



数据分析与知识发现
Data Analysis and Knowledge Discovery
ISSN 2096-3467, CN 10-1478/G2

《数据分析与知识发现》网络首发论文

题目：融合特征与注意力的跨领域产品评论情感分析
作者：祁瑞华，简悦，郭旭，关菁华，杨明昕
网络首发日期：2020-09-04
引用格式：祁瑞华，简悦，郭旭，关菁华，杨明昕. 融合特征与注意力的跨领域产品评论情感分析. 数据分析与知识发现.
<https://kns.cnki.net/kcms/detail/10.1478.g2.20200903.1320.006.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

融合特征与注意力的跨领域 产品评论情感分析

祁瑞华^{1,2}, 简悦^{1,2}, 郭旭², 关菁华², 杨明昕^{1,2}

¹(大连外国语大学语言智能研究中心 大连 116044)

²(大连外国语大学软件学院 大连 116044)

摘要:

[目的] 本文研究针对跨领域情感分类任务中标注数据资源相对匮乏以及从源领域到目标领域情感分类特征重要性区分问题。

[方法] 提出基于特征融合表示方法与注意力机制的跨领域双向长短时记忆产品评论情感分类模型, 融合 Bert 词向量和跨领域词向量生成跨领域统一特征空间, 通过双向长短时记忆网络结合注意力机制提取全局特征和局部特征的重要性权重。

[结果] 在亚马逊产品公开评论数据集上的对照实验结果表明, 该模型跨领域评论情感分类平均准确率达到对照模型的最高值 95.93%, 比文献中对照模型最高准确率高出 9.33%。

[局限] 还需在多领域大规模数据集上进一步检验模型的泛化性, 探究源领域知识对目标领域评论情感分类贡献规律。

[结论] 通过双向长短时记忆网络层学习融合特征能够有效获取情感语义信息, 对照实验中对目标领域最有帮助的源领域基本一致。

关键词: 特征融合; 注意力机制; 跨领域; 情感分类

分类号: TP393, G250

DOI: 10.11925/infotech..2020.0535

Sentiment Analysis of Cross-Domain Product Reviews Based on Feature Fusion

and Attention Mechanism

Qi Ruihua^{1,2}, Jian Yue^{1,2}, Guo Xu², Guan Jinghua², Yang Mingxin^{1,2}

¹(Research Center for Language Intelligence of Dalian University of Foreign Languages, Dalian 116044, China)

² (School of Software Engineering of Dalian University of Foreign Languages, Dalian 116044, China)

Abstract:

[Objective] This paper focus on the shortage of labeled data resources and the difficulty of distinguishing the weights of sentiment characteristics from the source domain to the target domain in the task of cross-domain sentiment classification.

[Methods] This paper proposed a feature fusion representation method and the attention mechanism based cross-domain product reviews sentiment classification model. This model integrates Bert and cross-domain word vector to generate cross-domain unified feature space, and extracts the importance weight of global and local features through attention mechanism.

[Results] The results of the controlled experiment on public Amazon reviews data set show that the average accuracy of the model reaches the highest value of 95.93% among the control models, 9.33% higher than that of the best control model.

[Limitations] It is necessary to further test the generalization of the model in large-scale multi-domain data sets and explore the contribution of source domain knowledge to target domain review sentiment classification.

[Conclusions] It is feasible to effectively acquire sentiment semantic information by learning fusion features through BiLSTM and attention mechanism. The source fields that are most helpful to the target fields in the controlled experiments are basically the same.

Keywords: feature fusion; attention mechanism; cross-Domain; sentiment classification

1 引言

情感分析技术自动探究用户对事件、议题或人物等的态度观点，在舆情分析、社交媒体和电商评论挖掘等任务中起到重要作用。由于领域差异、语义差异、特征差异和极性差异导致的数据稀疏始终是精准情感分析面临的主要挑战^[1]，以社交媒体为例，海量用户数据分散于不同平台和各个领域，现有情感分析技术对领域高度敏感^[2]，需要大量领域标注数据以达到好的分析效果，然而标注数据资源需要耗费大量人力，如果为每个领域都准备大规模标注数据，成本将会非常高昂。在这种情况下，运用跨领域信息共享迁移学习^[3]或者领域适配^[4]技术进行跨领域情感分析成为当前的研究热点。

跨领域情感分析技术的目标是充分利用相近领域的标注数据，来完成标注数据匮乏的目标领域情感分析任务，主要思路是基于源领域数据构建可应用于目标领域的知识模型。传统的跨领域情感分类主要基于词汇方法和领域相似度，此类情感分析算法处理高维特征的性能有限，难以实现语义匹配。近年来，运用向量、张量等代数结构综合表示文本单元的词嵌入分布式表示和深度学习算法倍受关注^[5-7]，但深度学习算法的黑盒特性欠缺可解释性，无法解读从源领域到目标领域的情感分类过程中哪些信息更值得学习。为此，本文提出适用于跨领域情感分类的特征融合表示方法和基于注意力机制的双向长短时记忆情感分类模型。

2 相关工作

情感分类任务与领域高度相关，直接将从源领域学习的情感分类模型应用于目标领域时通常表现不佳。跨领域情感分类领域现有研究主要围绕两个关键问题：跨领域情感特征表示和情感分类算法，主要方法包括：实例迁移、特征迁移、词嵌入、深度学习和注意力机制等。

