



基于语法和语义分割的跨领域方面级情感分类

吴震, 戴新宇*

南京大学计算机软件新技术国家重点实验室, 南京 210023

* 通信作者. E-mail: daixinyu@nju.edu.cn

收稿日期: 2021-05-16; 修回日期: 2022-07-10; 接受日期: 2022-10-14; 网络出版日期: 2023-06-30

国家自然科学基金 (批准号: 62206126, 61936012, 61976114) 资助项目

摘要 神经网络在方面级情感分类任务上已经取得了良好的性能. 然而, 由于复杂且耗时的数据标注流程, 方面级情感分类在很多领域上是低资源甚至是零资源的, 这限制了该任务在实际场景中的应用. 为了解决这个挑战性的问题, 本文关注跨领域的方面级情感分类, 并提出一种基于语法和语义分割的跨领域方面情感分类方法. 具体而言, 针对不同领域用词差异造成的领域漂移和注意力泛化问题, 本文首次提出利用单纯的语法信息来获取可在领域之间迁移的语法注意力, 并引入与目标领域相近的文档情感分类任务来增强神经网络模型对目标领域的情感识别能力, 最终从语法和语义两个层面分别提升模型的注意力机制和文本上下文表示. 实验在 6 个跨领域方面级情感分类任务上进行, 结果表明, 与其他 9 种基线方法相比, 本文的方法在 6 个任务上都取得了最先进的性能, 在平均准确率和平均 macro-F1 两个指标上比之前最好的模型 DIFD 分别提升 7.14% 和 7.6%. 此外, 即使以大规模预训练模型 BERT, BERT-ADA, RoBERTa 等作为骨干网络, 本文的方法仍能实现 3.5% 以上的平均准确率提升和平均 macro-F1 提升.

关键词 方面级情感分类, 跨领域, 神经网络, 注意力, 语法和语义

1 引言

近年来, 随着在线观点文本的快速增长, 情感分析受到了学术界和工业界的广泛关注. 典型的情感分析任务旨在判断一个文档或者句子整体的情感极性 (正面、中性、负面) ^[1,2]. 然而, 在实际场景中, 用户通常在一段文本中评论多个观点目标, 并且对它们表达不同的情感极性. 显然, 文档级或句子级情感分类不能帮助我们分析用户对不同目标的情感态度. 为此, 一个细粒度的情感分析任务——方面级情感分类 (aspect-level sentiment classification, ALSC) ^[2,3] 应运而生. 不同于文档级和句子级情感分类, 方面级情感分类旨在分析文本中特定方面目标 (aspect) 的情感极性. 举个例子, 餐厅评论 “The

引用格式: 吴震, 戴新宇. 基于语法和语义分割的跨领域方面级情感分类. 中国科学: 信息科学, 2023, 53: 1299–1313, doi: 10.1360/SSI-2021-0166
Wu Z, Dai X Y. Separated syntax and semantics modeling for cross-domain aspect-level sentiment classification (in Chinese). Sci Sin Inform, 2023, 53: 1299–1313, doi: 10.1360/SSI-2021-0166

tastes are great, but the **service** is dreadful” 包含两个方面目标, 分别是 “tastes” 和 “service”, 用户对它们的情感分别是正面的 (positive) 和负面的 (negative).

为了解决这个细粒度的情感分析任务, 早先的方法利用语义分析、情感词典等资源作为特征, 并结合机器学习算法进行情感预测^[4~6]. 然而, 基于特征工程的方法需要消耗较多的人力来设计和收集特征. 相比之下, 深度学习方法可以自动地学习低维连续的文本表示, 从而大大节省了人力资源. 在深度学习基础上, 大量工作采用注意力机制 (attention mechanism)^[7] 来捕捉文本中每个方面目标对应的情感表达^[8~10], 并取得良好的效果. 尽管如此, 人们忽视了方面级情感分类任务的一个现实问题——资源稀缺性. 在方面级情感分类任务中, 一段文本可能包含多个方面目标, 为了标注数据, 标注者需要找出所有的方面目标并判断它们在文本中的情感极性, 整个过程复杂耗时, 这导致方面级情感分类任务在很多领域上是零资源的. 显然, 如何利用好已标注的资源并泛化到新领域是非常有意义的. 因此, 本文关注一个现实的细粒度情感分析问题——跨领域场景下的方面级情感分类.

跨领域任务中存在一个共识, 不同领域之间会共享一些知识, 而这些共享的知识可以从源领域 (source domain) 迁移到目标领域 (target domain), 从而帮助目标任务. 沿着这个思路, 最近的跨领域方面级情感分类方法 DIFD (domain-invariant feature distillation)^[11] 以双向长短期记忆网络 (bidirectional long short-term memory, BiLSTM)^[12, 13] 作为编码器, 在注意力机制的基础上引入对抗学习^[14] 和正交的辅助任务来蒸馏可迁移的特征. 尽管如此, DIFD 模型中的注意力仍然是基于单词的上下文语义计算出来的, 而单词的使用和分布与领域密切相关, 这些分布差异会造成领域漂移^[15], 从而影响注意力机制在跨领域方面情感分类任务上的准确性和泛化性. 举个例子, “delicious” 这个词是餐厅 (restaurant) 领域的常用词, 但基本不会在出现笔记本电脑 (laptop) 评论中, 因此在笔记本电脑领域训练的注意力模型迁移到餐厅领域时可能很难关注到 “delicious”, 从而造成错误的情感分类.

为此, 本文首次提出一种基于语法的方面级注意力方法. 不同于以往基于单词语义的注意力模型, 基于语法的注意力完全利用语法信息 (词性信息、依存信息) 和方面目标的位置信息来获取注意力权重, 这些语法信息和语言本身的特性相关, 和领域没有明显的相关性, 因此具备更好的领域迁移性. 举个例子, 表 1^[16] 展示了一个跨领域的方面级情感分类样例, 源领域句子来自笔记本电脑评论, 目标领域句子来自餐厅评论, 单词后的中括号包含该单词对应的词性标签, 加粗的单词表示方面目标. 可以看到, 虽然这两个句子来自不同的领域并且包含不同的方面目标和观点词, 尤其是餐厅句子中的单词 “tasty” 基本不会用于笔记本电脑评论, 但它们拥有完全相同的词性序列, 因此在源领域句子上训练的语法注意力完全可以迁移到目标领域的句子上. 为了进一步验证语法信息和领域没有明显的相关性, 本文统计了 SemEval2014 任务 4^[3] 餐厅评论数据集和笔记本电脑评论数据集的二元词性标签, 表 2 列出了两个数据集最频繁的 20 个二元词性标签. 可以看到, 虽然餐厅领域和笔记本电脑领域差异性很大, 但前 20 的二元词性标签中有 17 个是相同的. 实际上, 当统计两个数据集最频繁的 100 个二元词性标签时, 我们发现 93 个是相同的, 这说明词性模式在同一种语言的不同领域之间具有较好的迁移性¹⁾.

虽然基于语法的注意力在跨领域场景下具有较强的鲁棒性, 但还遗留了一个挑战需要解决. 在表 1 的样例中, 即使语法注意力关注到了目标领域方面目标 “pizza” 对应的观点词 “tasty”, 但由于 “tasty” 基本不在源领域的语料中出现, 模型可能无法学会它的情感语义, 因此无法做出准确的情感判断. 为此, 本文引入与目标领域相近的文档级情感分类 (document-level sentiment classification, DSC)

1) 在不同语言之间, 词性模式的语法不变性不一定成立, 例如中文主谓宾结构的句子对应到日文中通常是主宾谓结构. 因此, 本文关注同一种语言内的跨领域方面级情感分类.

