

基于高效 Transformer 的中文在线课程评论方面情感分析

潘芳¹ 张会兵² 董俊超² 首照宇³

1 广西师范大学职业技术师范学院 广西 桂林 541004

2 桂林电子科技大学广西可信软件重点实验室 广西 桂林 541004

3 桂林电子科技大学信息与通信学院 广西 桂林 541004

(panfang@mailbox.gxnu.edu.cn)

摘要 准确挖掘在线课程评论中蕴涵的情感信息对在线课程的健康发展极具价值。现有中文在线课程评论情感分析研究大多为分析整条评论句子情感极性的粗粒度模型,无法准确表达课程评论句子中各个方面的细粒度情感。为此,提出一种基于高效 Transformer 的中文在线课程评论方面情感分析模型。首先,通过 ALBERT 预训练模型获得评论文本方面和上下文的动态字向量编码;然后,采用可以并行输入字向量的高效 Transformer 分别对课程评论文本的方面和上下文进行语义表征;最后,使用交互注意力机制交互地学习课程评论文本中方面和上下文的重要部分,并输入方面和上下文的最终表示到情感分类层进行在线课程评论情感极性预测。在中国 MOOC 网真实数据集上的实验结果表明,高效 Transformer 中文在线课程评论方面情感分析模型与基线模型相比,在更低的时间开销下准确率达到了 80% 以上。

关键词: 在线课程;课程评论;方面情感分析;预训练语言模型;注意力机制

中图法分类号 TP391

Aspect Sentiment Analysis of Chinese Online Course Review Based on Efficient Transformer

PAN Fang¹, ZHANG Hui-bing², DONG Jun-chao² and SHOU Zhao-yu³

1 Teachers College for Vocational and Technical Education, Guangxi Normal University, Guilin, Guangxi 541004, China

2 Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China

3 School of Information and Communication, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China

Abstract It is of great value for the healthy development of online courses that accurately mine the emotional information contained in online course reviews. Most of the existing research on sentiment analysis of Chinese online course reviews is a coarse-grained model, which cannot accurately express the fine-grained sentiment for all aspects of the review sentence. The paper puts forward an efficient Transformer based sentiment analysis model for Chinese online course review. Firstly, the dynamic word vector coding of the review's aspect and context is obtained by the Albert pre-training model. Then, the semantic representation of the review's aspect and context is carried out by the efficient Transformer which can input the word vector in parallel. Finally, it uses the interactive attention mechanism to learn the important parts of the context and aspects in the course review, and puts its final representation into the sentiment classification layer to predict the sentiment polarity. Experimental results on real datasets of MOOC in China show that the accuracy of the proposed model achieves more than 80% at lower time cost compared with the baseline model.

Keywords Online course, Course review, Aspect-based sentiment analysis, Pre-training language model, Attention mechanism

1 引言

在线课程作为信息时代新兴的教育方式,凭借其免费、便利的优势吸引了国内外众多师生的广泛参与。评论作为在线课程最直接的互动反馈方式,深入挖掘其蕴涵的情感信息,对提升教学质量、降低辍学率和促进在线课程的可持续发展等多个方面具有重要意义^[1-2]。

目前在线课程情感分析方面主要有情感词典、机器学习和深度学习 3 类方法^[2]。在早期的在线课程情感分析研究中,众多学者提出基于情感词典的解决方案^[3-4]。由于此类方

法的性能严重依赖规则词典的质量且扩展性不好,为此,研究人员开始采用机器学习算法构建课程评论情感分析模型:标注一部分课程评论的情感极性作为训练集,使用训练完成的模型来对未标注的课程评论情感进行预测^[5-6]。其中被广泛使用的机器学习算法有逻辑回归、支持向量机(SVM)^[6]、随机森林等。此类方法在特征工程中消耗的人力资源和时间成本过高,目前已逐步被自动抽取特征的深度学习取代。在深度学习课程评论情感分析中,长短时记忆网络(LSTM)成为主流模型,如 LSTM-CNN^[7-8]和双向 LSTM(Bi-LSTM)^[9]等。此外,一些研究表明使用集成学习结合多种机器学习或深度

基金项目:国家自然科学基金(61662013,61967005,U1811264)

This work was supported by the National Natural Science Foundation of China(61662013,61967005,U1811264).