2.1 实例迁移和特征迁移

早期跨领域情感特征表示方法主要基于实例迁移或特征迁移，情感分类方法主要采用情感词典或机器学习算法。实例迁移方法的思路是利用源领域的标注数据扩充目标领域数据，如 Gamon 等^[8]运用多领域标注数据集训练朴素贝叶斯 Bootstrapping 多分类器，Dai^[9]提取源领域中与测试集分布相同的标注数据，基于 TrAdaBoost 算法构建目标领域分类模型。为了进一步改善跨领域情感特征迁移效率，Zhang 等^[10]将领域语料根据文档各部分对情感倾向的贡献分为关键情感句和细节描述句，通过情感纯度、关键词和位置属性建立启发规则抽取关键情感句改善跨领域情感分类性能。总体上，实例迁移法适用于领域相似度较高的情况，当领域差异大或样本噪声高时性能不佳。

特征迁移方法假设源领域与目标领域部分特征一致，通过特征变换将多领域稀疏特征映射到低维空间，如 Blitzer 等^[11]提出领域特征与标注之间的结构对应学习 SCL 算法和互信息 SCL-MI 方法^[12]，基于共现核心词和非核心词构建支持向量机实现跨领域学习，Pan 等^[13]提出谱特征对齐算法 SFA，通过共现

双边图对齐领域特征和领域无关特征，基于谱聚类将连通特征加入到原始特征集，利用领域互信息进行特征提取和情感分类。以核心词为基础的方法和情感词典方法一般基于词袋模型的原理，性能高度依赖于核心词或情感词选择的质量和覆盖度，难以分析一词多义和语义匹配。为此，Wei 等^[14]提出跨领域语义关联自动匹配方法，使用情感极性稳定的词汇表示强情感特征，抽取高频核心词用于发现跨领域语义相似的情感特征，通过 Word2Vec 构建相似核心词特征改善了跨领域情感特征匹配效果和情感分类性能。

2.2 词嵌入情感特征

近年基于深度学习的特征表示学习技术发展迅速，突破了传统情感特征表示的局限。跨领域词嵌入情感特征的代表性研究主要有：Glorot 等^[15]运用堆栈去噪自编码器 SDA 算法以随机梯度下降的贪婪分层方式从多领域评论语料抽取高层特征，经由整流器单元生成新特征空间的维度非线性映射特征，通过中间概念实现跨领域特征学习。Chen 等^[16]改进了 Glorot 等^[15]的研究，提出边缘化 SDA 的 mSDA 算法提取主特征，提高了算法运行效率和可扩展性。上述以 SDA 为基础的方法受限于计算复杂度，仅考虑采用高频 Unigrams 和 Bigrams 作为输入特征，并未考虑低频词和语义表示。

Mikolov 等^[17]证明了基于连续空间的词嵌入表示有利于捕捉句法和语义规律。由于句法和语义规律是情感分析的基本要素，词向量情感特征在跨领域情感分类任务中的效果日益显现，如 Akhtar 等^[18]提出基于混合深度学习卷积神经网络学习词向量，从原始文本捕捉语义关联，在多目标优化框架的基础上采用支持向量机对情感增强优化向量进行分类。Yu 等^[19]基于权重对数似然比改进了 SCL 算法^[11]，借助辅助任务和卷积神经网络生成句嵌入向量超特征表示，在领域相关情感词抽取和情感分类任务上取得明显效果。Li 等^[20]结合卷积神经网络和对抗监督，基于领域特有信息和领域通用信息生成跨领域共享的特征表示。Zhang 等^[21]基于 ELMo 词嵌入表示通过隐含变量和高斯混合向量捕捉情感子类型，发现情感语法表达优于传统神经编码器。现有研究表明，词嵌入情感特征能够揭示内在语义信息，基于词嵌入情感特征的方法在跨领域情感分类任务上表现优于对照方法。

2.3 深度学习情感分类

深度学习算法具有局部特征抽象和特征记忆的优势，Tai 等^[22]指出长短时记忆网络(Long Short-Term Memory, LSTM)具有处理变长序列的能力，能够有效捕捉长距离依存关系中蕴含的语义信息，并构建了 LSTM 树型结构成功应用于语义相关性分析和情感分类任务。Qian 等^[23]利用情感词典、否定词和强度词构建语言学启发序列正则化 LSTM 模型，无需依赖句子解析和短语级标注，简洁有效地实现了句子级情感分类。

跨领域情感分类算法方面, Ganin 等^[7]受领域适应理论的启发, 认为有效率的跨领域特征不应仅从源领域学习, 而应同时学习源领域和目标领域的共同特征, 由此提出结合领域对抗神经网络 DANN 和 mSDA 的领域适应学习方法, 通过域对抗使领域深层特征分布一致, 训练生成混合领域神经网络模型并成功应用于文档情感分析和图像分类任务。Wei 等^[24]基于卷积神经网络和长短时记忆模型提出迁移学习框架, 在 MOOC 论坛情感分类、急迫性识别和 Amazon 产品评论情感分类等跨领域任务上效果良好。

2.4 注意力机制

注意力机制能够增强超特征学习过程中底层单元的权重表示^[25], 清晰地阐明情感分类任务中每句话或每个单词在分类预测中的权重, 受到越来越多学者的关注, 如 Yang 等^[26]提出文档层次注意力网络表示单词和句子两级注意力权重, Gu 等^[27]建立层次多通道结构引入注意力机制从文本和音频中抽取高层情感特征和主观信息。Li 等^[28]提出端到端的对抗记忆网络 AMN 自动捕捉领域核心词, 并进一步提出分层注意力转移网络 HATN^[29], 通过注意机制的层次结构反映文档结构, 自动表示领域特征词和句子级信息, 实现了核心特征和非核心特征的定位。

注意力机制在抽取全局关系的同时关注元素的局部联系, 通过并行计算降低时间复杂度, 有效改善了跨领域情感分类性能, 但其并行处理序列元素的过程忽略了输入序列顺序, 在处理文本时缺失自然语言语序等重要信息。