表 1 跨领域方面级情感分类样例^{a)}Table 1 Examples of cross-domain aspect-level sentiment classification^{a)}

Sentence	Given aspects	Sentiment
Source domain: the[DET] cpu [NOUN] is[VERB] very[ADV] warm[ADJ]	cpu	Negative
Target domain: the[DET] pizza [NOUN] is[VERB] quite[ADV] tasty[ADJ]	pizza	Positive

a) 单词后中括号为该单词对应的词性标签, 词性标签来自于 Biaffine Parser^[16].

表 2 SemEval 2014 任务 4 中两个数据集最频繁的 20 个二元词性标签^{a)}Table 2 Top 20 2-gram POS tags in the datasets of SemEval 2014 task 4^{a)}

Domain	Top20 2-gram POS tags
restaurant	PRON VERB, NOUN PUNCT, ADJ NOUN, NOUN NOUN, ADP NOUN, VERB DET, NOUN AUX, VERB ADP, NOUN CCONJ, NOUN ADP, PRON AUX, ADP DET, AUX ADJ, DET NOUN, ADV ADJ, DET ADJ, PART VERB, AUX DET, AUX ADV, ADJ PUNCT
laptop	PRON VERB, NOUN PUNCT, ADJ NOUN, NOUN NOUN, ADP NOUN, VERB DET, NOUN AUX, VERB ADP, NOUN CCONJ, NOUN ADP, PRON AUX, ADP DET, AUX ADJ, DET NOUN, ADV ADJ, DET ADJ, PART VERB, AUX VERB, VERB PRON, AUX PART

a) 两个数据集最频繁的 20 个二元词性标签中有 17 个是相同的.

作为辅助任务²⁾, 提升模型对目标领域的情感识别能力. 相比于稀缺的方面级情感分类数据, 互联网评论网站中包含大量带标签的文档级情感分类数据, 并且易于收集^[17,18].

最终, 本文提出一种显式的基于语法和语义分割 (separated syntax and semantics modeling, S3M) 的跨领域方面情感分类方法, 它能从语法和语义层面分别提升跨领域方面情感分类任务的性能. 在语法层面上, S3M 首次完全利用语法信息来获得方面目标的注意力权重, 从而提升注意力机制在跨领域方面情感分类任务中的鲁棒性. 在语义层面上, S3M 引入资源丰富的文档级情感分类来提升模型的文本上下文表示, 从而增强模型对目标领域情感语义的识别能力. 为了验证 S3M 方法的性能, 本文在 3 个数据集, 共 6 个跨领域方面级情感分类任务上进行了实验. 结果表明, S3M 显著优于之前最好的方法 DIFD. 即使以预训练语言模型 BERT (bidirectional encoder representations from transformers)^[19], BERT-ADA (BERT adaptation)^[20] 和 RoBERTa (robustly optimized BERT pretraining approach)^[21] 等作为编码器, S3M 方法仍能实现显著的提升, 进一步的分析也验证了 S3M 方法的有效性.

2 相关工作

2.1 方面级情感分类

方面级情感分类任务旨在判断文本中特定方面目标的情感极性^[3]. 早先的工作采用特征工程的方式, 以情感词典、词袋模型等特征作为支撑向量机 (support vector machine, SVM) 的输入来进行情感分类^[4]. 虽然取得了较好的结果, 但这类方法的性能严重依赖特征采集的质量, 并且需要消耗大量的人力资源. 为此, 最近的工作将深度学习应用到方面级情感分类任务中, 自动地学习文本中的特征表示. 在这些工作中, 大量方法采用注意力机制来捕获方面目标对应的观点表达, 然后聚合它们来进行情感分类. 例如, Wang 等^[8] 首次将注意力机制引入到方面级情感分类任务中. 在此基础上, Ma 等^[9] 提出一个交互式的注意力网络来交互地学习方面目标和文本的表示. Fan 等^[10] 则将粗粒度的注意力机制和细粒度的注意力机制融合在一起, 最终提升情感分类的性能. 此外, 也有一些工作采用多次注

2) 文档级情感分类数据来自于与目标领域相近的领域.

注意力过程来更充分地挖掘文本中和方面目标相关的情感表达^[22, 23]. 虽然这些注意力模型在单领域方面级情感分类任务上取得了良好的效果, 但它们没有考虑跨领域的场景, 因此通常不具备较好的泛化性能.

2.2 迁移学习

迁移学习 (transfer learning) 旨在将源任务的知识或模型迁移到目标任务中, 从而提升目标任务的性能^[24]. 当源任务和目标任务属于同类任务但两者的领域不同时, 这时候的迁移学习称之为跨领域迁移学习, 本文关注的跨领域方面级情感分类就属于此类任务. 在情感分析研究中, 早期的跨领域工作主要关注粗粒度的情感分析任务, 如句子级、文档级情感迁移. Pan 等^[25] 提出谱特征对齐算法 (spectral feature alignment, SFA) 来建模领域特定词和领域无关词之间的共现关系, 从而减小跨领域情感分析时的领域差异. Glorot 等^[26] 在神经网络基础, 提出用无监督学习自动抽取评论文本中可迁移的情感表示. 随后, 梯度反转层 (gradient reversal layer, GRL)^[27]、对抗学习 (adversarial learning)^[28]、辅助任务^[29] 等技术也被用来抽取领域无关的特征. 和粗粒度的跨领域情感分析任务相比, 跨领域方面级情感分类任务更具挑战性, 因为它的数据获取更加困难, 需要分析的情感粒度也更加细致. 为此, Hu 等^[11] 提出 DIFD 模型, 在注意力机制的基础上, 额外地添加对抗学习和正交的方面目标抽取任务来蒸馏领域无关的特征, 从而进行跨领域的方面级情感分类. 遗憾的是, DIFD 中的注意力仍然基于单词的语义表示计算得出, 而单词在不同领域中分布差异会造成领域漂移, 从而影响注意力机制在跨领域任务中的准确性和鲁棒性. 除了针对性的跨领域情感分析方法外, 大规模预训练语言模型 BERT^[19] 也常被用于跨领域任务中. 在 BERT 的基础上, Rietzler 等^[20] 提出 BERT-ADA 模型, 他们先在大量目标领域文本上微调语言模型, 然后再去微调下游任务, 在跨领域情感分类上取得了优异的性能. RoBERTa^[21] 去掉了 BERT 中的下一句预测任务, 使用更大的批处理量和语料来训练语言模型, 从而获得更强大的文本上下文表示能力.

相比之下, 本文提出的语义和语法分割方法 (S3M) 将注意力机制和情感语义分开建模, 首次利用纯粹的语法信息 (词性信息、依存信息) 和方面目标的位置信息来获得更鲁棒的注意力权重, 然后利用容易收集的文档级情感分类数据作为辅助任务, 增强模型对目标领域的情感识别能力, 从语法和语义层面分别提升跨领域方面情感分类的性能.

3 基于语法和语义分割的跨领域方面情感分类方法

3.1 问题定义

假定源领域有 N_s 条带情感标注的样本 $D_s = \{(x_s^k, a_s^k), y_s^k\}_{k=1}^{N_s}$, 目标领域有 N_t 条无标注的样本 $D_t = \{(x_t^k, a_t^k)\}_{k=1}^{N_t}$, 这里 x 表示文本, a 是文本 x 中的方面目标, y 表示方面目标 a 在文本中 x 的情感极性 (正面、负面、中性). 跨领域的方面级情感分类任务 (cross-domain aspect-level sentiment classification) 旨在源领域的标注数据集 D_s 上学习一个映射函数 $\mathcal{F}: \{(x_s, a_s)\} \rightarrow \{y_s\}$, 然后将这个映射函数 $\mathcal{F}$ 迁移到目标领域, 来预测目标领域中无标注样本 (x_t, a_t) 对应的情感标签 y_t .

