

基于评价对象类别的跨领域情感分类方法研究

张 慧 李寿山 李培峰 朱巧明

(江苏省计算机信息处理技术重点实验室 苏州 215006) (苏州大学计算机科学与技术学院 苏州 215006)

摘 要 情感分类任务具有领域相关性,即使用某一个领域的标注样本训练出的分类模型在对其他领域样本进行分类时性能表现往往会非常差。情感分类的跨领域学习旨在减少跨领域的性能损失。提出一种基于评价对象类别的跨领域学习方法。首先,将评价对象分为4大类:整体、硬件、软件和服务;然后,人工标注源领域中属于以上4类评价对象的句子,并构建评价对象类别分类器;最后,将不同的评价对象类别当作不同的视图,进而使用协同学习(Co-training)进行跨领域情感分类。实验结果表明,提出的方法有效地改进了跨领域学习性能。

关键词 评价对象,协同训练,最大熵,跨领域情感分类

中图分类号 TP391 文献标识码 A

Cross-domain Sentiment Classification with Opinion Target Categorization

ZHANG Hui LI Shou-shan LI Pei-feng ZHU Qiao-ming

(Provincial Key Lab of Computer Information Processing Technology of Jiangsu, Suzhou 215006, China)

(School of Computer Science and Technology, Soochow University, Suzhou 215006, China)

Abstract The task of sentiment classification is domain-specific, i. e., a classifier learning from the annotated data from a domain often performs dramatically badly on the data from a different domain. We presented a novel approach for cross-domain sentiment classification. Specifically, we first generalized four general categories of the opinion targets: overall, software, hardware, and service and classified all sentences into these categories. Then, some sentences with the category information were annotated in the source domain and a classifier for opinion target categorization was developed with the annotated data to classify all the sentences in both the source and target domain. Third, the four categories of opinion targets were considered as four different views which are employed in a standard co-training algorithm to perform cross-domain sentiment classification. Experimental results across several domains of Chinese reviews demonstrate the effectiveness of the proposed approach.

Keywords Opinion target, Co-training, Maximum entropy, Cross-domain sentiment classification

1 引言

随着互联网的快速发展,网络评论已经成为人们网络生活的重要组成部分,针对这些海量的评论数据,如何从中挖掘出用户的情感倾向等有价值的信息成为研究人员关注的课题。情感分类是挖掘这些重要信息的一个基本方法,即将待分析文本按照感情色彩(褒义或者贬义等)进行自动分类。目前广泛应用的情感分类方法是基于监督学习的方法,该方法以已标注的语料为基础,训练出一个分类器模型,然后对测试集进行测试。然而,这种方法仅适用于在单一领域的情感分类,即在一个领域中训练出的模型往往不能在另外一个领域有很好的表现。由于对现实中的每个领域都进行标注是一件代价非常昂贵的事情,因此,很多的研究者开始对跨领域的情感分类进行研究^[1-3]。

情感分类是一个领域相关问题。通常来讲,源领域的标注样本很难很好地直接应用到目标领域,造成领域间差异的一个很重要的原因在于每个领域的评价对象存在很大差异(例如,酒店领域的评价对象一般为设施、环境、服务、价格等;笔记本领域的评价对象一般为硬盘、显示器、操作系统等)。虽然不同的领域存在不同的评价对象,但是每个评价对象总能归属于某几个特定的类别(例如,硬件,软件等)。因此,我们可以通过基于评价类别的平滑策略为不同领域但属于同一类别的评价对象构建联系,从而达到领域适应的目的。例如,在各个领域中,不管是酒店的服务人员,还是产品领域的客服、物流等都属于服务这一类别。我们可以使用同一类别的评论信息来获得高精度的跨领域分类信息,帮助整体提高跨领域的分类性能。

具体来讲,本文基于上面的思想提出了一种新的跨领域

到稿日期:2012-03-13 返修日期:2012-06-09 本文受国家自然科学基金(60970056,61070123,61003155),高等学校博士学科点专项科研基金(20093201110006),模式识别国家重点实验室开发课题基金,江苏省自然科学基金(BK2011282),江苏省高校自然科学基金重大项目(11KJ520003),教育部科技发展中心网络时代的科技论文快速共享专项研究资助。