综上所述, 在语料规模有限的情况下无法保证源领域与目标领域的相似度, 当源领域和目标领域差异较大时, 实例迁移方法的情感分析准确率明显下降。而特征迁移方法无法区分源领域核心词和情感词在目标领域表达的含义是否相同甚至相反, 例如: 图书产品评论中的“这本书纸质很厚”中的“厚”表达的是正面情感, 而在电子产品评论中的“厚”通常表达的是负面情感。

为更好地表示源领域和目标领域共同的语义信息, 本文提出通过跨领域预训练学习构建词嵌入情感特征融合表示模型。为充分学习跨领域情感特征, 考虑注意力机制在全局关系和局部关系抽取上具有优势但忽略语序信息, 本文采用深度学习双向长短时记忆算法增强局部特征和语序记忆, 提出面向跨领域词嵌入情感特征融合表示的注意力双向长短时记忆情感分类模型。

3 基于特征融合与注意力的跨领域情感分类模型

3.1 任务描述

假设给定的评论数据源领域 $D_s = \{(X_{s_i}, Y_{s_i})\}_{i=1}^{n_s}$ 中包含具有情感倾向标注的评论文本 n_s 个, X_{s_i} 表示 D_s 中第 i 条标注评论文本, Y_{s_i} 表示 X_{s_i} 对应的情感倾向,

$Y_{s_i} \in \{0,1\}$, 其中 0 表示积极情感倾向, 1 表示消极情感倾向; 给定与 D_S 不同来源的目标领域 $D_T = \{(X_{T_i})\}_{i=1}^{n_T}$ 中包含未标注评论文本 n_T 个, X_{T_i} 表示 D_T 中第 i 个未标注文本。跨领域情感二分类任务可以形式化为: 在 D_S 上拟合一个情感分类模型来预测 D_T 中的 X_{T_i} 的情感倾向 Y_{T_i} , 表示为 $D_S \rightarrow D_T$ 。

本文基于领域适应理论^[30]提出基于特征融合与注意力机制的跨领域情感分类模型, 首先通过 Bert(Bidirectional Encoder Representation from Transformers) 词向量和 Skip-gram 词表示模型将源领域 D_S 和目标领域 D_T 投影到相同的特征空间, 获得源领域和目标领域的特征融合后的超特征表示; 然后通过双向长短时记忆网络和注意力机制对 D_S 监督学习建立模型, 将模型和参数设置共享到 D_T , 对 D_T 的数据进行跨领域情感分类。模型架构分为特征融合表示学习层、双向长短时记忆网络层、注意力计算层、全连接层和情感预测层五个部分, 如图 1 所示。

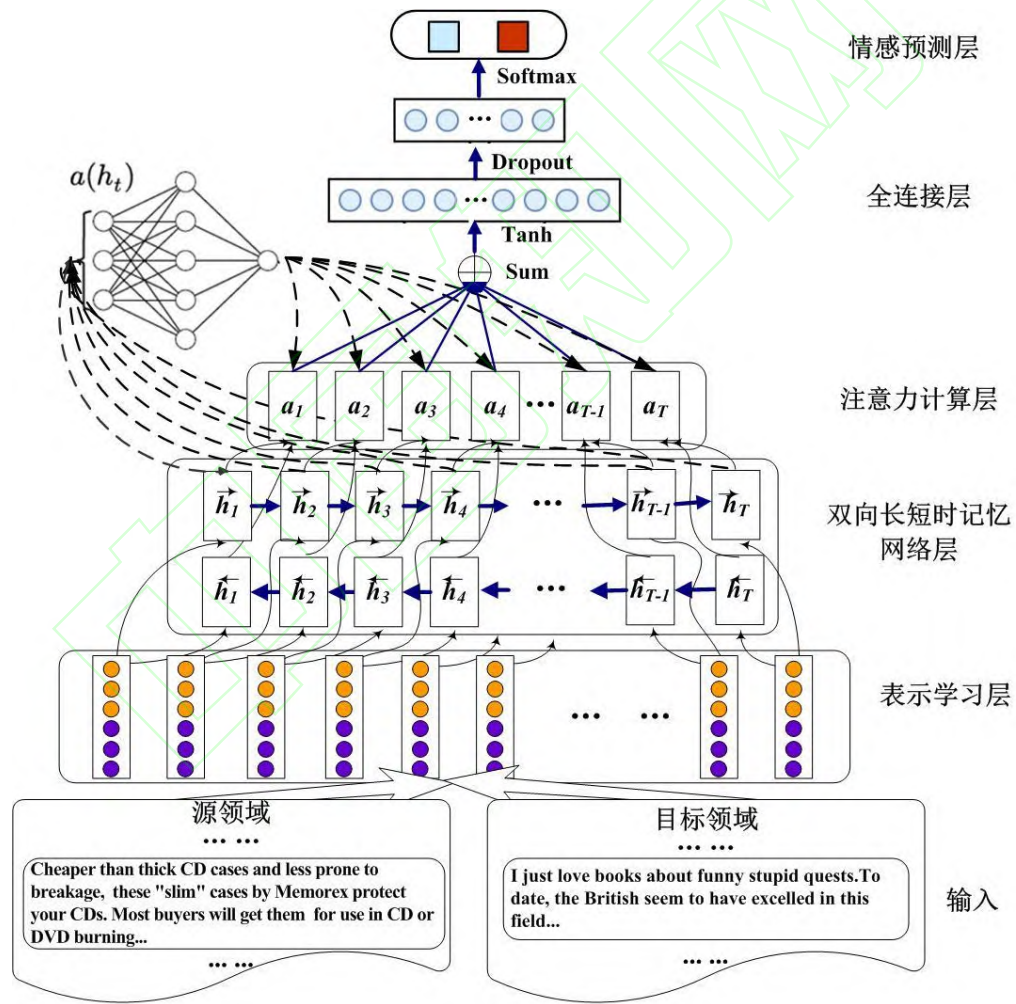


图 1 基于特征融合与注意力的跨领域评论情感分类模型

Fig1. Sentiment Classification Model for Cross-Domain Product Reviews Based on Feature Fusion and Attention Mechanism