通信作者:张会兵(zhanghuibing@guet.edu.cn)

学习模型,能够提升在线课程评论情感分析模型的性能^[10-11]。现有的这3类方法均对整个课程评论句子进行粗粒度情感分析,不能细粒度分析在线课程中各个方面的情感。在一条包含多个情感不一致的方面的在线课程评论中往往会出分析错误的问题。例如:针对在线课程评论“课程资料丰富,老师讲的不给力啊!”,它包含“课程”和“老师”两个方面,并且这两个方面有相反的情感极性,但现有的情感分类模型会将其判断为积极的情感极性,而忽略了句子中“老师”方面的消极极性。

为此,本文提出一种基于高效 Transformer (E-Transformer)的在线课程评论方面情感分析模型。该模型采用 ALBERT 预训练模型动态编码在线课程评论方面和上下文的汉字^[12];使用 E-Transformer 并行地自动抽取在线课程评论中方面及上下文的语义特征;通过注意力机制交互学习评论文本中方面和上下文的重要汉字,并生成最终表示;采用情感分类层对在线课程评论中方面对应的情感极性进行分类。我们在中国 MOOC 网真实在线课程评论数据集上进行实验,结果表明 E-Transformer 模型与现有基线方面情感分析模型相比,在中文在线课程评论情感分析中有更高的准确率和更低的时间消耗。

2 相关工作

2.1 在线课程情感分析

与商品评论情感分析相比,在线课程评论数据的情感分析仍处于起步阶段。文献[13]通过 OpenNLP 词性标注分析器和 sentiWordNet 词汇资源计算词汇情感得分,对学生反馈评论进行情绪分析。Yang 等利用 TextRank 和 word2vec 模型构建在线课程评论情感词典,显著提高了课程评论领域的情绪识别和分类效果,为在线课程评价提供了分析工具和基础语料库^[14]。基于词典的课程情感分析方法可移植性差,受情感词典质量的影响较大,构建情感词典需要花费大量的时间和精力。基于机器学习的方法凭借泛化能力强、对短文本分类效果好的优势,吸引了许多研究者的关注。最大熵 (MaxEnt)^[2]、朴素贝叶斯 (NB)^[5] 和 SVM^[6] 等方法被证明具有较好的性能。此外,已有研究发现融合了词典知识的机器学习算法比单纯的机器学习模型准确率更高^[15]。然而,基于传统机器学习的性能对特征工程依赖较大,特征工程的构建需要消耗大量的人力资源,故不适用现如今海量在线课程评论数据的情感分析任务。因此,基于深度学习的课程评论情感分析模型成为现在的研究热点。文献[7]使用 NB 和 CNN-LSTM 两种模型分类智能学习系统中的情感极性,发现 CNN-LSTM 神经网络模型的效果比 NB 机器学习模型好。文献[9]使用 Bi-LSTM, MaxEnt 等多种模型对越南某大学学生的反馈语料库进行研究,根据多种评价标准对这些模型进行比较和评价,发现 Bi-LSTM 取得了最高的准确率。Kandhro 等通过预训练词向量模型表征评论文本中的语义和句法信息,采用 LSTM 神经网络模型分析学生反馈文本表达的情感^[16]。该模型克服了词袋模型和机器学习算法组合的情感分类模型存在句子序列信息丢失的问题,同时也验证了通过改进预训练词向量提升情感分类模型性能的可行性。该研究对本文采用 ALBERT 预训练模型动态编码在线课程评论方面和上下文具有重要启发,但该研究的静态预训练词向量模

型的精确度远低于根据上下文动态编码的 ALBERT 模型。

2.2 方面情感分析

方面情感分析 (Aspect sentiment analysis) 旨在分析方面实体的情感极性^[17]。近年来以循环神经网络 (RNN) 为代表的端到端神经网络模型,凭借其优秀的自动化特征抽取能力和良好的分类效果,吸引了众多研究者的关注^[18]。Dong 等首次将 RNN 应用到方面情感分析上,并借助句法结构信息辅助 RNN 提高情感分类准确率。该模型容易受句法分析质量的影响,在处理如 Twitter、评论等非书面表达文本时性能不稳定^[19]。为了进一步加强方面词与句子中每个令牌 (token) 的关联,Tang 等提出了方面相关 LSTM (TD-LSTM) 和方面连接 LSTM (TC-LSTM)^[20]。TD-LSTM 通过使用 LSTM 和 LSTM 两个 LSTM 网络,分别学习表示给定方面的左上文和右下文。连接 LSTM 最后的隐藏层向量输入到 softmax 层,预测句子所含方面的情感极性。TC-LSTM 通过合并方面连接组件进一步扩展 TD-LSTM,增强了 TD-LSTM 捕获方面与上下文的交互能力。Wang 等提出一种基于注意力机制的 ATAE-LSTM 模型,该模型输入方面表示和单词嵌入的连接,然后采用包含注意力机制的 LSTM 计算注意力向量与方面词向量,拼接后传入 softmax 层,求得方面的情感极性表达^[22]。Ma 等提出一种交互注意网络 (IAN),考虑了方面的注意力和所处的上下文语境。该网络使用两个基于注意力的 LSTM 交互地捕获方面短语和上下文的关键词^[23]。然而,上述这些方面情感分析模型并不完全适用于中文在线课程情感分析,原因包括:1) 现有的基于神经网络的目标情感分析模型需要大量的标注训练集用于模型训练泛化,而中文在线课程评论方面及其情感极性的人工标注成本极高;2) LSTM 无法并行抽取方面和上下文的语义信息,模型训练时间开销较大。因此,本文采用 BERT 预训练模型和高效 Transformer 对 IAN 模型进行改进,在少量标注数据的情况下实现中文课程评论方面情感的高效分析。