张 慧(1985-),女,硕士生,主要研究方向为自然语言处理、情感分析,E-mail: huihuizhang2009@gmail.com;李寿山(1980-),男,副教授,主要研究方向为自然语言处理、情感分析;李培峰(1971-),副教授,硕士生导师;朱巧明(1963-),男,教授,博士生导师。

学习方法。首先,将评价对象划分为4个大类,即:整体、硬件、软件、服务;在源领域中人工标注评价对象类别的观点句并构建评价对象类别分类器对目标领域的所有句子进行评价对象类别分类;最后,将评价对象类别当作不同的视图,使用协同训练(Co-training)的方法进行跨领域情感分类学习。与以往跨领域情感分类不同的是,文中提出的方法是第一个从评价对象的角度来解决跨领域情感分类。

本文第2节主要介绍跨领域情感分类及评价对象抽取方法的相关工作;第3节主要介绍评价对象类别的语料标注;第4节介绍所提方法在跨领域情感分类中的应用;第5节主要介绍和分析实验结果;最后进行总结。

2 相关工作

2.1 跨领域情感分类相关研究

由于在不同的领域中,表达情感的方式往往差异很大,因此情感分类问题是一个领域非常相关的问题,即在一个领域中训练出来的分类器往往不适合直接应用于另外一个领域。除此之外,因为分类模型的训练需要大量的标注数据,而标注数据是非常耗时耗力的,所以很多研究者开始转向跨领域情感分析的研究。Aue等^[1]在4个不同的数据集上进行了多领域对一个领域的跨领域情感分析,文中提出了4种多领域情感分类算法,并对4种算法在不同领域之间的分类效果上作了比较和分析。Wu等^[2]将文本的情感倾向性与图排序算法结合起来,在图排序算法基础上,利用训练样本的准确标签与测试样本的伪标签来迭代进行倾向性分析。Blitzer等^[3]基于互信息进行枢轴选择对SCL(Structural Correspondence Learning)进行改进,并探讨了不同领域的相识度来选择合适的训练语料。Pan等^[4]提出了特征对齐算法SFA(Spectral Feature Alignment),即用领域无关的词作为桥梁,将不同领域的领域相关的词对齐到统一的集簇,从而找到不同领域特征之间的映射关系。Li等^[5]提出使用组合分类器的方法,即通过多个分类器的融合来实现多领域的情感分类。

2.2 评价对象标注与应用相关研究

评价对象抽取的任务是给定一个句子,从中抽取出其的评价对象,其是细粒度的情感分析任务^[6]。如Hu等^[7]设计了一个用来挖掘产品评价对象并对用户评论进行总结的系统,文中主要采用了关联挖掘的方法从用户评论中找到被评论的产品特征,并对其倾向性进行判断,最后给出一个基于特征的总结。Zhuang等^[8]在IMDb的电影评论语料的基础上,抓取了20个不同的电影,每个电影100条评论,以句子为标注单位,对含有评价对象-观点对的句子进行标注,并提出了用依存语法图抽取评价对象-观点词的方法。在Zhuang的基础上,Jakob等^[9]在DSRC语料上(其中包括services和universities两个领域共234和256条评论)进行了句子级以及表达级别标注,在此标注语料基础上,提出了基于CRF模型的抽取评价对象方法,并且分别应用于单领域与跨领域中。

3 评价对象类别的定义与检测

3.1 4类评价对象的定义

在用户的评论中存在很多的评价对象,并且各个领域中的评价对象是不尽相同的,所以将某一特定的评价对象作为多领域中通用的评价对象是不可能的。为了解决过细粒度化

带来的问题,提出了将评价对象划分为某几个大类。图1为将评价对象分为几个类别的框架图,即,首先将评价对象分为整体和非整体两大类。整体描述的是针对某一产品的总体评价;而与之相对应的非整体则描述的是产品的某些属性、组成部分或者一些其他的间接评价对象。然后,产品的属性和组成部分又分为软件与硬件两大类;产品相关的评价对象又分为与产品的附属服务和其他。总之,将评价对象分为以下4大类:整体、软件、硬件和服务。除了以上定义的几类评价对象,把不属于以上任意类别的一类统一定义为其他类。