3.2 表示学习层

在表示学习层，采用 Bert 词向量和源领域 D_s 与目标领域 D_T 评论文本预训练词向量相连接作为单词表示，融合单词的全局语义信息和跨领域局部语义信息构成特征表示模型。Bert 词向量基于 Transformer 结构采用掩码语言模型和后句预测方法在大规模数据上训练，融合词上下文信息的同时捕获句子之间关系，获取双向特征表示。源领域与目标领域评论文本预训练词向量通过神经语言模型 Skip-gram 将文档中的词预先训练为词向量，形成领域融合的分布式词向量，投射到统一的 k 维空间 D 。在这个跨领域向量空间中，将源领域与目标领域的样本表示为领域无关的词向量，语义相近的词在向量空间中的距离也相近，从而改善情感特征不匹配导致的跨领域语义匹配问题。Skip-gram 词表示模型复杂度和时间复杂度低，运行效率高，适用于本任务。预训练的具体过程为：每次从目标词 w 的上下文 c 中选择一个词，将其词向量作为模型的输入，对于整个语料的优化目标为最大化公式(1)^[31]。

$$\sum_{(w,c) \in D} \sum_{w_j \in c} \log P(w | w_j) \quad (1)$$

设 V 为词表， $e^{(w)}$ 为词 w 作为上下文时的表示， $e'(w)$ 为词 w 作为目标词时的表示，则(1)中 $P(w | w_j)$ 计算如下^[31]：

$$P(w | w_j) = \frac{\exp(e'(w)^T e(w_j))}{\sum_{w' \in V} \exp(e'(w')^T e(w_j))} \quad (2)$$

本文的特征融合表示模型由 Bert 词向量和源领域与目标领域评论文本预训练词向量相连接构成，如图 1 所示。将每条评论文本视作一个长度为 l 的文档，超过长度 l 的长评论文本取前 l 个词，长度不足 l 的短评论文本后面用零补齐到长度 l 。评论文本中的每个词 x_i ，分别被投射到对应的词向量 $v_i \in \mathbf{R}^k$ ，其中 k 表示词向量维度。每条评论文本 $W \in \mathbf{R}^{lk}$ 由词向量 $W_{1:l} = v_1 \oplus v_2 \dots \oplus v_l$ 连接而成，其中 \oplus 代表连接符号，对于预训练集合词汇表外的单词进行随机初始化处理。生成的词向量序列按顺序逐个输入到下一层的双向长短时记忆网络层。

3.3 双向长短时记忆网络层

双向长短时记忆网络(Bi-directional Long Short-Term Memory, BiLSTM)^[32] 由两个 LSTM 组成，每个 LSTM 由一系列处理器单元构成，包括输入门 i_t 、遗忘门 f_t 、输出门 o_t 和记忆单元 c_t ，可以看作前馈神经网络中的神经元，它们完成了精细变换后激活函数的计算。在学习过程中根据规则将有价值的信息留下，不符合规则的信息则通过遗忘门被抛弃。这些控制门和记忆单元的结合同样能够更好地表示和控制输入序列的长期记忆和短期记忆，在任意一个时刻，每个单元会接收到来自于前面隐层的输入和此刻的输入单元，提高了模

型处理远距离依赖问题的能力。

设输入词向量为 $\{x_1, x_2, \dots, x_n\}$ ，其中 x_t 表示一个单元的输入向量，即表示学习层生成的一个文档中一个单词的词向量，则 LSTM 单元前向传播的输入门、遗忘门和输出门和记忆单元计算公式如下^[32]：

$$\begin{cases} f_t = \sigma(w_f x_t + U_f h_{t-1} + b_f) \\ i_t = \sigma(w_i x_t + U_i h_{t-1} + b_i) \\ o_t = \sigma(w_o x_t + U_o h_{t-1} + b_o) \\ c_t = f_t * c_{t-1} + i_t * \tanh(w_c x_t + U_c h_{t-1} + b_c) \\ h_t = o_t * \tanh(c_t) \end{cases} \quad (3)$$

其中， h_t 表示网络中的隐藏层状态， w_f, w_i, w_o 为权重矩阵， b_f, b_i, b_o 为权重矩阵的偏置值， σ 表示激活函数。 $*$ 表示点乘积。

BiLSTM 中每个时刻状态如公式(4)所示，输出由两个方向的 LSTM 共同决定。其中 u_t 和 v_t 分别表示正向输出和反向输出的权重矩阵， b_t 表示 t 时刻的偏置值^[32]。

$$\begin{cases} \bar{h}_t = LSTM(\bar{x}_t, \bar{h}_{t-1}) \\ \tilde{h}_t = LSTM(\tilde{x}_t, \tilde{h}_{t-1}) \\ h_t = u_t \bar{h}_t + v_t \tilde{h}_t + b_t \end{cases} \quad (4)$$

3.4 注意力计算层

模型中注意力计算层的任务是提取输入信息中的主要部分，有效突出情感分类预测过程中关键信息的权重。本文的注意力计算层采用前馈注意力机制^[26]，对于每个记忆单元使用前馈神经网络计算其与情感类别的关联。本层的输入是长短时记忆网络层输出的隐藏状态 h_t ，设 T 为输入序列时间的总步数，注意力机制计算状态序列 h 的权重并从整个序列生成表示注意力权重的上下文向量 c_t ， c_t 计算公式如下^[26]：

$$c_t = \sum_{i=1}^T \alpha_i h_i \quad (5)$$

设 a 为计算 h_t 梯度重要性的学习函数，公式(5)中 α_i 计算公式如下^[26]：

$$\alpha_i = \frac{\exp(e_i)}{\sum_{k=1}^T \exp(e_k)}, \quad e_i = a(h_i) \quad (6)$$