2.3 预训练语言模型

预训练语言模型在自然语言处理领域的多项任务中均有良好表现,受到了学术界和工业界的广泛关注。其中 Google 在 2018 年推出的 BERT 模型为目前使用广泛的预训练模型之一,被广泛应用于阅读理解、情感分析、序列标注等众多任务^[24]。该模型采用含有 self-attention 机制的 Transformer 编码器,以实现所有层中语言表示的双向学习,解决了 GPT 和 ELMo 等这类单向语言模型无法双向结合上下文有效信息的局限性。但 BERT 模型庞大的参数量和资源消耗在中文在线课程评论方面情感分析中并不适用。因此本文采用一种轻量版的 BERT 模型 (ALBERT) 编码中文在线课程评论的方面和上下文。ALBERT 通过层参数共享、嵌入因式分解、句内连贯损失 3 种机制在保证 BERT 模型性能的同时大大降低了参数量和资源消耗^[12]。

3 问题定义与模型

3.1 问题定义

在分析在线课程评论方面情感时,给出一条包含 n 个汉字的评论句子 $s = \{w_1, w_2, \dots, w_n\}$, 其中 $a = \{w_1, w_2, \dots, w_m\}$ 为句子中包含 m 个汉字的方面。方面可以是一个汉字或者多个汉字组成的短语。本文的主要目标是推断课程评论句子 s 中方面 a 的情感极性。情感极性可能是 {积极、消极、中

性)。例如在课程评论句子“课程资料丰富,老师认真负责,如果能加上字幕就更加完美了。”中,包含“课程”“老师”和“字幕”3个方面。我们的目标是分析该句中3个方面的情感极性,期望输出结果分别是积极、积极和消极。

3.2 E-Transformer 中文在线课程评论方面情感分析模型

E-Transformer 中文在线课程评论方面情感分析模型的总体架构如图1所示。它主要由将中文在线课程评论方面和上下文编码为字向量的字嵌入层、分别抽取方面和上下文语义信息的语义表征层、交互学习方面与上下文情感信息相关性的注意力层和分类输出情感极性的分类输出层4部分组成。

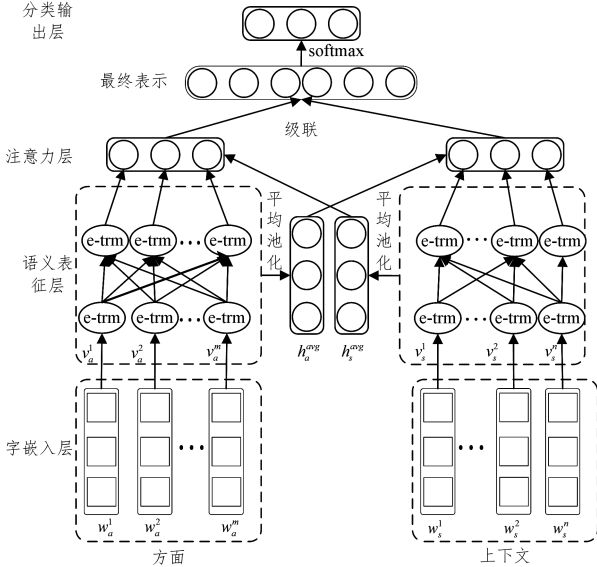


图1 E-Transformer 中文在线课程评论方面情感分析模型
Fig.1 E-Transformer for sentiment analysis model of Chinese online course review

3.2.1 字嵌入层

字嵌入层通过 ALBERT 预训练模型,将在线课程评论句子中每个汉字 w_i 映射为一个低维向量 $v_i \in \mathbb{R}^{d_w}$, 其中 d_w 为字向量维度。字嵌入层的输出为在线课程评论上下文向量 $\{v_1^c, v_2^c, \dots, v_n^c\} \in \mathbb{R}^{n \times d_w}$ 和方面向量 $\{v_1^a, v_2^a, \dots, v_m^a\} \in \mathbb{R}^{m \times d_w}$ 。