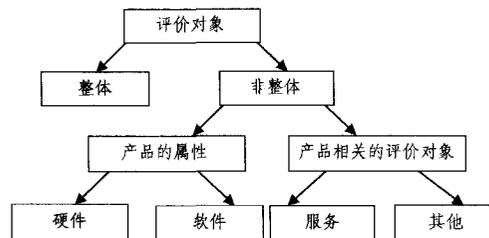


图1 评价对象分类层次图

1. 整体(Overall)

这类评价对象一般是指对产品等的总体评价,不会细化到哪一属性或者组成部分。例如:

- E1: 总体感觉还不错。
- E2: 总之令我很失望。
- E3: 整体满意。

在评价这类评价对象时,一般有常见的特征。往往出现一些指示词如“总体”、“总之”、“整体”等表示总结的指示词。

2. 硬件(Hardware)

这类评价对象一般指比较具体的产品的属性或者某一组成部分,硬件与软件的不同主要表现在硬件具有具体的形状等特征,是能够看得见、摸得着的实体。例如:

- E4: 设施太陈旧。
- E5: USB接口太少。
- E6: 就发现键盘已经坏了。

在评价这类评价对象时,常常使用一些描述实体属性的形容词,例如“太少”、“太大”等,同时行为描述也是评价这类评价对象的一种特殊方式,例如,E6中的“坏了”。

3. 软件(Software)

这类评价对象一般指产品比较抽象的某一属性或者部分。跟硬件对比,它不具有具体的大小、形状等,是客观存在但看不见、摸不着的实体。

- E7: 价格特贵。
- E8: 性价比高。
- E9: 自带系统太臭了。

在评价这一类评价对象时,常常使用一些描述抽象性的形容词,例如“高”、“贵”等。评价的对象往往也是比较抽象的物体,例如“价格”、“性价比”等。

4. 服务(Service)

这类评价对象的划分标准:是否与人打交道,主要是指消费者在购买产品等时所得到的包括客服、物流等服务。例如:

- E10: 前台服务很周到。
- E11: 服务态度差。
- E12: 客服挺耐心。

在评价这一类评价对象时,总会使用关于“服务”这一独特评价对象的评价词,例如“周到”、“耐心”。

文中分别对酒店和笔记本两个领域的语料进行标注,标注规模为正、负各 300 篇评论。并对这两个领域中的各个评价对象类别的分布作了统计,如表 1 所列。

表 1 各类评价对象统计结果表

	整体	软件	硬件	服务	其他
酒店	630	1166	983	644	3311
笔记本	193	1196	541	154	1588

3.2 评价对象类别的检测

3.2.1 面向评价对象类别的多类分类方法

根据 3.1 节的定义,将评价对象分为了 4 大类,所以对评价对象的检测就要用到多分类器。一般情况下,各个分类器如支持向量机、最大熵最开始都是被设计用来解决二分类问题的。目前,解决多分类问题主要有两种方法:第一种方法是将多分类器看作二分类的组合,将多分类的问题转化为多个二分类问题;第二种方法则是通过修改目标函数,从根本上解决多分类问题。对于后者代价过高,只适合小规模问题,所以大多数采用第一类方法。

3.2.2 基于欠采样的样本不平衡分类方法

在经过对标注语料后的统计显示,各个类别的样本分布比较不平衡。因为不进行处理直接应用到分类器,会影响分类器的分类精度,所以对不平衡现象进行处理非常有必要。本文采用欠采样方法处理类别中的样本不平衡问题。