3.5 全连接层

全连接层的任务是将学到的分布式特征表示映射到样本标注空间，本文模型全连接层采用 \tanh 激活函数。为防止过拟合，在训练过程中加入 DropOut 层，在更新参数时随机丢弃 50% 的输入神经元。

3.6 情感预测层

情感预测层将特征向量通过全连接方式采用 Softmax 激活函数映射输出情感预测二分类结果，激活函数计算公式如下^[33]：

$$y = \text{softmax}(w^T \cdot z + \tilde{b}) \quad (7)$$

其中， w 为全连接层权重， \tilde{b} 为偏置项。最后计算出的 $p(y_i|x)$ 代表评论情感极性的概率分布，其中 y_1 表示积极情感， y_2 表示消极情感。

模型采用二元交叉熵作为损失函数，优化目标是最小化训练样本预测输出值和实际样本值的交叉熵。交叉熵损失函数公式如下^[34]：

$$\text{loss} = - \sum_{i=1}^n y_i \log y_i + (1 - y_i) \log(1 - y_i) \quad (8)$$

其中， y_i 为第 i 个样本的实际类别， \hat{y}_i 为预测类别。情感预测层输出对每一条评论的情感预测结果。

4 实验及结果分析

4.1 实验数据

本文采用文献^[12]提供的亚马逊网站产品评论作为实验数据集，该数据集中包含 Books、DVD disk、Electronics、Kitchen appliances 和 Videos 五个领域的产品评论文本，每个领域包含 6000 条有标注的数据，积极评论和消极评论各 3000 条。除此之外，还包括原始无标注评论语料共 82638 条，平均每个领域 16527 条。数据集中各个领域语料统计信息如表 1 所示。

表 1 实验数据集

Table1 Experimental data set

领域	积极评论	消极评论	无标注文本
Books	3000	3000	9750

DVD disk	3000	3000	11843
Electronics	3000	3000	17009
Kitchen appliances	3000	3000	13856
Videos	3000	3000	30180

实验每次将五个领域中的一个领域数据集作为源领域 D_S ，另一个领域数据集作为目标领域 D_T ，并将 D_S 和 D_T 分割为训练集和测试集，训练集使用 D_S 和 D_T 中 2800 条正向评论和 2800 条负向评论，测试集使用目标领域全部 6000 条有标注评论进行情感分类预测。

4.2 实验参数

基于特征融合与注意力的跨领域评论情感分类实验参数设置如表 2 所示，其中基于 Bert 实验的词向量维度为 768，Bert 融合跨领域特征以及 Bert 融合跨领域特征注意力机制实验的词向量维度为 1068。

表 2 跨领域情感分类实验参数

Table2 Experimental parameters of cross-domain sentiment classification

参数名称	值	参数名称	值
最大长度	120	优化器	Adam
词向量维度	768,1068,1068	损失函数	binary_crossentropy
LSTM 隐藏单元	128	batch_size	32
全连接层 1	128, activation="tanh"	全连接层 2	2, activation="softmax"
dropout	0.5	输出层激活函数	Softmax

4.3 对照实验

本文实验主要探究两个问题：一是本文特征融合表示模型和跨领域情感分类模型预测的效果如何；二是对于特定的目标领域，从哪些源领域学习的效果最好。为验证本文模型效果，与以下文献中的方法进行对照：

(1) S-only(Source-only): 基线方法，基于神经网络的非自适应使用三层全连接神经网络构建的逻辑回归分类器模型，取频次最高的 5000 个 Unigram 和 Bigram 作为两个领域的特征。

(2) SFA: 文献[13]中通过谱特征来对齐核心特征和非核心特征的线性方法，利用两个领域之间的互信息进行特征提取和分类。

(3) mSDA: 文献[16]中的 SDA 模型的改进模型, 该方法在通过 SDA 进行主特征提取的过程中提高了算法运行的效率。

(4) DANN: 文献[7]中的域对抗神经网络模型, 此方法先与 mSDA 结合提取特征, 然后通过域对抗使领域深层特征分布一致, 再使用源领域深层特征进行训练, 最终对目标领域进行情感分类。

(5) HATN: 发表于 AAAI2018 的文献[29]中的分层注意力转移网络模型, 通过注意力机制层次结构反映文档层次结构, 优化核心特征和非核心特征的定位。

(6) CDSA-B: 本文跨领域评论情感分类模型的对照模型, 表示学习层采用 Bert 词向量, 无注意力机制。

(7) CDSA-F: 本文提出的特征融合跨领域评论情感分类模型, 表示学习层采用融合跨领域特征。

(8) CDSA-F-Att: 本文提出的特征融合与注意力跨领域评论情感分类模型, 表示学习层采用融合跨领域特征词向量, 加入注意力计算层。

4.4 实验结果与分析

对照实验在 Amazon 多领域评论数据集上的结果如表 3 所示, 本文方法在跨领域情感分类任务上的平均准确率为 95.93%, 最高准确率达到 97.1%, 明显高于文献中的对照方法, 比对照方法中准确率最高的 HATN 高出 9.33%。对照方法中 S-only 的准确率最低, 平均准确率只有 75.92%, 说明在不同领域中的评论数据上训练得到的模型, 并不适合直接用于新的领域的情感分析; SFA 的平均准确率为 78.69%, 主要原因是方法的离散性比较差, 使用较简单的线性分类方法, 性能高度依赖于手动选择核心特征的准确性; DANN 和 mSDA 算法的平均准确率分别达到 79%和 82.36%, 准确率有一定的改善; HATN 算法在 GAN 基础上增加了注意力机制, 自动选择领域特征词并且能得到句子级权重信息, 效果有较大改进, 准确度达到 86.6%。