3.2.2 语义表征层

语义表征层由 n 层 E-Transformer 编码器组成^[25],如图2所示。其中,E-Transformer 编码器主要包括位置编码、稠密合成器^[26]、可逆残差网络^[27]和分块前馈全连接网络^[27]4部分。

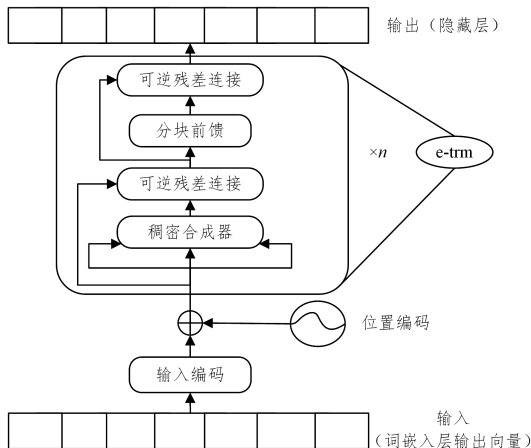


图2 E-Transformer 编码器结构图
Fig.2 Structure of E-Transformer encoder

(1)位置编码

由于 E-Transformer 编码器不能编码在线课程评论句子序列的顺序,字嵌入层的输出向量在传入 E-Transformer 编码器进行语义表征时需额外引入维度相同的位置编码,标记在线课程评论句子中汉字的位置信息。位置编码通过式(1)正弦函数和式(2)余弦函数所得。

$$P_{(p,2i)} = \sin(p/1000^{2i/d}) \tag{1}$$

$$P_{(p,2i+1)} = \cos(p/1000^{2i/d}) \tag{2}$$

其中, p 为在线课程评论句子中汉字位置, i 为字向量维度。

(2)稠密合成器

Transformer 编码器主要依靠多头自注意力(multi-head attention)机制表征在线课程评论句子中每个汉字与其余汉字的相互关系,但这种汉字对汉字的注意力由查询向量和关键字向量的数量积获得,计算复杂度较高。为了进一步提升计算效率,E-Transformer 采用稠密合成器(Dense synthesizer)替换 Transformer 多头自注意力机制。

首先使用如式(3)所示的参数化函数 $F(\cdot)$ 学习输入在线课程评论句子中每个汉字的权重 B_i 。

$$B_i = F(X_i) = W(\sigma_R(W(X_i) + b)) + b \tag{3}$$

其中, X_i 为输入在线课程句子中第 i 个汉字, $\sigma_R(\cdot)$ 为 ReLU 激活函数, b 为偏置。

然后,计算在线课程句子中每个汉字在稠密合成器的输出,如式(4)所示:

$$Y_i = \text{Softmax}(B_i)G(X_i) \tag{4}$$

其中, $G(X_i)$ 是类似于 Transformer 编码器值向量的参数化函数。

为了防止稠密合成器带来的参数规模增加问题,引入矩阵分解的方法降低参数规模,提升稠密合成器的防止过拟合能力。稠密合成器的因式分解变量可以表示为:

$$A, B = F_A(X_i), F_B(X_i) \tag{5}$$

其中, $F_A(\cdot)$ 和 $F_B(\cdot)$ 表示将 X_i 分别映射到维度 a 和 b , $a \times b = l$, l 为在线课程评论句子长度。

因式分解后稠密合成器的输出如式(6)所示:

$$Y_i = \text{Softmax}(C_i)G(X_i) = \text{Softmax}(H_A(A_i) * H_B(B_i))G(X_i) \tag{6}$$

其中, $H_A(\cdot)$ 和 $H_B(\cdot)$ 为平铺(tiling)函数。

(3)可逆残差连接

在 Transformer 的基本单元中,注意力层和全连接层通过残差网络连接。随着 Transformer 层数的增加,残差网络需要记录的每层输入结果太多,从而导致内存消耗过大。E-Transformer 采用可逆残差网络连接稠密合成器和全连接层。可逆残差网络根据最终层输出结果反推出网络结构中间层的结果,无需记录每一层的输入,内存消耗低。其计算公式如下:

$$Y_1 = X_1 + DS(X_1) \tag{7}$$

$$Y_2 = X_2 + FFN(X_2) \tag{8}$$

其中, $DS(\cdot)$ 为稠密合成器, $FFN(\cdot)$ 为前馈神经网络。

(4)分块前馈

E-Transformer 基本单元中的前馈神经网络的深度远大于稠密合成器的深度,会消耗大量的内存资源。前馈神经网络

络的计算在句子中是完全相互独立的,为了进一步降低内存消耗,对其在所有位置的批处理操作进行分块处理,如式(9)所示:

$$Y_2 = [Y_2^{(1)}; \dots; Y_2^{(c)}] \\ = [X_2^{(1)} + FFN(Y_1^{(1)}); \dots; X_2^{(c)} + FFN(Y_1^{(c)})] \quad (9)$$

