA-BERT	87.05	81.68	80.25	76.66	75.72	74.62

普通基线模型中,基于特征工程的 SVM 提出时间最早,但其效果却出人意料得好,超过了多数不含 BERT 加持的基线模型。因此,如果不考虑人力成本, SVM 在特定数据集上将会是一个不错的选择。TD-LSTM 模型的结构最为简单,只使用了 2 个 LSTM,但其效果也是最不理想的。而引入注意力机制的模型,例如 IAN、AOA,通过精心设计的注意力机制来引导网络进行语义特征提取,从而获得比 TD-LSTM 优秀很多的结果。

根据表 2 可以观察到,就餐评论数据集而言,80% 的分类准确度已经接近模型效果的天花板了,而 BERT 模型的加持,则使情感分析结果获得了惊人的提升。即使是最简单的 BERT-base 模型,其效果已经超过了绝大部分普通基线模型。AEN-BERT、BERT-SPC 和 BAT 为了充分发挥

BERT 的强大优势,设计了合适的网络结构和子模块来引导整个模型满足方面词情感分析任务的要求,取得了令人瞩目的成绩。DFAOA-BERT 则凭借双特征 AOA 机制,无论是在准确度还是 Macro-F₁ 分数上,都获得了比上述基线更理想的结果,证明了整体模型在解决方面词情感分析任务上的优秀能力。

3.4 子模块的有效性实验

为验证 DFAOA-BERT 模型中主要子模块的设计合理性,本文进行了相应的补充实验。结果如表 3 所示。其中,“w/o”(without)代表去除,“G”代表全局特征提取器,“L”代表局部特征提取器,“cdw”代表语境动态加权模块。所有实验均在显卡 NVIDIA Tesla P100 上进行。最优秀的结果加粗表示。

表 3 DFAOA-BERT 模型与其变种的实验结果对比

Table 3 Comparison of experimental results between DFAOA-BERT model and its variants

%

模型	餐馆数据集		笔记本数据集		Twitter数据集	
	Acc	MF ₁	Acc	MF ₁	Acc	MF ₁
DFAOA-BERT w/o cdw	84.64	76.93	79.00	75.06	75.00	73.82
DFAOA-BERT w/o L	84.55	77.49	79.47	76.14	73.70	72.48
DFAOA-BERT w/o G	83.21	75.56	79.78	76.38	73.27	71.10
DFAOA-BERT	87.05	81.68	80.25	76.66	75.72	74.62

可以清楚地看到,缺少任一模块, DFAOA-BERT 模型的效果都将有着明显的下降。全局特征提取器的缺失对整体效果影响最大,这说明了全局语义特征对于本任务的重要性。最难判断的句子往往具有一个特点,即情感指示词会出现在模型意想不到的位置,因此,仅依靠局部特征提

取器将会因位置问题而在很大概率上忽视情感指示词的作用,导致情感判断产生关键性错误。

在笔记本评论数据集和 Twitter 评论数据集上,局部特征提取器和全局特征提取器的缺失所带来的效果下降幅度相近。这两者在整体模型中有着相辅相成的关系,全局特征提取器能够顾及

全面的语义信息但却没有突出重点, 局部特征提取器则相反, 可以充分捕捉方面词邻近语境的语义特征但却难以考虑全面。两个提取器对 DFAOA-BERT 的作用都非常关键。而语境动态加权模块作为局部特征提取器所独有的注意力机制设计, 较为显著地提升了局部特征提取器的表现效果。

通过上述补充实验, 可以验证 DFAOA-BERT 各子模块在本任务上的有效性。

3.5 结果分析与改进

根据上述的实验结果, 可以发现, 与之前的模型相比, DFAOA-BERT 的综合表现较为优异。无论是全局、局部特征提取器的设计, 还是语境动态加权模块的引入, 都对整体模型的效果有着显著的正面作用。