3.2 方法总览

图 1 展示了本文提出的基于语法和语义分割的跨领域方面情感分类方法 (S3M) 的整体架构, 它主要包含两个组件: 基于语法注意力的方面级情感分类模块和基于文档级情感分类的语义增强模块. 在 S3M 方法中, 这两个组件被显式分隔开, 分别从语法和语义层面来提升跨领域方面情感分类任务中

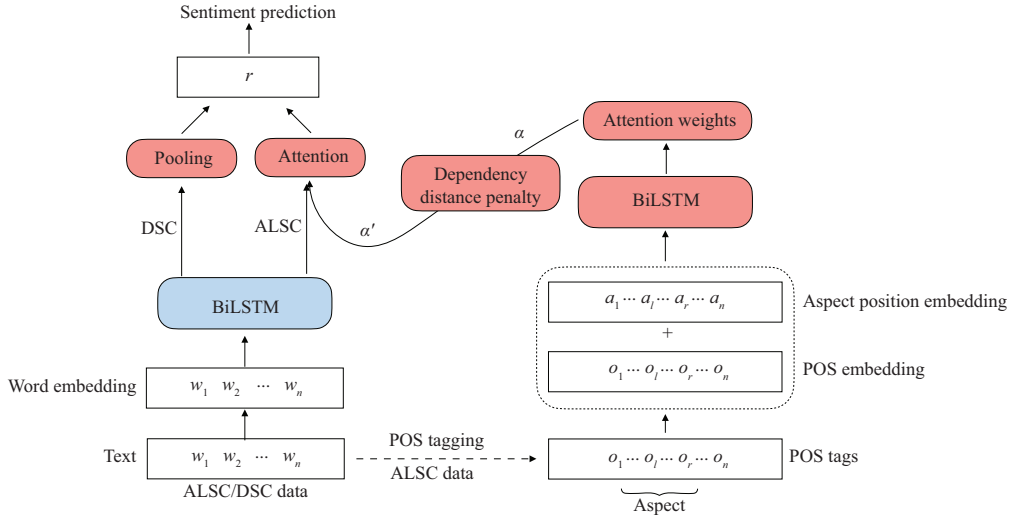


图 1 (网络版彩图) 基于语法和语义分割的跨领域方面情感分类模型 S3M 总体框架图

Figure 1 (Color online) The overall architecture of S3M for cross-domain aspect-level sentiment classification

的注意力和文本上下文表示. 形式上看, 对于注意力过程 $r = \alpha H$ (α 为方面目标对文本中每个词的注意力权重, H 为单词对应的上下文表示, r 用来判断方面目标的情感极性), 不同于以往基于语义注意力的模型, S3M 利用单纯的语法信息 (词性信息、依存信息) 和方面目标的位置信息来获得注意力权重 α , 从而确保模型的注意力能在不同的领域之间迁移. 为了提升模型对目标领域的情感识别能力, S3M 引入与目标领域接近的文档级情感分类数据作为辅助任务, 使得模型的文本上下文表示 H 和目标领域的语义更加接近.

3.3 基于语法注意力的方面级情感分类

3.3.1 文本编码

给定一个包含 n 个词的输入文本 $x = \{w_1, w_2, \dots, w_n\}$, 本文尝试两种不同的编码方式来获得输入文本 x 的上下文表示 $H_w \in \mathbb{R}^{n \times d_h}$, 从而验证 S3M 方法对不同编码器的兼容性和有效性, d_h 为文本上下文表示的维度:

(1) 词向量 + 双向长短期记忆网络 (BiLSTM). S3M 首先将其映射成一个词向量序列 $E_w = \{w_1, w_2, \dots, w_n\} \in \mathbb{R}^{n \times d_w}$, d_w 为词向量的维度, 然后利用一个 BiLSTM 网络来编码 E_w , 从而获得文本的上下文表示 $H_w = \{h_1^w, w_2^w, \dots, w_n^w\}$.

(2) 预训练语言模型. 在这个模式下, S3M 直接使用预训练语言模型编码输入文本 x , 进而生成 x 对应的文本上下文表示 $H_w = \{h_1^w, w_2^w, \dots, w_n^w\}$.

3.3.2 基于语法信息的注意力机制

面向方面目标的词性编码. 对于文本 $x = \{w_1, w_2, \dots, w_n\}$ 和文本中的方面目标 $a = \{w_l, \dots, w_r\}$, l 和 r 分别表示方面目标 a 在文本 x 中的开始位置和结束位置, S3M 使用 Biaffine Parser^[16] 来获得 x 中每个词的词性标签 $\{o_1, o_2, \dots, o_n\}$, 然后将词性标签映射为对应的词性向量序列 $E_o = \{o_1, o_2, \dots, o_n\} \in \mathbb{R}^{n \times d_o}$, d_o 为词性向量的维度. 为了获得面向方面目标的注意力权重, S3M 将方面目标的位置信息融入到词性编码中. 具体而言, 对于文本中每个单词 w_i , 如果 w_i 是方面目标词则用 1 表示, 否则则用 0

表示, 这样输入样本会被编码成一个能表示方面目标位置的序列 $\{0, 1\}^n$, 然后 S3M 将这个 $\{0, 1\}^n$ 序列映射为编码方面目标的位置向量 $\mathbf{E}_a = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\} \in \mathbb{R}^{n \times d_a}$, d_a 为位置向量的维度. 举个例子, 对于样本 “The **hot** **dog** is very tasty” (粗体表示方面目标), 单词的位置编码为 $\{0, 1, 1, 0, 0, 0\}$, 对应的位置向量为 $\{\mathbf{0}, \mathbf{1}, \mathbf{1}, \mathbf{0}, \mathbf{0}, \mathbf{0}\}$ (粗体 $\mathbf{1}$ 和 $\mathbf{0}$ 分别表示 1 和 0 对应的向量编码, 两个位置向量编码随机初始化并随着模型一起训练). 在获取能够表示方面目标的位置向量后, S3M 将词性向量 $\mathbf{E}_o$ 和位置向量 $\mathbf{E}_a$ 融合在一起, 从而获得融合方面目标位置的词性向量: $\hat{\mathbf{E}}_o = \mathbf{E}_o + \mathbf{E}_a$. 同样地, S3M 使用另一个 BiLSTM 网络³⁾编码 $\hat{\mathbf{E}}_o$ 来生成词性的上下文表示 $\mathbf{H}_o = \{\mathbf{h}_1^o, \mathbf{w}_2^o, \dots, \mathbf{w}_n^o\}$.

基于词性的注意力权重. 已有的工作通常基于单词的上下文表示来计算注意力权重, 而不同领域的用词分布差异容易造成领域漂移, 这会影响语义注意力在跨领域场景下的泛化性. 相比之下, S3M 模型的注意力完全来自于文本单词的词性特征 (part-of-speech, POS) 和方面目标的位置信息, 与领域信息几乎无关, 因此能较好地由源领域泛化到目标领域上. 其注意力权重 α 计算方式如下:

$$\alpha = \text{softmax}(\mathbf{W}_{o2} \tanh(\mathbf{W}_{o1} \mathbf{H}_o + \mathbf{b}_o)). \quad (1)$$

依存距离惩罚. 一些研究表明, 距离方面目标越近的词越有可能是方面目标对应的观点词^[30, 31], 这些观点词对判断方面目标的情感极性至关重要. 为了提升语法注意力权重 α 对方面目标的敏感性, S3M 进一步引入依存距离信息^[32] 对注意力 α 进行惩罚 (dependency distance penalty, DDP):

$$c_i = \begin{cases} 1 & \text{if } l \leq i \leq r, \\ \frac{1}{d_i} & \text{otherwise,} \end{cases} \quad (2)$$

$$\alpha'_i = \frac{c_i \times \alpha_i}{\sum_{j=1}^n c_j \times \alpha_j}, \quad (3)$$

上式中 d_i 表示文本中第 i 个词 w_i 和方面目标 a 之间的依存距离. 可以看到, 在依存距离上离方面目标越近的词具有更大的保留权重 c_i , 距离越远的词权重越小, 越有可能被忽视.