其中, c 为分块数量。

在获得在线课程评论句子上下文和方面的字向量编码后,将这两组向量分别输入到 E-Transformer 编码器中表征汉字在句子和方面中蕴涵的语义,分别输出式(10)和式(11)的隐藏层表征。在线课程评论句子上下文和方面中的每个汉字经过 E-Transformer 编码器中稠密合成器后分别含有线课程评论句子上下文和方面中任意其余汉字的语义信息。

$$h_s = [h_s^1, h_s^1, \dots, h_s^n] \\ = E-Transformer(v_s^1, v_s^2, \dots, v_s^n) \quad (10)$$

$$h_a = [h_a^1, h_a^1, \dots, h_a^m] \\ = E-Transformer(v_a^1, v_a^2, \dots, v_a^m) \quad (11)$$

其中, $h_i^s \in \mathbb{R}^{d_h}$, $h_i^a \in \mathbb{R}^{d_h}$, d_h 为隐藏层向量维度。

3.2.3 注意力层

在中文在线课程评论方面情感分析中,在线课程评论句子上下文与方面间的交互关系至关重要^[23]。使用注意力机制捕捉在线课程评论句子上下文与方面的交互信息。首先采用平均池化方法求得方面和上下文的表示 h_a^{avg} 和 h_s^{avg} , 如式(12)和式(13)所示。然后依据式(14)和式(15)分别计算在线课程评论上下文和方面的注意力向量。

$$h_a^{avg} = \frac{1}{m} \times \sum_{i=1}^m h_a^i \quad (12)$$

$$h_s^{avg} = \frac{1}{n} \times \sum_{i=1}^n h_s^i \quad (13)$$

$$\alpha_i = \frac{\exp(\gamma(h_s^i, h_a^{avg}))}{\sum_{j=1}^n \exp(\gamma(h_s^j, h_a^{avg}))} \quad (14)$$

$$\beta_i = \frac{\exp(\gamma(h_a^i, h_s^{avg}))}{\sum_{j=1}^m \exp(\gamma(h_a^j, h_s^{avg}))} \quad (15)$$

其中, $\gamma(\cdot)$ 为得分函数。 h_s^i 表示在线课程评论句子上下文第 i 个汉字对方面的重要性; h_a^i 表示在线课程评论句子方面第 i 个汉字对上下文的重要性。

$$\gamma(h_s^i, h_a^{avg}) = \tanh(\mathbf{W} \cdot h_s^i \cdot h_a^{avg} + \mathbf{b}) \quad (16)$$

$$\gamma(h_a^i, h_s^{avg}) = \tanh(\mathbf{W} \cdot h_a^i \cdot h_s^{avg} + \mathbf{b}) \quad (17)$$

其中, \mathbf{W} 为权重矩阵, \mathbf{b} 为偏置矩阵。

注意力层的最终输出分别为 $r_s = \sum_{i=1}^n \alpha_i h_s^i$ 和 $r_a = \sum_{i=1}^m \beta_i h_a^i$ 。

3.2.4 分类输出层

分类输出层的主要任务是级联注意力层的输出,如式(18)所示。输入级联后的结果到 softmax 层来预测在线课程评论中方面对应的最终情感极性,如式(19)所示。

$$r = r_s \oplus r_a \quad (18)$$

$$p = \text{softmax}(\mathbf{w} \cdot r + \mathbf{b}) \quad (19)$$

其中, $p \in \mathbb{R}^C$ 为在线课程评论方面情感极性的概率分布, \mathbf{w} 和 \mathbf{b} 分别为权重矩阵和偏置, $C=3$ 表示在线课程评论方面所对应的 3 类情感极性。

模型训练通过交叉熵损失函数和 L_2 正则化,如式(20)所示。

$$L = - \sum_{i \in C} \sum_{c \in C} [((y_i = c) \cdot \log(p(y_i = c)) + \lambda \|\theta\|^2)] \quad (20)$$