但 DFAOA-BERT 依然存在一定的改进空间。对实验结果进行横向对比, 可以观察到, DFAOA-BERT 在餐馆数据集上的表现非常优异, 而在处理笔记本数据集和 Twitter 数据集时会呈现出明显的准确度下降。出现该问题的主要原因在于, 笔记本数据集和 Twitter 数据集中存在较多拼写错误与无意义的标记, 从而对模型的语义特征抽取和情感分类造成了很大的困难。将 AEN-BERT 与 DFAOA-BERT 进行对比, 可以观察到, 虽然 DFAOA-BERT 在餐馆数据集上有着显著的优势, 但在笔记本和 Twitter 数据集上的表现却和 AEN-BERT 差别不大。这是因为 AEN-BERT 在语义编码和特征抽取模块上有着更为合适的设计, 它针对方面词情感分析任务专门设计了一套注意力机制来生成语义表示并提取特征, 而非使用大多数模型所选取的 LSTM。因此, DFAOA-BERT 在处理低质量数据集方面仍存在上升空间, 可通过设计合理的注意力机制来改进语义编码和特征抽取模块, 减小模型对数据质量的依赖程度。

此外, BERT 的领域内精调以及方面词抽取模块的引入, 也是提升 DFAOA-BERT 效果的两种改进方向。BERT 的领域内精调是指对基础的 BERT 模型进行一定调整, 使用额外的领域内语料资源来增强 BERT 在该领域的语义表示能力, 从而帮助整体模型获得更优秀的情感判断能力。方面词抽取模块的引入则相当于使用一个模型来同时完成方面词情感分析任务和方面词抽取任务, 而这两个任务所涉及的模型参数存在相辅相成的效应, 在 Li 等^[26] 以及 Hu 等^[27] 的实验中被证明能够有效提升整体模型的效果。因此, DFAOA-BERT 可通过以上两种改进思路, 增强方面词情感分析能力。

4 结束语

本文针对当前方面词情感分析任务研究中存在的不足, 设计了 DFAOA-BERT 模型。该模型首次将 AOA 与 BERT 预训练语言模型相结合, 能够有效提取指定方面词的情感特征。而全局和局部特征提取器的设计, 则使得模型既能顾及整个句子的全局语义信息, 还能有目的地弱化情感干扰项的不利影响。实验证明, DFAOA-BERT 在 SemEval 2014 任务 4 中的餐馆评论数据集、笔记本评论数据集以及 ACL-14 Twitter 社交评论数据集上均取得了优异的成绩, 而子模块的有效性实验也充分说明了 DFAOA-BERT 各个模块设计的合理性, 任何部分的缺失都将对模型的效果产生显著的不利影响。

尽管 DFAOA-BERT 模型在多个数据集上都还有着不错的表现, 但它仍存在一定的提升空间。本文分析了一些模型的设计优势, 阐释了 DFAOA-BERT 存在的不足, 并从多个角度提出了改进方向与策略方案, 以期增强 DFAOA-BERT 的情感判断能力与处理低质量数据时的稳定性。

参考文献:

- [1] LIU Bing. Sentiment analysis and opinion mining[J]. *Synthesis lectures on human language technologies*, 2012, 5(1): 1-167.
- [2] PANG B, LEE L. Opinion mining and sentiment analysis[M]. *Foundations and Trends in Information Retrieval*, 2008: 1-135.
- [3] HUANG Binxuan, OU Yanglan, CARLEY K M. Aspect level sentiment classification with attention-over-attention neural networks[C]//*Proceedings of the 11th International Conference on Social, Cultural, and Behavioral Modeling*. Washington, USA, 2018: 197-206.
- [4] SUTSKEVER I, VINYALS O, LE Q V. Sequence to sequence learning with neural networks[C]//*Proceedings of the 27th International Conference on Neural Information Processing Systems*. Montreal, Canada, 2014: 3104-3112.
- [5] TANG D, QIN B, FENG X, et al. Effective LSTMs for target-dependent sentiment classification[C]//*Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics*. Osaka, Japan, 2016: 3298-3307.
- [6] MA D, LI S, ZHANG X, et al. Interactive attention networks for aspect-level sentiment classification[C]//*Proceedings of the 26th International Joint Conference on Artificial Intelligence*. Melbourne, Australia, 2017:

- 4068–4074.
- [7] XU Weidi, TAN Ying. Semi-supervised target-oriented sentiment classification[J]. *Neurocomputing*, 2019, 337: 120–128.
- [8] PETERS M, NEUMANN M, IYYER M, et al. Deep contextualized word representations[C] // Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics. New Orleans, Louisiana, 2018 : 2227–2237.
- [9] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C] // Proceedings of the 31st International Conference on Neural Information Processing Systems. Los Angeles, USA, 2017: 6000–6010.
- [10] RADFORD A, NARASIMHAN K, SALIMANS T, et al. Improving language understanding by generative pre-training[EB/OL]. [2019-5-10]. https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/language_understanding_paper.pdf.
- [11] DEVLIN J, CHANG M- W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[C] // Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics. Minneapolis, Minnesota, USA, 2019 : 4171–4186.
- [12] KARIMI A, ROSSI L, PRATI A. Adversarial training for aspect-based sentiment analysis with BERT[EB/OL]. [2019-5-10]. <https://arxiv.org/abs/2001.11316>.
- [13] SONG Youwei, WANG Jiahai, JIANG Tao, et al. Attentional encoder network for targeted sentiment classification[EB/OL]. [2019-5-10]. <https://arxiv.org/abs/1902.09314>.
- [14] JIANG Long, YU Mo, ZHOU Ming, et al. Target-dependent twitter sentiment classification[C] // Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. Portland, USA, 2011: 151–160.
- [15] CUI Y, CHEN Z, WEI S, et al. Attention-over-attention neural networks for reading comprehension[C] // Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. Vancouver, Canada, 2017 : 593 – 602.
- [16] PENNINGTON J, SOCHER R, MANNING C. Glove: global vectors for word representation[C] // Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar, 2014: 1532–1543.
- [17] ZENG Biqing, YANG Heng, XU Ruyang, et al. LCF: a local context focus mechanism for aspect-based sentiment classification[J]. *Applied sciences*, 2019, 9(16): 1–22.
- [18] PONTIKI M, GALANIS D, PAVLOPOULOS J, et al. SemEval-2014 task 4: aspect based sentiment analysis[C] // Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014). Dublin, Ireland, 2014: 27–35.
- [19] DONG Li, WEI Furu, TAN Chuanqi, et al. Adaptive recursive neural network for target-dependent twitter sentiment classification[C] // Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. Baltimore, Maryland, 2014: 49–54.
- [20] KIRITCHENKO S, ZHU Xiaodan, CHERRY C, et al. NRC-Canada-2014: detecting aspects and sentiment in customer reviews[C] // Proceedings of the 8th International Workshop on Semantic Evaluation. Dublin, Ireland, 2014: 437–442.
- [21] CHEN Peng, SUN Zhongqian, BING Lidong, et al. Recurrent attention network on memory for aspect sentiment analysis[C] // Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Copenhagen, Denmark, 2017: 452–461.
- [22] XU H, LIU B, SHU L, et al. BERT post-training for review reading comprehension and aspect-based sentiment analysis[C] // Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics. Minneapolis, Minnesota, USA, 2019 : 2324 – 2335.
- [23] GLOROT X, BENGIO Y. Understanding the difficulty of training deep feedforward neural networks[C] // Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. Chia Laguna Resort, Italy, 2010: 249–256.
- [24] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. *The journal of machine learning research*, 2014, 15(1): 1929–1958.
- [25] KINGMA D P, BA J. Adam: a method for stochastic optimization[C] // The 3rd International Conference for Learning Representations, San Diego. <http://arxiv.org/abs/1412.6980>.
- [26] LI Xin, BING Lidong, LI Piji, et al. A unified model for opinion target extraction and target sentiment prediction[J]. *Proceedings of the AAAI conference on artificial intelligence*, 2019, 33(1): 6714–6721.