3.3.3 方面级情感分类损失函数

进行依存距离惩罚之后, S3M 用优化后的语法注意力权重 α' 来聚合文本的上下文表示 $\mathbf{H}_w$, 从而生成面向方面目标的情感分类表示: $\mathbf{r} = \sum_{i=1}^n \alpha'_i \mathbf{h}_i^w$. 最终, 方面目标对应的情感预测概率 $\hat{\mathbf{y}}$ 和交叉熵损失 $\mathcal{L}_a$ 如下式所示:

$$\hat{\mathbf{y}} = \text{softmax}(\mathbf{W}_p \mathbf{r} + \mathbf{b}_p), \quad (4)$$

$$\mathcal{L}_a = -\frac{1}{N_s} \sum_{k=1}^{N_s} y_s^k \log \hat{y}_s^k. \quad (5)$$

3.4 基于文档级情感分类的语义增强模块

虽然语法注意力能够泛化到目标领域上, 但在上述过程中, 文本上下文表示 $\mathbf{H}_w$ 仍然是在源领域语料上训练的, 这会使 $\mathbf{H}_w$ 对目标领域某些特有情感词的语义不敏感, 如餐厅评论中的 “tasty”, “delicious” 基本不会出现在笔记本电脑评论中. 此外, 一些词在不同领域中可能表达不同的情感语义,

3) 对于文本而言, 不同单词可能会映射到相同的词性, 这个过程不仅降低了用词差异导致的单词序列模式稀疏性, 也使得词性序列具有更强、更通用的序列信息. 例如, 词性序列 “NOUN VERB ADJ” 通常构成系表结构, 用于表达名词主语的特征、状态等.

如餐厅评论 “the service was warm” 和笔记本电脑评论 “the cpu was very warm” 中的 “warm” 就表达不同的情感极性, 因此直接将文本上下文编码器迁移到目标领域上可能会伤害目标任务的性能. 为此, S3M 引入与目标领域接近的文档级情感分类数据作为辅助任务, 在语义层面上增强文本上下文编码器, 使其适合目标领域的情感语义.

具体而言, 对于辅助文档 $x^d = \{w_1^d, w_2^d, \dots, w_n^d\}$, S3M 使用 3.3.1 小节中的文本编码器来编码文档, 从而获得上下文表示 $\mathbf{H}_w^d = \{\mathbf{h}_{w_1}^d, \mathbf{h}_{w_2}^d, \dots, \mathbf{h}_{w_n}^d\} \in \mathbb{R}^{n \times d_h}$, 最终使用平均池化来获得文档表示: $\mathbf{r}_d = \text{Pooling}(\mathbf{H}_w^d)$. 文档级情感分类任务的交叉熵损失如下式所示 (N_d 表示文档级情感分类数据样本的数量):

$$\hat{y}_d = \text{softmax}(\mathbf{W}_d \mathbf{r}_d + \mathbf{b}_d), \quad (6)$$

$$\mathcal{L}_d = -\frac{1}{N_d} \sum_{k=1}^{N_d} y_d^k \log \hat{y}_d^k. \quad (7)$$

3.5 模型训练

最终, S3M 采用多任务学习的方式, 联合优化方面级情感分类任务和辅助的文档级情感分类任务, 模型最终的优化目标如下式所示:

$$\mathcal{L} = \mathcal{L}_a + \lambda \mathcal{L}_d, \quad (8)$$

λ 是一个超参数, 用来度量辅助任务文档级情感分类的重要程度.

4 实验与分析

4.1 实验设置

数据集设置. 参考 Hu 等^[11] 的工作, 本文使用了 3 个方面级情感分类数据集: Restaurant (R), Laptop (L) 和 Twitter (T), 共 6 个跨领域方面级情感分类任务: R→L, L→R, R→T, T→R, L→T, T→L, 箭头的左边和右边分别表示源领域和目标领域. Restaurant (R) 和 Laptop (L) 数据集来自 SemEval 2014 任务 4^[3], Twitter (T) 数据集来自 Dong 等^[33] 的工作. 在每个跨领域任务中, 源领域的训练集和测试集分别作为跨领域任务的训练集和验证集, 目标领域的训练集和测试集被组合在一起作为任务的测试集, 因此本文的任务属于目标领域零资源的跨领域情感分析任务. 此外, 本文使用 3 个文档级情感分类数据: Yelp, Amazon, Twitter, 它们来自于 Chen 等^[34] 的工作, 分别用来增强跨领域方面情感分类任务中目标领域 Restaurant (R), Laptop (L) 和 Twitter (T) 的情感语义. 因此, 跨领域方面情感分类任务 R→L 和 T→L 使用的辅助数据集是 Amazon 电商评论数据, L→R 和 T→R 的辅助数据集是 Yelp 商户评论数据, R→T 和 L→T 使用的辅助数据集是 Twitter 情感分类数据. 表 3 展示了不同数据集的统计信息.

模型设置. 本文采用 300 维预训练的 GloVe 向量^[35] 来初始化 S3M 模型的词向量. 词性向量和位置向量由正态分布 $N(0, 1)$ 随机初始化, 维度为 50 维. LSTM 单元的隐层维度也设为 50, 因此 BiLSTM 网络的输出维度为 100. 为了避免过拟合, 模型的词嵌入层和最终输出层添加了 dropout 技术^[36], 随机 dropout 的概率为 0.5. 优化器采用 Adam^[37], 学习率为 $1e-3$. 当使用预训练语言模型作为 S3M 方法的文本编码模块时 (参见 3.3.1 小节), 预训练模型部分的参数学习率设为 $1e-5$, 输出层的 dropout 概率为 0.1, 其他部分的参数设置保持不变. 为了获得统计上可靠的结果, 每组实验重复 5 次并报告 5 次实验的平均结果.

表 3 数据集样本分布
Table 3 Statistics of datasets

Tasks	Datasets		Positive	Neutral	Negative	Total
ALSC	Restaurant (R)	Train	2164	633	805	3502
		Test	728	196	196	1120
	Laptop (L)	Train	987	460	866	2313
		Test	341	169	128	638
	Twitter (T)	Train	1561	3127	1560	6248
		Test	173	346	173	692
DSC	Yelp	Train	10k	10k	10k	30k
	Amazon	Train	10k	10k	10k	30k
	Twitter	Train	10k	10k	10k	30k

评价指标. 参考 Hu 等^[11]的工作, 本文采用标准的分类准确率 (accuracy, Acc.) 和 macro-F1 作为模型的评价指标, 指标越高, 表示分类性能越好.

4.2 基线系统

为了验证 S3M 方法在跨领域方面级情感分类任务上的有效性, 本文和两类方法进行了对比. 第 1 类是非迁移的方法 (non-transfer methods), 即经典的单领域方面级情感分类方法; 第 2 类是迁移方法 (transfer methods), 可以进行较好的跨领域方面级情感分类.