其中, $y_i = c$ 为在线课程评论方面真实情感极性, $p(y_i = c)$ 为在线课程评论方面预测情感极性, λ 为正则化参数。

4 实验

4.1 数据集

采用 Python 爬虫程序收集了国内用户数量最多的中国 MOOC 网计算机、外语两大门类的 121 门国家精品课的真实中文课程评论数据 15000 条。在中文在线课程评论方面情感分析中,关键的前提条件就是要有一个无超链接、俚语等噪音数据的规则数据集。因此,对收集的中国 MOOC 中文在线课程评论的数据集进行预处理操作,主要包括:1) 标点符号和特殊字符。在线课程评论中存在滥用标点符号和特殊字符的问题,使用正则表达式删除评论文本中所有的标点符号和特殊字符。2) 超链接。观察获得的评论数据发现存在与情感表达无关的超链接,采用正则表达式进行删除。3) 缩写和俚语。中文在线课程评论中存在很多非正式的缩写和俚语。如“666”“稀饭”等。为了保证模型能够正确识别评论想要表达的正确含义,使用含义相同的规范表达替换评论文本中所有的缩写和俚语。4) 感叹词。对中文在线课程评论中存在的“哈哈哈哈哈”“好好好好”等感叹词,采用正则表达式来规范化处理。预处理完成后我们邀请了计算机专业和外语专业的 30 位同学,分为 6 组对中文在线课程评论的方面及其对应情感极性进行手动标注,标注示例如表 1 所列。每位标注人员对每条评论中含有的方面{老师,课程,平台}及其对应情感{积极,中性,消极}给出一个明确的结论。如果 5 位标注人员结论高度一致,就保留该条评论到最终的实验数据集中。数据集的详细信息如表 2 所列。

表 1 中文在线课程评论方面及对应情感标注示例

Table 1 Examples of Chinese online course reviews aspect and sentiment annotations

Review text	Aspect	Aspect-based Sentiment polarity
老师棒棒哒	老师	积极
课程覆盖面很全,定义等理论讲解清晰明了证明客观全面条理清晰简单易懂	课程	积极
速度有点快平台没有提供课件需要反复暂停记一些东西	平台	消极

表 2 实验数据集统计

Table 2 Experimental dataset statistics

Dataset	Positive	Neutral	Negative	Total
Computer training set	3048	517	1053	4618
Computer test set	1012	139	969	2120
Foreign language training set	2534	280	507	3321
Foreign language test set	1483	100	335	1918

4.2 评价指标与参数设置

实验评价指标选用准确率,如式(21)所示:

$$Accuracy = \frac{TP + TN}{(TP + FP) + (TN + FN)} \quad (21)$$

其中,真正例(TP)、假正例(FP)、真负例(TN)、假负例(FN)为根据样例 E-Transformer 预测情感极性类别与真实情感极性类别组合划分的 4 种类型。

E-Transformer 的参数设置如表 3 所列。实验所选基线模型的参数设置除将原始英语词向量模型统一改为 GloVe 中文词向量外,其余参数均采用模型的默认参数。

表 3 实验参数设置

Table 3 Experimental parameter setting

Parameter name	Parameter value
Word vector	ALBERT-Base ^[12]
Vector dimension	768
Maximum sequence length	128
Batch size	64
Learning rate	1×10^{-4}
training step	5
Layer number	4

4.3 实验结果与分析

4.3.1 准确性分析

为了验证 E-Transformer 中文在线课程评论方面情感分析模型的准确性,将其与以下 5 种基线模型进行比较。

1) LSTM: 对句子中的令牌进行编码后作为模型的输入,得到隐藏状态的平均向量,将其输入 softmax 进行情感极性分类^[20]。

2) TD-LSTM: 使用两个 LSTM 网络从左往右和从右往左分别对方面的左上文和右下文的信息建模,连接两个 LSTM 最后一个隐藏状态,输入 softmax 进行情感极性分类^[20]。

3) AT-LSTM: 首先学习每个方面的编码,然后通过注意力机制结合进 LSTM 模型进行情感极性的分类^[22]。

4) ATAE-LSTM: AT-LSTM 的扩展模型,与 AT-LSTM 模型不同之处在于模型输入时将方面嵌入附加到句子中每个词的嵌入中^[22]。

5) IAN: 使用两个包含注意力机制的 LSTM 模型对方面词和上下文进行交互学习,融合方面词与上下文表征,预测方面在上下文语境中的情感极性^[23]。

表 4 列出了 6 种方面情感分析模型在两个中文在线课程评论数据集上的实验结果。可以看出,E-Transformer 模型比基线模型中表现最好的 IAN 在两个数据集上的准确率分别提升了约 6.2% 和 9.7%。这是因为 E-Transformer 中的稠密合成器能够捕捉中文在线课程评论句子中每两个汉字间的相互关系,语义抽取能力强于 IAN 模型中的 LSTM。此外,本文采用的 ALBERT 预训练模型能够根据上下文语义对同一汉字进行含义不同的动态编码,比 IAN 模型中采用的 GloVe 静态编码词向量模型准确性更好。