欠采样方法的思想是减少多数类的样本,从而降低多数类的特征对分类器的影响。将多数类的样本减少到与少数类的样本一致。

3.2.3 评价对象类别自动分类算法流程

评价对象检测算法流程图如图 2 所示。

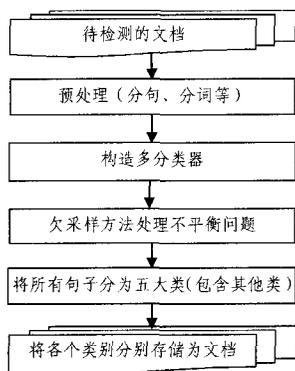


图 2 评价对象检测算法流程图

4 基于评价对象类别的跨领域情感分类

跨领域情感分类主要是应用源领域中的训练数据训练一个分类器模型,通过逐步加入目标领域的样本来实现不同领域间的情感分类,然而通过什么样的标准加入怎样的样本是研究的关键。本章主要提出了将评价对象分为 4 大类,并将每类作为一个视图,然后采用协同训练的方法在每个视图中加入置信度比较高的样本,通过这种方法来进行跨领域的情感分类。图 3 详细说明了基于评价对象分类的跨领域情感分类的流程。

输入:

源领域:标注的数据集 L ,包括属于整体类别的句子 $S_{L-overall}$,属于硬件类别的句子 $S_{L-hardware}$,属于软件类别的句子 $S_{L-software}$,属于服务类别的句子 $S_{L-service}$ 。

目标领域:未标注的数据集 U ,包括属于整体类别的句子 $S_{U-overall}$,属于硬件类别的句子 $S_{U-hardware}$,属于软件类别的句子 $S_{U-software}$,属于服务类别的句子 $S_{U-service}$ 。

输出:

更新的标注数据集 L 。

流程:

循环迭代 N 次直到 $U=\emptyset$ 。

- (1)使用 $S_{L-overall}$ 训练第一个分类器 f_1 。
- (2)使用分类器 f_1 标注未标注 U 中的 $S_{U-overall}$ 。
- (3)分别在正、负文本中选择 n_1 个置信度高的样本 A_1 。
- (4)使用 $S_{L-hardware}$ 训练第二个分类器 f_2 。
- (5)使用分类器 f_2 标注未标注集 U 中的 $S_{U-hardware}$ 。
- (6)分别在正、负文本中选择 n_2 个置信度高的样本 A_2 。
- (7)使用 $S_{L-software}$ 训练第三个分类器 f_3 。
- (8)使用分类器 f_3 标注未标注集 U 中的 $S_{U-software}$ 。
- (9)分别在正、负文本中选择 n_3 个置信度高的样本 A_3 。
- (10)使用 $S_{L-service}$ 训练第三个分类器 f_4 。
- (11)使用分类器 f_4 标注未标注集 U 中的 $S_{U-service}$ 。
- (12)分别在正、负文本中选择 n_4 个置信度高的样本 A_4 。
- (13)将 $A_1 \cup A_2 \cup A_3 \cup A_4$ 加入到标注数据集 L 。
- (14)更新 $S_{L-overall}$ 、 $S_{L-hardware}$ 、 $S_{L-software}$ 和 $S_{L-service}$ 。

图 3 基于评价对象类别的跨领域情感分类的流程图

5 实验结果分析

本节首先介绍了人工标注评价对象类别的一致性分析,然后说明了评价对象类别探测的实验结果,最后介绍了基于提出的在评价对象类别基础上进行跨领域情感分类的性能以及与其他跨领域情感分类方法性能的比较。

5.1 实验设置

在中文 4 个领域:酒店(Hotel)、笔记本(Notebook)、家具(Furniture)、数码相机(Digital camera)进行相应的实验。将酒店和笔记本领域分别作为源领域,其他的两个领域作为目标领域。实验中采用的数据规模为:源领域正、负各 1000 篇作为初始训练集,目标领域正、负各 800 篇作为未标注样本,目标领域的正、负各 200 篇作为测试集。

在实验中采用最大熵分类器,使用 Mallet(<http://mallet.cs.umass.edu/>)工具包,所有参数全部都采用默认值。在进行分类之前,首先采用中国科学院计算技术研究所的分词软件 ICTCLAS(<http://ictclas.org/>)对中文文本进行分词操作。给定分好词的文本后,选取词的一元特征,用以获得文本向量的表示^[10]。