非迁移方法. ATAE (attention with aspect embedding)^[8] 将词向量和方面目标向量拼接作为 LSTM 网络的输入, 然后采用注意力机制生成方面目标对应的情感表示. IAN (interactive attention networks)^[9] 提出交互式的注意力机制分别学习上下文表示和方面目标表示, 然后拼接它们进行情感分类. MemNet (memory network)^[22] 在记忆网络基础上, 使用多次注意力机制来聚合文本的上下文表示. RAM (recurrent attention memory)^[23] 在 GRU (gated recurrent unit) 单元中进行多次注意力机制, 从而非线性地得到方面目标的表示. GCAE (gated convolutional network with aspect embedding)^[38] 是首个基于卷积神经网络的方面级情感分类模型. BiLSTM-ATT (BiLSTM-attention)⁴⁾ 是本文的基线方法, 与 ATAE 类似, 同样采用基于单词上下文的注意力机制来获取方面目标对应的情感表示.

迁移方法. IATN (interactive attention transfer network)^[39] 将方面目标信息和句子表示融合, 从而进行跨领域的情感分类. DIFD^[11] 是非预训练方法中最好的跨领域方面级情感分类方法, 它将方面目标检测作为辅助任务来蒸馏领域不变的特征. 不同于本文方法, DIFD 训练时需要目标领域的方面目标标注数据. DIFD(S) 是 DIFD 的一个简化, 只在源领域数据上进行训练. BERT^[19] 是一个在大规模语料上预训练的语言模型, 通过微调的方式来适配下游任务, 在很多任务上都取得了良好的性能. BERT-ADA^[20] 是一个专为领域迁移开发的 BERT 模型. 它先在大量目标领域的文本上微调语言模型, 使得 BERT 能够很好地建模目标领域的上下文语义, 然后微调下游任务进行领域迁移. 在 BERT 的基础上, RoBERTa^[21] 使用更大的批处理量和语料来训练语言模型, 具有更强大的文本上下文编码能力. 本文中的预训练模型均为 base 版本.

4) 为了比较公平, BiLSTM-ATT 中隐层单元数设为 200, 和 S3M 模型使用同等规模的参数.

表 4 模型性能的总体对比 (未使用预训练语言模型的方法)^{a)}Table 4 Performance comparison of different methods without pre-trained language model^{a)}

Model	R→L		L→R		R→T		T→R		L→T		T→L		Average	
	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1
Non-transfer methods														
ATAE-LSTM	56.56	47.71	63.21	46.01	34.97	33.82	50.78	44.06	40.43	39.82	42.39	40.73	48.05	42.02
IAN	60.39	50.69	66.50	47.97	35.09	33.02	51.14	44.46	44.16	44.34	43.88	40.16	50.19	43.44
MemNet	57.17	47.79	65.97	49.28	34.47	32.37	50.36	42.68	44.94	44.37	36.36	34.96	48.21	41.90
RAM	58.32	43.98	54.68	26.81	33.18	29.96	48.18	43.23	44.87	44.06	42.73	42.51	46.99	38.42
GCAE	61.74	50.39	66.60	49.07	32.87	29.34	45.62	37.52	44.93	45.18	47.98	42.39	49.95	42.31
BiLSTM-ATT	62.32	52.52	69.44	52.41	35.33	34.00	47.52	40.21	43.41	43.23	47.68	46.90	50.95	44.89
Transfer methods														
IATN	62.05	46.37	67.13	51.64	34.14	30.55	40.13	37.79	44.80	44.69	46.12	42.91	49.06	42.32
DIFD(S)	63.81	57.74	68.30	53.66	36.89	33.74	55.63	44.10	45.73	46.17	44.53	43.76	52.48	46.52
DIFD	64.86	60.51	68.53	57.31	40.13	38.85	57.60	46.59	47.32	47.31	48.97	47.56	54.56	49.68
BiLSTM-POSATT	66.07	60.90	71.47	57.91	40.28	39.47	61.18	51.46	47.05	47.35	50.77	49.09	56.14	51.03
S3M	67.32	61.33	73.09	62.53	51.34	51.42	66.17	59.35	52.36	52.71	59.97	56.36	61.70	57.28

a) POSATT 表示基于词性上下文的注意力机制. S3M 显著优于之前最好的模型 DIFD ($p < 0.01$), S3M 模型的参数量 (156k) 也少于 DIFD 模型的参数量 (187k).

4.3 主要实验结果

表 4 展示了不同方法在 6 个跨领域方面级情感分类任务上的结果, 加粗的数值表示每个任务上最好的性能. 可以看到:

- 在非迁移的方法中, 简单的注意力模型 IAN, BiLSTM-ATT 在不同任务上取得了不错的结果. 相比之下, 反而是在单领域上性能较好的 RAM 模型在跨领域场景下取得了最差的性能. 可能的原因是 RAM 采用多次注意力机制并用 GRU 单元非线性地聚合信息, 有更强的数据拟合能力, 在跨领域场景下过分拟合源领域的标注数据, 从而在目标领域上表现出较差的泛化性能.

- 和非迁移的方法相比, 除了 IATN, 基于迁移的方法总体上取得了更好的性能. 相比于跨领域的基线方法 IATN, 只采用源领域标注数据的模型 DIFD(S) 已经取得了明显的提升, 这验证了辅助的方面抽取任务对跨领域方面情感分类的有效性. 在 DIFD(S) 的基础上, DIFD 进一步引入目标领域的数据进行对抗训练, 来蒸馏可迁移的知识, 从而成为了之前跨领域方面级情感分类任务中性能最好的模型.

- BiLSTM-POSATT 模型是本文 S3M 模型的一个变种, 它不采用依存距离惩罚和文档级情感分类作为辅助任务, 只是将基线模型 BiLSTM-ATT 中基于单词上下文的语义注意力机制替换成基于词性上下文的注意力机制. 结果表明 BiLSTM-POSATT 在 BiLSTM-ATT 基础上取得了平均 5.19% 的准确率提升和 6.14% 的 macro-F1 提升, 甚至超过之前最好的模型 DIFD, 这证明了基于语法的注意力机制在跨领域场景下的有效性. 在添加依存距离惩罚和文档级数据进行语义增强之后, 本文的方法 S3M 获得了进一步的性能提升, 在平均准确率和 macro-F1 上比 DIFD 分别提升了 7.14% 和 7.6%.

表 5 展示了基于预训练语言模型的方法在跨领域方面级情感分类任务上的性能, 可以看到:

- BERT 以 Transformer^[40] 作为编码器, 在大规模语料上进行预训练, 具有很强的上下文感知能

表 5 基于预训练语言模型方法的性能对比^{a)}

Table 5 Performance comparison of different methods using pre-trained language models^{a)}

Model	R→L		L→R		R→T		T→R		L→T		T→L		Average	
	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1
BERT	73.87	67.46	72.51	61.70	49.47	50.49	60.21	51.13	47.92	47.84	49.88	48.93	58.97	54.59
BERT-POSATT	74.22	67.12	75.98	67.43	50.31	50.76	60.80	54.70	53.56	53.95	57.68	55.68	62.09	58.27
BERT-S3M	74.70	68.81	77.62	69.47	52.32	53.00	61.69	57.51	52.21	52.35	60.52	58.22	63.18	59.89
BERT-ADA	75.84	70.54	75.35	64.39	50.97	51.38	65.37	54.67	50.36	50.76	55.01	52.70	62.14	57.40
BERT-ADA-POSATT	78.85	75.73	78.01	70.84	51.97	52.25	70.80	63.72	51.82	51.89	66.57	62.61	66.34	62.69
BERT-ADA-S3M	79.50	75.07	80.20	72.78	54.72	55.18	71.86	63.19	53.97	54.27	67.29	63.13	67.92	63.93
RoBERTa	76.55	71.00	79.52	70.59	53.05	52.57	65.46	60.47	51.83	51.32	60.89	58.85	64.55	60.80
RoBERTa-POSATT	78.37	73.88	80.28	71.50	54.38	53.52	70.16	64.02	52.47	52.83	66.05	62.69	66.95	63.07
RoBERTa-S3M	78.21	75.29	82.47	73.75	55.66	56.09	71.74	65.11	54.48	54.70	66.16	63.60	68.12	64.75

a) POSATT 表示基于词性上下文的注意力机制. BERT-S3M, BERT-ADA-S3M, RoBERTa-S3M 显著优于对应的基础模型 ($p < 0.01$).