表 4 6 种方面情感分析模型在两个数据集上的准确率对比

Table 4 Comparison of accuracy of six sentiment analysis models in two datasets

Model	Computer	Foreign language
LSTM	0.683	0.655
TD-LSTM	0.704	0.676
AT-LSTM	0.732	0.689
ATAE-LSTM	0.759	0.717
IAN	0.786	0.741
E-Transformer	0.835	0.813

在所选的基线模型中,忽略方面词的 LSTM 表现最差。考虑方面词的 TD-LSTM 相比 LSTM 模型在两个数据集上的准确率分别提升约 3.0% 和 3.1%。其原因是忽略方面词的 LSTM 无法准确区分句子中的方面词和上下文词,TD-LSTM 充分考虑了方面词的上下文信息。此外,由于注意力机制的引入,AE-LSTM 和 ATAE-LSTM 的效果优于 TD-LSTM。究其原因是融合注意力机制的模型能够很好地挖掘与方面密切相关的上下文信息。IAN 能取得比 AE-LSTM 和 ATAE-LSTM 更好的效果,这是因为 IAN 通过注意力机制将方面和上下文进行交互建模,获取到的句子语义更加准确。

4.3.2 时间性能分析

为了验证 E-Transformer 在线课程评论方面情感分析模型的时间性能,将 E-Transformer 模型与 5 种基线模型在相同的实验环境下进行比较。6 种模型在计算机数据集一次迭代训练花费时间的实验结果如图 3 所示。

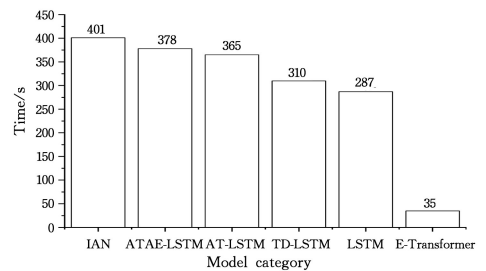


图 3 6 种模型迭代训练一次花费的时间

Fig. 3 Iterative training of six models takes time at a time

从图 3 可以看出,E-Transformer 模型训练一次花费时间仅为 LSTM 模型的 1/8。这是因为 E-Transformer 模型中稠密合成器具有并行表征语义信息的能力,而 LSTM 只能按顺序表征语义信息,并且每个神经元的运行复杂。此外,TD-LSTM 和 AT-LSTM 等融合注意力的模型分别比 LSTM 的时间花费多 23s 和 94s。这表明,在方面情感分析中对方面单独建模、融合注意力机制等方式虽然有效提升了模型的预测准确率,但增大了模型的时间开销。

结束语 针对现有中文在线课程方面情感分析无法准确捕获方面对应细粒度情感问题,本文提出一种面向中文在线课程评论方面情感分析的 E-Transformer 模型,该模型可以并行表征中文在线课程评论方面及其上下文的语义信息,改进了传统方面情感分析模型过度依赖标注数据规模和训练时间开销大的问题。此外,模型通过交互注意力机制对中文在线课程评论方面和上下文语境进行交互建模,重点关注课程评论文本中方面和上下文中的重要部分,生成更加准确的方面和上下文表示。实验结果表明,E-Transformer 模型能够在中文在线课程评论方面情感分析中有较好的准确率,同时与现有模型相比时间开销更小。在未来的研究中,计划融合单词注意力到 E-Transformer 模型中^[28],以期进一步提升模型性能。

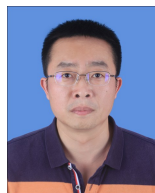
参考文献

- [1] WANG L, HU G, ZHOU T. Semantic analysis of learners' emotional tendencies on online MOOC education [J]. Sustainability, 2018, 10(6): 1921.
- [2] MITEBAIDAL K, DELGADOVERA C, SOLÍS AVILÉS E, et al.