本文提出的跨领域特征融合方法 CDSA-F 在表示学习层加入融合特征共享不同领域之间的特征表示, 通过双向长短时记忆网络层学习获取语义信息, 情感分类平均准确率达到 95.57%, 比只采用 Bert 词向量作为特征表示的对照方法提升了 1.15%; 加入注意力机制的 CDSA-F-Att 方法在此基础上较好地抽取了底层单元的跨领域情感特征权重信息, 平均准确率进一步提高了 0.36%, 达到了八种对照模型的最高值 95.93%, 证明了通过双向长短时记忆网络层学习融合特征能够有效获取情感语义信息。

表 3 亚马逊评论数据集上的实验准确率

Table3 Experimental accuracy on Amazon review data set

源领域	目标领域	S-only	SFA	DANN	mSDA	HATN	CDSA-B	CDSA-F	CDSA-F-Att
Book	DVD	0.8057	0.8285	0.8342	0.8612	0.8707	0.932	0.9496	0.9546
Book	Electronic	0.7365	0.7638	0.7627	0.7902	0.8575	0.9368	0.9546	0.9543
Book	Kitchen	0.7163	0.7810	0.7790	0.8105	0.8703	0.9465	0.952	0.9611
Book	Video	0.8145	0.8295	0.8323	0.8498	0.8780	0.9556	0.961	0.9653
DVD	Book	0.7645	0.8020	0.8077	0.8517	0.8778	0.9341	0.9455	0.9486
DVD	Electronic	0.7312	0.7600	0.7635	0.7617	0.8632	0.9286	0.9508	0.9535
DVD	Kitchen	0.7343	0.7750	0.7815	0.8260	0.8747	0.9476	0.9628	0.9651
DVD	Video	0.8275	0.8595	0.8595	0.8380	0.8912	0.9615	0.9665	0.971
Electronic	Book	0.6887	0.7235	0.7353	0.7992	0.8403	0.9315	0.947	0.9486
Electronic	DVD	0.7260	0.7593	0.7627	0.8263	0.8432	0.9338	0.9568	0.9613
Electronic	Kitchen	0.8463	0.8650	0.8453	0.8580	0.9008	0.957	0.9641	0.966
Electronic	Video	0.7248	0.7565	0.7720	0.8170	0.8408	0.9548	0.9623	0.9671
Kitchen	Book	0.7153	0.7397	0.7417	0.8055	0.8488	0.9368	0.9403	0.95
Kitchen	DVD	0.7332	0.7567	0.7532	0.8218	0.8472	0.934	0.9581	0.9603
Kitchen	Electronic	0.8315	0.8538	0.8553	0.8800	0.8933	0.9463	0.96	0.9601
Kitchen	Video	0.7608	0.7797	0.7637	0.8147	0.8485	0.959	0.9691	0.968
Video	Book	0.7703	0.7948	0.8003	0.8300	0.8710	0.9388	0.9418	0.9419
Video	DVD	0.8243	0.8365	0.8415	0.8590	0.8790	0.953	0.9596	0.965
Video	Electronic	0.7187	0.7593	0.7572	0.7767	0.8598	0.94	0.9498	0.9586
Video	Kitchen	0.7133	0.7478	0.7522	0.7952	0.8645	0.958	0.9628	0.9671
平均准确率		0.7592	0.7869	0.7900	0.8236	0.8660	0.9442	0.9557	0.9593

为探究对于特定目标领域从哪些源领域学习达到的效果最好,计算表 3 中八个模型在源领域-目标领域实验中的平均准确率,生成准确率矩阵。并通过如图 2 所示的热力图直观地呈现跨领域情感分类效果。热力图中的列标记表示源领域,行标记表示目标领域,矩阵中的色块颜色越深表示分类准确率越高,反之颜色越浅表示准确率越低。

从图 2 中可以看出对于 Book 目标领域,最高准确率的源领域是 DVD 和

Video, Video 源领域的效果最好；对于 DVD 目标领域，Book 和 Video 作为源领域的准确率最高，Video 源领域的效果最好；对于 Electronics 目标领域，最高准确率的源领域是 Kitchen；对于 Kitchen 目标领域，源领域 Electronic 的准确率最高；对于 Video 目标领域，获得最高准确率的源领域是 DVD 和 Books，其中 DVD 源领域的效果最好。