力, 因此在跨领域方面级情感分类任务上也展示出良好的领域迁移性. 在 BERT 的基础上, BERT-ADA 进一步在大量目标领域的语料上微调语言模型来建模目标领域的语义, 进而完成领域迁移. 结果表明, 在大量目标领域数据的语义增强下, BERT-ADA 的平均准确率和 macro-F1 比 BERT 分别提升了 3.17% 和 2.89%. 由于使用更大规模的预训练语料和批处理量, RoBERTa 在所有预训练语言模型中取得了最佳的迁移性能.

• 当使用强劲的预训练语言模型作为骨干网络后, 基于语法的注意力机制 POSATT 仍然带来了显著的性能提升. 在引入文档级情感分类数据进行语义增强之后, 模型的性能得到了进一步提升. 最终, S3M 方法在预训练语言模型 BERT 和 BERT-ADA 的基础上实现了 4% 以上的平均准确率提升和超过 5% 的 macro-F1 提升. 即使以强大的 RoBERTa 作为编码器, S3M 仍实现 3.5% 以上的平均准确率提升和 macro-F1 提升. 有力地证明了基于语法的注意力机制和基于文档情感分类的语义增强策略在跨领域方面级情感分类任务上的有效性.

4.4 消融实验

为了验证不同组件对最终模型 S3M 的贡献, 本文进行了消融实验, 结果如表 6 所示. 模型 S3M w/o POSATT, w/o DDP, w/o DSC 分别表示从 S3M 模型中移除基于词性上下文的语法注意力 (替换为基于单词上下文的语义注意力)、依存距离惩罚 DDP 或者文档级情感分类任务 DSC.

可以看到, 从 S3M 模型中移除任意一个组件, 模型性能都出现了一定程度的下降, 这验证了词性注意力机制、依存距离惩罚和辅助的文档级情感分类任务对跨领域方面级情感分类的有效性. 在 3 个组件中, 移除文档级情感分类任务后, 模型性能出现了最明显的下降, 这表明有效建模目标领域的情感语义是非常重要的. 相比之下, 依存距离惩罚带来的提升最小, 但它仍然能够帮助调整注意力机制使其具有更强的鲁棒性, 最终为跨领域方面级情感分类作出贡献.

4.5 权重参数 λ 的影响

式 (8) 中的权重参数 λ 反应了辅助的文档级情感分类任务对目标领域的语义增强程度. 本文通过改变参数 λ 的数值来探究 S3M 方法中基于文档级情感分类的语义增强模块对跨领域方面级情感分类

表 6 不同组件对实验结果的影响^{a)}
Table 6 Effects of different components on results^{a)}

Model	R→L		L→R		R→T		T→R		L→T		T→L		Average	
	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1
S3M	67.32	61.33	73.09	62.53	51.34	51.42	66.17	59.35	52.36	52.71	59.97	56.36	61.70	57.28
w/o POSATT	65.93	58.81	71.81	60.79	48.01	47.43	64.50	57.30	47.61	46.98	57.26	55.17	59.18	54.41
w/o DDP	66.11	59.28	72.46	61.70	50.91	50.92	63.16	57.53	52.08	52.40	57.83	55.44	60.42	56.20
w/o DOC	66.22	60.51	71.35	56.56	40.45	39.68	62.19	52.49	47.39	47.65	52.60	50.74	56.69	51.27

a) POSATT, DDP, DSC 分别表示基于词性上下文的注意力机制、依存距离惩罚和文档级情感分类任务。

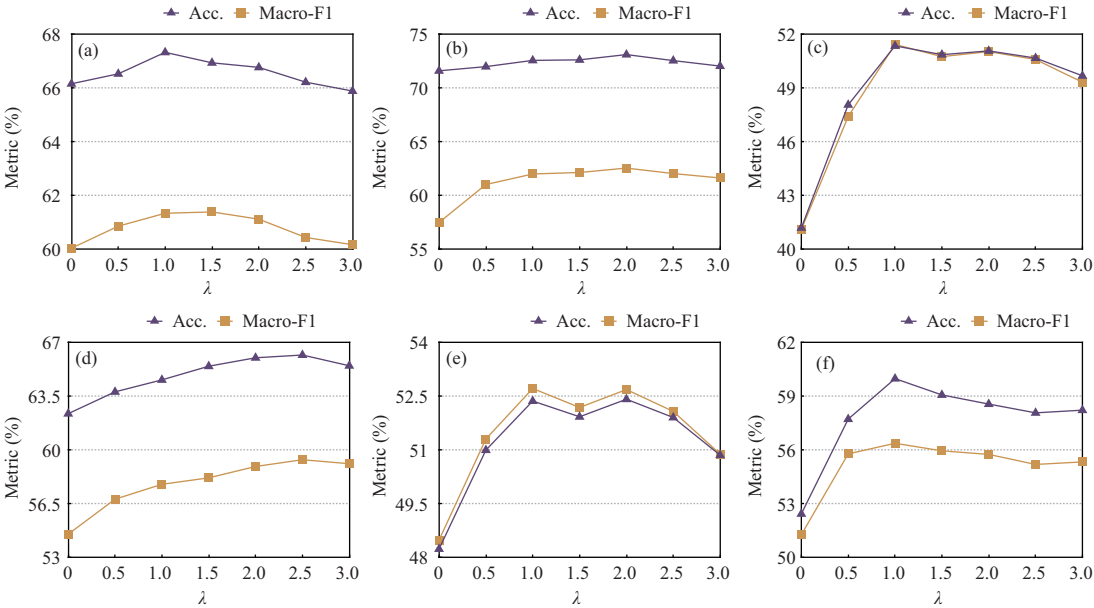


图 2 (网络版彩图) 不同的权重参数 λ 对 S3M 性能的影响

Figure 2 (Color online) The effects of different λ on S3M performance. (a) R→L; (b) L→R; (c) R→T; (d) T→R; (e) L→T; (f) T→L

的影响, λ 的取值范围为 $[0, 3]$, 取值间隔为 0.5. 图 2 展示了 S3M 在取不同 λ 时 6 个跨领域方面级情感分类任务的性能, 横轴为 λ 的取值, 纵轴为准确率指标和 macro-F1 指标.

随着参数 λ 的增加, S3M 模型的性能总体上展示出先上升后下降的趋势. 在上升阶段, λ 越大, 模型对目标领域的情感语义越敏感, 因此在跨领域方面级情感分类时表现越好. 这验证了建模目标领域的情感语义对跨领域方面级情感分类任务的必要性, 也展现了基于文档级情感分类的语义增强模块在建模目标领域情感语义方面的有效性. 当参数 λ 大于一定数值后, 和最优的数值的相比, 模型性能都出现一定程度的下降. 这是因为过大的 λ 会使得 S3M 更偏向学习辅助的文档级情感分类任务, 从而影响主任务跨领域方面级情感分类的性能. 即使如此, 与 λ 为 0 相比, 使用基于文档级情感分类的语义增强模块总体上都为跨领域方面级情感分类带来了提升. 此外, 图 2(b) 和 (d) 表明当目标领域为 Restaurant (R) 时, S3M 需要更大的权重参数 λ 才能实现更好的性能, 这可能与辅助的情感分类任务的收敛速度有关. 从训练过程看, Yelp 文档情感分类的收敛速度更慢, 因此更大的权重参数 λ 能使模型更快更好地学会目标领域的情感语义.