- [27] HU M, PENG Y, HUANG Z, et al. Open-domain targeted sentiment analysis via span-based extraction and classification[C] // Proceedings of the 57th Conference of the Association for Computational Linguistics. Florence, Italy, 2019: 537–546.

作者简介:



肖宇晗, 硕士研究生, 主要研究方向为深度学习、数据挖掘和自然语言处理。



林慧苹, 副教授, 博士, 主要研究方向为大数据分析、企业信息服务, 主持和参与国家 863 计划、国家自然科学基金项目、国家重点研发计划项目等多项。发表学术论文 20 余篇。



谭莹, 教授, 博士生导师, 主要研究方向为智能科学、计算智能与群体智能、机器学习、人工神经网络、群体机器人、大数据挖掘。烟花算法发明人。吴文俊人工智能科学技术成就奖创新三等奖获得者。发表学术论文 330 余篇, 出版学术专著 12 部。

陈小平主编《人工智能伦理导引》出版

人工智能是 21 世纪最引人注目的重大科技进展, 被普遍认为是第四次工业革命的引导力量。人工智能与大数据、网络、信息、自动化、物联网和云计算等新技术相结合, 正渗透到工业、农业、服务业、交通运输、能源、金融、安全和国防等各个领域, 成为这些产业部门转型升级的核心驱动力。在这种新形势下, 人工智能及大数据等相关技术将对各行各业、各种岗位的从业人员产生广泛、深刻、长期的影响。由此出现一种普遍的现象: 所有社会成员都需要更多、更深入地了解人工智能, 全面、客观、准确地认识人工智能带给我们的正面效应和负面效应。

社会的上述巨大变化必然反映到教育实践中, 引出本科生和研究生教育的一种全新的重大需求——人工智能及其应用的快速发展, 要求对现有学校教育内容进行及时的扩展, 将人工智能和人工智能伦理纳入各个学科的教育体系, 开设相关的课程和培训, 使学生得到必要的培养和训练。为此中国人工智能学会人工智能伦理专委会(筹)主任陈小平领衔法律、哲学等社会科学相关领域一线专家共同编写了国内第一部跨学科、导引性人工智能伦理教材——《人工智能伦理导引》。本书面向所有专业的本科生和研究生, 对相关专业的研究者、相关领域的管理者和相关行业的从业者亦有普遍参考价值。

对人工智能基本原理、主要技术进展和当前挑战进行专业性、通俗化的概括与解读, 为人工智能伦理和治理的学习与探索奠定可靠的技术基础; 对当前人工智能伦理与治理的五个重点领域——数据、企业、AI 科研、传媒、法制进行专业解剖和案例分析, 为人工智能伦理与治理的理论研究和应用实践提供专业基础; 对人机社会技术伦理进行系统性的梳理与反思, 为人工智能伦理专业研究深造提供知识基础。

本书重要特色主要体现在:

跨学科与导引性 国内第一部由人工智能与法律、科技伦理、哲学等相关领域专家共同编写的跨学科、导引性人工智能伦理教材。

内容专业且有趣, 深刻且通俗 本书针对人工智能的广泛应用以及人工智能伦理和治理的普遍迫切需求, 采用通俗的语言, 传授精当的知识。内容专业深刻的同时, 与普遍应用和生活实际相联系, 由浅入深地进行探究。

章前导读、章后讨论, 阅读指引实用 每章开始前由案例导入, 引出章节主题。在进行细致学习后, 每章最后附有讨论与思考题, 发人深省, 激发读者对有关人工智能及科技伦理相关问题进行思索和探究。章节结构的设置为读者提供了实用的阅读和学习指引。



识别二维码, 即可购买