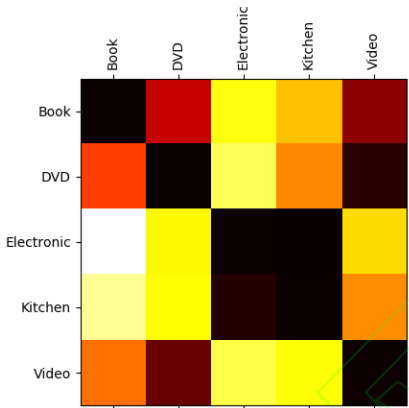


图 2 .八个模型在源领域-目标领域实验中的平均准确率

Fig2. Average accuracy of eight model in source-target domain experiments

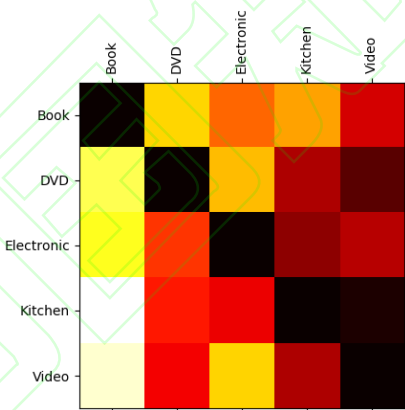


图 3 .CDSA-F 模型在源领域-目标领域实验中的平均准确率

Fig3. Average accuracy of CDSA-F model in source-target domain experiments

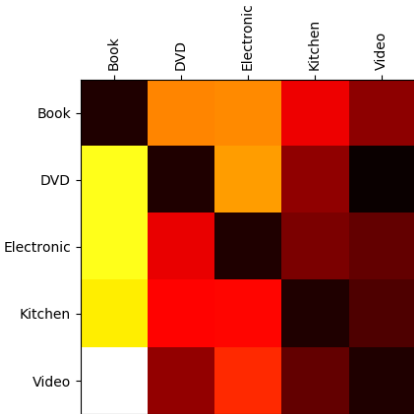


图 4 .CDSA-F-Att 模型在源领域-目标领域实验中的平均准确率

Fig4. Average accuracy of CDSA-F-Att model in source-target domain experiments

为进一步验证对于特定目标领域有帮助的源领域，计算 CDSA-F 模型和 CDSA-F-Att 模型在源领域-目标领域实验中的平均准确率矩阵，对应的平均准确率热力图分别如图 3 和图 4 所示。分析实验结果发现，各个算法结果数据中对目标领域最有帮助的源领域基本一致，如图 3 和图 4 所示的跨领域特征融合、跨领域特征融合与注意力机制实验结果的源领域-目标领域配对准确率对照，与图 2 中的八个对照模型平均准确率的分布基本一致，即跨领域评论情感分类互补效果好的领域有 Book 和 DVD，Book 和 Video，Video 和 DVD，Kitchen 和 Electronic，其中 Book、DVD 和 Video 三个领域的跨领域信息对彼此的分类贡献最大。

5 总结与展望

本文选择亚马逊网站在书籍、数字视频光盘、电子产品、厨房用品和录像带五个类目下的英文评论作为实验数据，针对跨领域情感分类任务中标注数据资源相对匮乏以及从源领域到目标领域情感分类特征重要性区分问题，提出基于特征融合表示方法与注意力机制的跨领域双向长短时记忆产品评论情感分类模型，融合 Bert 词向量和跨领域词向量生成跨领域统一特征空间，通过双向长短时记忆网络结合注意力机制提取全局特征和局部特征的重要性权重，在亚马逊产品公开评论数据集上的对照实验结果表明该模型明显提高了跨领域评论情感分类的准确率，实验结果证实了本文模型的有效性，并进一步探究了源领域和目标领域的搭配效果。

在未来的工作中，将在多领域大规模数据集上进一步检验模型的泛化性，进一步探究源领域知识对目标领域评论情感分类贡献规律，并结合更多的无标签评论数据来提高模型的预测能力。

参考文献:

- [1]Al-Moslmi T, Omar N, Abdullah S, et al. Approaches to cross-domain sentiment analysis: A systematic literature review[J]. IEEE Access, 2017, 5: 16173-16192.
- [2]Lai S, Liu K, He S, et al. How to generate a good word embedding[J]. IEEE intelligent systems, 2016, 31(6): 5-14.
- [3]Weiss K, Khoshgoftaar T M, Wang D D. A survey of transfer learning[J]. Journal of big data, 2016, 3(1): 1-40.
- [4]Tahmoresnezhad J, Hashemi S. Visual domain adaptation via transfer feature learning[J]. Knowledge and information systems, 2017, 50(2): 1-21.

- [5]Sun M, Tan Q, Ding R, et al. Cross-domain sentiment classification using deep learning approach.[C]//2014 IEEE 3rd International conference on cloud computing and intelligence systems. IEEE.2014: 60-64.
- [6]余传明. 基于深度循环神经网络的跨领域文本情感分析[J].图书情报工作,2018,62(11):23-34.(Yu C,A cross-domain text sentiment analysis based on deep recurrent neural network[J].Library and information service, 2018,62(11):23-34.)
- [7]Ganin Y, Ustinova E, Ajakan H, et al. Domain-adversarial training of neural networks[J]. The Journal of machine learning research, 2016, 17(1): 2096-2030.
- [8]Gamon M, Aue A. Automatic identification of sentiment vocabulary: exploiting low association with known sentiment terms[C]//Proceedings of the ACL workshop on feature engineering for machine learning in natural language processing.USA: Association for Computational Linguistics.2005: 57-64.
- [9]Dai W, Yang Q, Xue G R, et al. Boosting for transfer learning[C]//Proceedings of the 24th international conference on Machine learning. USA:Association for Computing Machinery.2007: 193-200.
- [10]Zhang S, Liu H, Yang L, et al. A cross-domain sentiment classification method based on extraction of key sentiment sentence[C]//Proceedings of the 4th CCF conference on natural language processing and Chinese computing-Volume 9362. Germany:Springer-Verlag. 2015: 90-101.
- [11]Blitzer J, McDonald R, Pereira F. Domain adaptation with structural correspondence learning[C]//Proceedings of the 2006 conference on empirical methods in natural language processing. USA: Association for Computational Linguistics.2006: 120-128.
- [12]Blitzer J, Dredze M, Pereira F. Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification[C]//Proceedings of the 45th annual meeting of the association of computational linguistics. USA: Association for Computational Linguistics.2007: 440-447.
- [13]Pan S J, Ni X, Sun J T, et al. Cross-domain sentiment classification via spectral feature alignment[C] // Proceedings of the 19th International conference on World Wide Web.USA: Association for Computing Machinery. 2010: 751-760.
- [14]Wei X, Lin H, Yang L. Cross-domain sentiment classification via constructing semantic correlation[J]. IAENG International Journal of Computer Science, 2017, 44(2): 172-179.
- [15]Glorot X, Bordes A, Bengio Y. Domain adaptation for large-scale sentiment classification: A deep learning approach[J]. 2011.[C]// Proceedings of the 28th international conference on Machine Learning. USA: Omnipress.2011:513-520.
- [16]Chen M, Xu Z, Sha F, et al. Marginalized denoising autoencoders for domain adaptation [C]//The 29th international conference on machine learning,USA:Omnipress. 2012: 1627-1634.