表 7 基线方法和本文方法的下界与上界
Table 7 The lowerbound and upperbound results of the baseline BiLSTM-ATT and our methods

Model	R→L		L→R		R→T		T→R		L→T		T→L		Average	
	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1
LowerBound														
BiLSTM-ATT-LowerBound	62.82	53.20	71.96	50.00	37.86	37.52	55.50	45.94	42.77	41.91	54.07	45.64	54.16	45.70
BiLSTM-POSATT-LowerBound	68.81	62.96	74.75	56.74	39.42	38.63	69.64	57.00	44.82	44.72	60.87	53.93	59.72	52.33
S3M-LowerBound	69.75	64.17	76.76	65.46	52.02	52.41	70.46	62.06	50.23	50.34	64.89	58.51	64.02	58.82
UpperBound														
BiLSTM-ATT-UpperBound	71.44	65.16	77.85	65.93	69.13	66.99	76.62	62.28	68.90	66.61	69.62	63.53	72.26	65.08
BiLSTM-POSATT-UpperBound	72.41	67.52	78.76	67.61	70.69	68.74	78.57	66.11	70.37	68.47	70.66	64.01	73.58	67.08
S3M-UpperBound	73.51	69.06	79.62	68.38	71.15	69.81	78.84	66.35	70.81	68.88	71.03	64.10	74.16	67.77

4.6 基线方法和本文方法的下界与上界

为了进一步验证基于语法的注意力机制和基于文档情感分类的语义增强模块的有效性, 本文对比了基线方法 BiLSTM 和本文方法 BiLSTM-POSATT, S3M 在 6 个跨领域方面级情感分类任务上的下界性能与上界性能, 结果如表 7 所示. 测试下界性能时, 源领域训练集被视为跨领域任务的训练集, 源领域的测试集为跨领域任务的验证集, 目标领域的测试集为跨领域任务的测试集⁵⁾. 在测试上界性能时, 本文将源领域的完整训练集和目标领域训练集的 80% 组合在一起作为训练集, 目标领域训练集中剩下的 20% 作为验证集, 目标领域的测试集作为上界测试的测试集. 在测试下界性能和上界性能时, 模型的超参数保持不变.

观察下界结果可以发现, 在基线 BiLSTM-ATT 的基础上, 基于语法的注意力机制和基于文档情感分类的语义增强模块带来了显著的性能提升, 这验证了它们在零资源的跨领域方面级情感分类场景下的有效性. 值得注意的是, S3M 模型在跨领域任务 R→L 和 L→R 上的下界性能甚至与 BiLSTM-ATT 的上界接近, 这表明本文方法可以极大地降低对目标领域标注数据的要求. 在引入目标领域的标注数据进行上界测试时, 基于语法的注意力方法 BiLSTM-POSATT 比基于单词语义的注意力方法 BiLSTM-ATT 展示出更好的性能, 这是因为语法注意力从语法层面捕获方面目标对应的情感上下文, 具有更强的鲁棒性, 额外的标注数据会使得语法注意力更加准确. 此外, 在上界测试中, 辅助的文档级情感分类任务仍带来了一定的性能提升.

4.7 语法注意力可视化

为了分析基于词性的注意力机制和依存距离惩罚对注意力权重的影响, 本文以跨领域方面情感分类 R→L 作为分析任务, 可视化 BiLSTM-ATT, BiLSTM-POSATT 和 BiLSTM-POSATT+DDP 3 个模型在目标领域 Laptop 样本上的注意力权重, 颜色越深表示注意力权重越大, 结果如图 3 所示.

可以看到, 基于单词上下文的注意力模型 BiLSTM-ATT 倾向关注两个领域共享的常用词, 但未能捕捉方面目标对应的情感词. 相比之下, 采用词性注意力的模型 BiLSTM-POSATT 准确地捕获了对方面目标的情感极性至关重要的观点词. 在词性注意力的基础上, 依存距离惩罚进一步微调注意力

5) 在本小节中, 为了使同一种方法的上界性能与下界性能具有可比性, 本小节均将目标领域的测试集作为跨领域任务的测试集. 本文中的其他实验均将目标领域的训练集和测试集组合在一起作为跨领域任务的测试集, 这样观测样本更多, 结果更准确.

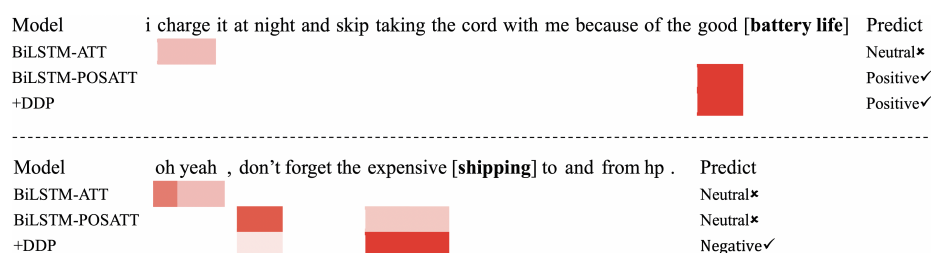


图 3 (网络版彩图) 不同模型的注意力可视化

Figure 3 (Color online) Attention visualization of different models over words

使其更加准确, 如在第 2 个样例中 BiLSTM-POSATT+DDP 更加关注观点词 “expensive”. 图 3 中的样例表明基于语法的注意力机制在跨领域场景下能更准确地捕获方面目标对应的情感表达.

5 结论与展望

由于标注困难, 方面级情感分类在很多领域上缺少标注数据, 从而限制了它在实际场景中的应用. 为此, 本文关注跨领域的方面级情感分类任务, 并提出了一种显式的基于语法和语义分割的方法 S3M 来解决这个目标领域零资源的跨领域任务. 具体而言, 本文: (1) 针对用词差异造成的领域漂移和注意力泛化问题, 首次提出利用词性、依存距离等语法信息来获得领域可迁移的注意力; (2) 提出利用与目标领域相关的文档级情感分类数据作为辅助任务帮助模型识别目标领域的情感语义. 最终, 本文提出的方法在多个跨领域方面情感分类任务上获得了最好的性能, 在平均准确率和平均 macro-F1 上比之前最好的模型 DIFD 分别提升了 7.14% 和 7.6%. 即使使用强大的预训练语言模型作为骨干网络, S3M 仍能实现显著的性能提升. 通过实验对比与分析, 可以得出结论: (1) 基于语法的注意力机制在跨领域场景能更鲁棒地捕获方面对应目标的情感表达; (2) 有效建模目标领域的情感语义对提升目标领域的方面情感分类是非常有帮助的.

尽管本文引入的与目标领域接近的文档级情感分类与对目标领域有效, 但很难保证引入的领域与目标领域的情感语义完全相同. 接下来, 作者们将考虑如何从大量无标注的目标领域文本中提取情感语义, 并与基于语法的注意力机制结合起来, 从而直接辅助目标领域的方面情感分类.