- Sentiment analysis in education domain: A systematic literature review [C]// International Conference on Technologies and Innovation. Springer, 2018: 285-297.
- [3] MORENOMARCOS P M, ALARIOHOYOS C, MUÑOZ-MERINO P J, et al. Sentiment analysis in MOOCs: A case study [C]// Global Engineering Education Conference (EDUCON). IEEE, 2018: 1489-1496.
- [4] SHARMA N, JAIN V. Evaluation and Summarization of Student Feedback Using Sentiment Analysis [C]// International Conference on Advanced Machine Learning Technologies and Applications. Springer, 2020: 385-396.
- [5] HARRIS S C, KUMAR V. Identifying student difficulty in a digital learning environment [C]// International Conference on Advanced Learning Technologies (ICALT). IEEE, 2018: 199-201.
- [6] BUENAÑO FERNÁNDEZ D, VILLEGASCH W, LUJÁN-MORA S. Using text mining to evaluate student interaction in virtual learning environments [C]// World Engineering Education Conference (EDUNINE). IEEE, 2018: 1-6.
- [7] ORAMAS B S R, ZATARAIN C R, BARRÓN E M L, et al. Opinion mining and emotion recognition in an intelligent learning environment [J]. Computer Applications in Engineering Education, 2019, 27(1): 90-101.
- [8] BARRONESTRADA M L, ZATARAINCABADA R, ORAMASBUSTILLOS R. Emotion Recognition for Education using Sentiment Analysis [J]. Research in Computing Science, 2019, 148(5): 71-80.
- [9] NGUYEN P X V, HONG T T T, VAN NGUYEN K, et al. Deep learning versus traditional classifiers on vietnamese students' feedback corpus [C]// NAFOSTED Conference on Information and Computer Science (NICS). IEEE, 2018: 75-80.
- [10] LALATA J P, GERARDO B, MEDINA R. A Sentiment Analysis Model for Faculty Comment Evaluation Using Ensemble Machine Learning Algorithms [C]// International Conference on Big Data Engineering. ACM, 2019: 68-73.
- [11] ONAN A. Sentiment analysis on massive open online course evaluations: A text mining and deep learning approach [J]. Computer Applications in Engineering Education, 2020, 1002(5): 22253.
- [12] LAN Z, CHEN M, GOODMAN S, et al. Albert: A lite bert for self-supervised learning of language representations [J]. arXiv: 1909.11942.
- [13] SOE N, SOE P T. Domain Oriented Aspect Detection for Student Feedback System [C]// International Conference on Advanced Information Technologies (ICAIT). IEEE, 2019: 90-95.
- [14] SHUOQIU Y, CHAOJUN X. Research on Constructing Sentiment Dictionary of Online Course Reviews based on Multi-source Combination [C]// International Conference on Data Science and Information Technology. ACM, 2019: 71-76.
- [15] YUAN X. Emotional tendency of online legal course review texts based on SVM algorithm and network data acquisition [J]. Journal of Intelligent & Fuzzy Systems, 2019, 37(5): 6253-6263.
- [16] KANDHRO I A, WASI S, KUMAR K, et al. Sentiment Analysis of Student's Comment by using Long-Short Term Model [J]. Indian Journal of Science and Technology, 2019, 12(8): 1-16.
- [17] DO H H, PRASAD P W C, MAAG A, et al. Deep learning for aspect-based sentiment analysis: a comparative review [J]. Expert Systems with Applications, 2019, 118(3): 272-299.
- [18] ZHOU J, HUANG J X, CHEN Q, et al. Deep learning for aspect-level sentiment classification: Survey, vision, and challenges [J]. IEEE Access, 2019, 7(5): 78454-78483.
- [19] DONG L, WEI F, TAN C, 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. Association for Computational Linguistics, 2014: 49-54.
- [20] TANG D, QIN B, FENG X, et al. Effective LSTMs for target-dependent sentiment classification [J]. arXiv: 1512.01100.
- [21] LIU Q, ZHANG H, ZENG Y, et al. Content attention model for aspect based sentiment analysis [C]// Proceedings of the World Wide Web Conference. ACM, 2018: 1023-1032.
- [22] WANG Y, HUANG M, ZHU X, et al. Attention-based LSTM for aspect-level sentiment classification [C]// Conference on Empirical Methods in Natural Language Processing. International World Wide Web Conferences Steering Committee, 2016: 606-615.
- [23] MA D, LI S, ZHANG X, et al. Interactive attention networks for aspect-level sentiment classification [J]. arXiv: 1709.00893.
- [24] DEVLIN J, CHANG M W, LEE K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding [J]. arXiv: 1810.104805.
- [25] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]// Advances in Neural Information Processing Systems. Curran Associates Inc. 2017: 5998-6008.
- [26] TAY Y, BAHRI D, METZLER D, et al. Synthesizer: Rethinking Self-Attention in Transformer Models [J]. arXiv: 2005.00743.
- [27] KITAIEV N, KAISER Ł, LEVSKAYA A. Reformer: The efficient transformer [J]. arXiv: 2001.04451v2.
- [28] XU Q, ZHU L, DAI T, et al. Aspect-based sentiment classification with multi-attention network [J]. Neurocomputing, 2020, 388(5): 135-143.



PAN Fang, born in 1982, master, associate professor. Her main research interests include educational big data and online learning behavior analysis.



ZHANG Hui-bing, born in 1976, Ph.D, is a member of China Computer Federation. His main research interests include educational big data and social computing.