- [17]Mikolov T, Yih W, Zweig G. Linguistic regularities in continuous space word representations[C]//Proceedings of the 2013 conference of the north American chapter of the association for computational linguistics: Human language technologies.USA: Association for Computational Linguistics. 2013: 746-751.
- [18]Akhtar M S, Kumar A, Ekbal A, et al. A hybrid deep learning architecture for sentiment analysis[C]//Proceedings of the 26th international conference on computational linguistics: Technical papers.The COLING 2016 Organizing Committee. 2016: 482-493.
- [19]Yu J, Jiang J. Learning sentence embeddings with auxiliary tasks for cross-domain sentiment classification[C]//Proceedings of the 2016 conference on empirical methods in natural language processing. USA: Association for Computational Linguistics. 2016: 236-246.
- [20]Li Y, Baldwin T, Cohn T. What's in a domain? Learning domain-robust text representations using adversarial training[C]//Proceedings of the 2018 conference of the north American chapter of the association for computational linguistics: human language technologies, Volume 2 (Short Papers). USA: Association for Computational Linguistics.2018:474-479.
- [21]Zhang L, Tu K,Zhang Y. Latent variable sentiment grammar[C]//Proceedings of the 57th annual meeting of the association for computational linguistics.USA: Association for Computational Linguistics.2019.4642-4651.
- [22]Tai K S, Socher R, Manning C D. Improved semantic representations from tree-structured long short-term memory networks[C]//Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing (Volume 1: Long Papers).USA: Association for Computational Linguistics.2015: 1556-1566.
- [23]Qian Q,Huang M, Lei J, et al. Linguistically regularized LSTM for sentiment classification[C]// proceedings of the 55th annual meeting of the association for computational linguistics (Volume 1: Long Papers).USA: Association for Computational Linguistics. 2017:1679 -1689.
- [24]Wei X, Lin H, Yang L, et al. A convolution-LSTM-based deep neural network for cross-domain MOOC forum post classification[J]. Information,2017,8(3): 2078-2489.
- [25]Raffel C, Ellis D P W. Feed-forward networks with attention can solve some long-term memory problems[J]. arXiv preprint arXiv:1512.08756, 2015.
- [26]Yang Z, Yang D, Dyer C, et al. Hierarchical attention networks for document classification[C]//Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies.USA: Association for Computational Linguistics. 2016: 1480-1489.
- [27]Gu Y , Yang K , Fu S , et al. Multimodal affective analysis using hierarchical attention strategy with word-level alignment[C]// Proceedings of the 56th annual meeting of the association for computational linguistics (Volume 1: Long Papers). USA: Association for Computational Linguistics.2018:2225–2235.

- [28]Li Z, Zhang Y, Wei Y, et al. End-to-end adversarial memory network for cross-domain sentiment classification [C]//International joint conference on artificial intelligence 2017. USA:AAAI Press,2017: 2237-2243.
- [29]Li Z, Wei Y, Zhang Y, et al. Hierarchical attention transfer network for cross-domain sentiment classification[C]//32th AAAI conference on artificial intelligence, USA: AAAI Press, 2018:5852-5859.
- [30]Ben-David S, Blitzer J, Crammer K, et al. A theory of learning from different domains[J]. Machine learning, 2010, 79(1-2): 151-175.
- [31]Jatnika D, Bijaksana M A, Suryani A A. Word2Vec model analysis for semantic similarities in English words [J]. Procedia Computer Science, 2019, 157: 160-167.
- [32]Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.
- [33]Zhou P , Shi W , Tian J , et al. Attention-based bidirectional Long Short-Term Memory networks for relation classification[C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). USA: Association for Computational Linguistics.2016:207-212.
- [34]Buja A,Stuetzle W,Shen Y. Loss functions for binary class probability estimation and classification: Structure and applications[EB/OL]. <http://www-stat.wharton.upenn.edu/~buja/>,2020-07-12.

(通讯作者: 祁瑞华, ORCID: 0000-0002-2583-3055, E-mail:rhqi@dlufl.edu.cn。)

基金项目：本文系国家自然科学基金一般项目“典籍英译国外读者网上评论观点挖掘研究”（项目编号：15BYY028），辽宁省高等学校创新人才项目（项目编号：WR2019005）和辽宁省社科规划基金项目一般项目“大数据环境下突发事件谣言预警研究”（项目编号：L17BTQ005）的研究成果之一。

作者贡献声明：

祁瑞华：提出研究思路，设计研究方案，进行实验，论文起草，论文最终版本修订；

简悦：采集、清洗和分析数据，进行实验，论文起草；

郭旭：清洗和分析数据，进行实验；

关菁华：清洗和分析数据，进行实验；

杨明昕：清洗和分析数据，进行实验。

所有作者声明不存在利益冲突关系。

支撑数据：

支撑数据由作者自存储，E-mail: 734975095@qq.com。

[1] 祁瑞华,简悦.跨领域预训练词向量.

books + dvd.vec

books+ electronics.vec

books + kitchen.vec

books +video.vec

dvd + electronics.vec

dvd + kitchen.vec

dvd +video.vec

electronics + kitchen.vec

electronics +video.vec

kitchen+ video.vec

video + electronics.vec

[2] 祁瑞华,简悦.CDSA.csv.跨领域对照实验结果

作者简介：祁瑞华，女，1974 年生，博士，教授，硕士生导师，研究领域：自然语言处理和文本挖掘，Email: rhqi@dlufl.edu.cn；简悦，女，1997 年生，硕士研究生；郭旭，男，1983 年生，硕士，讲师；关菁华，女，1979 年生，博士，副教授；杨明昕，女，1997 年生，硕士研究生。