参考文献

- Pang B, Lee L. Opinion mining and sentiment analysis. *FNT Inf Retrieval*, 2008, 2: 1–135
- Liu B. Sentiment analysis and opinion mining. In: *Synthesis Lectures on Human Language Technologies*. Berlin: Springer, 2012
- Pontiki M, Galanis D, Pavlopoulos J, et al. SemEval-2014 task 4: aspect based sentiment analysis. In: *Proceedings of International Workshop on Semantic Evaluation*, Dublin, 2014. 27–35
- Kiritchenko S, Zhu X D, Cherry C, et al. NRC-Canada-2014: detecting aspects and sentiment in customer reviews. In: *Proceedings of the 8th International Workshop on Semantic Evaluation*, Dublin, 2014. 437–442
- Wagner J, Arora P, Cortes S, et al. DCU: aspect-based polarity classification for SemEval task 4. In: *Proceedings of the 8th International Workshop on Semantic Evaluation*, Dublin, 2014. 223–229
- Vo D, Zhang Y. Target-dependent twitter sentiment classification with rich automatic features. In: *Proceedings of International Joint Conferences on Artificial Intelligence*, Buenos Aires, 2015. 1347–1353
- Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. In: *Proceedings of the 3rd International Conference on Learning Representations*, San Diego, 2015

- 8 Wang Y Q, Huang M L, Zhao L, et al. Attention-based lstm for aspect-level sentiment classification. In: Proceedings of Conference on Empirical Methods in Natural Language Processing, Austin, 2016. 606–615
- 9 Ma D H, Li S J, Zhang X D, et al. Interactive attention networks for aspect-level sentiment classification. In: Proceedings of International Joint Conferences on Artificial Intelligence, Melbourne, 2017. 4068–4074
- 10 Fan F, Feng Y, Zhao D. Multi-grained attention network for aspect-level sentiment classification. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, Brussels, 2018. 3433–3442
- 11 Hu M, Wu Y, Zhao S, et al. Domain-invariant feature distillation for cross-domain sentiment classification. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, Hong Kong, 2019. 5559–5568
- 12 Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*, 1997, 9: 1735–1780
- 13 Graves A, Mohamed A, Hinton G. Speech recognition with deep recurrent neural networks. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, 2013. 6645–6649
- 14 Ganin Y, Ustinova E, Ajakan H, et al. Domain-adversarial training of neural networks. *J Mach Learn Res*, 2016, 17: 2096–2030
- 15 Liu Q, Zhang Y, Liu J. Learning domain representation for multi-domain sentiment classification. In: Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), New Orleans, 2018. 541–550
- 16 Dozat T, Manning C D. Deep biaffine attention for neural dependency parsing. In: Proceedings of the 5th International Conference on Learning Representations, Toulon, 2017
- 17 He R, Lee W S, Ng H T, et al. Exploiting document knowledge for aspect-level sentiment classification. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, 2018. 579–585
- 18 Zhao F, Wu Z, Dai X. Attention transfer network for aspect-level sentiment classification. In: Proceedings of the 28th International Conference on Computational Linguistics, Barcelona, 2020. 811–821
- 19 Devlin J, Chang M W, Lee K, et al. BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Minneapolis, 2019. 4171–4186
- 20 Rietzler A, Stabinger S, Opitz P, et al. Adapt or get left behind: domain adaptation through BERT language model finetuning for aspect-target sentiment classification. In: Proceedings of the 12th Language Resources and Evaluation Conference, Marseille, 2019. 4933–4941
- 21 Liu Y, Ott M, Goyal N, et al. RoBERTa: a robustly optimized BERT pretraining approach. 2019. ArXiv:1907.11692
- 22 Tang D Y, Qin B, Liu T. Aspect level sentiment classification with deep memory network. In: Proceedings of Conference on Empirical Methods in Natural Language Processing, Austin, 2016. 214–224
- 23 Chen P, Sun Z Q, Bing L D, et al. Recurrent attention network on memory for aspect sentiment analysis. In: Proceedings of Conference on Empirical Methods in Natural Language Processing, Copenhagen, 2017. 452–461
- 24 Ruder S. Neural transfer learning for natural language processing. Dissertation for Ph.D. Degree. Galway: University of Galway, 2019
- 25 Pan S J, Ni X, Sun J T, et al. Cross-domain sentiment classification via spectral feature alignment. In: Proceedings of the 19th International Conference on World Wide Web, Raleigh, 2010. 751–760
- 26 Glorot X, Bordes A, Bengio Y. Domain adaptation for large-scale sentiment classification: a deep learning approach. In: Proceedings of the 28th International Conference on Machine Learning, Bellevue, 2011. 513–520
- 27 Ganin Y, Lempitsky V. Unsupervised domain adaptation by backpropagation. In: Proceedings of the 32nd International Conference on Machine Learning, Lille, 2015. 1180–1189
- 28 Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. In: Proceedings of Advances in Neural Information Processing Systems, Montreal, 2014. 2672–2680
- 29 Yu J, Jiang J. Learning sentence embeddings with auxiliary tasks for cross-domain sentiment classification. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, Austin, 2016. 236–246
- 30 He R, Lee W S, Ng H T, et al. Effective attention modeling for aspect-level sentiment classification. In: Proceedings of the 27th International Conference on Computational Linguistics, Santa Fe, 2018. 1121–1131
- 31 Wu Z, Zhao F, Dai X Y, et al. Latent opinions transfer network for target-oriented opinion words extraction. In: Proceedings of the AAAI Conference on Artificial Intelligence, New York, 2020. 9298–9305
- 32 Liu H T. Dependency distance as a metric of language comprehension difficulty. *J Cogn Sci*, 2008, 9: 159–191

- 33 Dong L, Wei F, Tan C, et al. Adaptive recursive neural network for target-dependent twitter sentiment classification. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short papers), Baltimore, 2014. 49–54
- 34 Chen Z, Qian T. Transfer capsule network for aspect level sentiment classification. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, 2019. 547–556
- 35 Pennington J, Socher R, Manning C D. Glove: global vectors for word representation. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, Doha, 2014. 1532–1543
- 36 Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res, 2014, 15: 1929–1958
- 37 Kingma D P, Ba J. Adam: a method for stochastic optimization. In: Proceedings of the 3rd International Conference on Learning Representations, San Diego, 2014
- 38 Xue W, Li T. Aspect based sentiment analysis with gated convolutional networks. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, 2018. 2514–2523
- 39 Zhang K, Zhang H, Liu Q, et al. Interactive attention transfer network for cross-domain sentiment classification. In: Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, 2019. 5773–5780
- 40 Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. In: Proceedings of Advances in Neural Information Processing Systems, Long Beach, 2017. 5998–6008

Separated syntax and semantics modeling for cross-domain aspect-level sentiment classification

Zhen WU & Xinyu DAI*

National Key Lab for Novel Software Technology, Nanjing University, Nanjing 210023, China

* Corresponding author. E-mail: daixinyu@nju.edu.cn

Abstract Recently, neural network-based methods have achieved significant success in aspect-level sentiment classification (ALSC). However, the ALSC task is low-resource or even zero-resource in many domains because of complex and time-consuming data annotations. Data scarcity limits the applications of the ALSC to most domains in real-world scenarios. We thus focus on the cross-domain ALSC task. In previous attention-based methods, the distribution difference of words among domains leads to domain shift and poor generalization of the attention mechanism. To address these issues, we propose a separate syntax and semantics modeling (S3M) method to improve the attention of the neural model from the perspective of syntax and context representations from the perspective of semantics. Specifically, we find that the syntax pattern can be transferred across domains and thus propose syntax-based attention to capture the opinion expression of the given aspect. Besides, we introduce document-level sentiment classification data similar to target domain to help the model capture the sentiment semantics of target domain. We conduct experiments across six cross-domain ALSC tasks. Experimental results show that our S3M method achieves state-of-the-art performance above the benchmark. This method outperforms the previous best method DIFD by 7.14% in average accuracy and 7.6% in average macro-F1. When using BERT, BERT-ADA, or RoBERTa as a backbone, S3M still achieves over 3.5% improvement in average accuracy and in average macro-F1.

Keywords aspect-level sentiment classification, cross-domain, neural network, attention, syntax and semantics