5.2 评价对象类别标注一致性分析

我们人工标注来自酒店和笔记本两个领域正、负各 300 篇评论中的所有句子,将每一个句子标注为以下 4 种评价对象类型:整体、硬件、软件和服务。在标注过程中,为了评价标注体系的有效性,抽取来自酒店领域的 80 篇评论让两名标注者对其进行标注并使用 Cohen's kappa(k)^[11] 进行一致性分析。表 2 给出了各个评价对象类别的句子数目及标注评价对象类别的一致性结果。从一致性结果可以看出,提出的评价对象分类体系比较合理,能够达到一定的一致程度。

表2 人工标注各评价对象一致性分析表

类别数	标注的句子数	k
Overall	41	0.65
Hardware	91	0.71
Software	102	0.72
Service	51	0.77

5.3 评价对象类别分类结果分析

如表1所列,各个评价对象类别里的样本数量是不平衡的,例如,在酒店领域,属于服务类别的句子有644句,而属于软件类别的句子有1166句,差不多是服务领域的两倍,如果对这种不平衡现象不处理就会严重影响到分类器的分类精度。所以本文主要使用欠采样(Under-sampling)的方法处理此不平衡问题。表3分别给出不使用任何处理不平衡方法(Full-Training)和采用欠采样方法(Under-sampling)进行分类的结果。

表3 采用欠采样方法处理不平衡的评价对象分类性能表

	Full-Training	Under-sampling
Hotel	0.660	0.689
Notebook	0.530	0.630
Hotel→ Notebook	0.364	0.402
Notebook →Hotel	0.360	0.401

表3中,第2行、第3行表示训练集和测试集都来自同一领域的实验结果,第4行、第5行表示训练集和测试集来自不同领域的实验结果。可以看出,一方面,训练集和测试集来自同一领域的实验结果明显好于跨领域的实验结果;另一方面,如果将不平衡的各类别数据直接训练分类器效果是比较差的(Full-Training),而使用欠采样方法后,分类的性能得到改善。因此在评价对象对类别分类中,采用的是欠采样方法。

表4给出了使用欠采样后,每个类别的分类详细结果。从表4可以看出,总体上,酒店领域的准确率、召回率和F值都要高于笔记本领域,究其原因可能是相对于酒店领域,笔记本领域的每个句子都比较短,特征空间不如酒店领域的充实,导致分类结果不如酒店领域。

表4 使用欠采样后的各类评价对象分类结果表

	Hotel			Notebook		
	P	R	F	P	R	F
Overall	0.69	0.81	0.75	0.57	0.68	0.59
Software	0.67	0.59	0.63	0.63	0.63	0.63
Hardware	0.69	0.67	0.68	0.63	0.60	0.61
Service	0.80	0.74	0.77	0.75	0.55	0.63
Other	0.56	0.60	0.58	0.41	0.38	0.39

在各个评价对象类别识别中,“服务”类别的准确率、召回率和F值均都较高,该类别的信息比较容易识别。

5.4 基于评价对象类别的跨领域情感分类结果分析

本节主要介绍了使用4种方法进行跨领域情感分类的实验结果。

基准系统(Baseline):直接利用源领域的语料训练分类器,不加入任何目标领域的样本,然后将目标领域的样本作为测试集进行测试。

自学习方法(Self-training):在基准系统的基础上,通过自学习方法,每次加入置信度较高的目标领域的样本,此方法只有一个视图,不能充分利用评论中所表现出的各个侧面的信息。

随机子空间的协同学习(Co-training with Random Fea-

ture Subspaces, CRFS):将所有的特征随机分为4个子空间,将每个子空间作为一个视图,分别训练分类器,通过协同学习的方法加入目标领域中置信度较高的样本。Ganchev等^[12]的实验结果表明:在情感分类任务中,基于随机子空间的协同学习方法往往优于SCL方法。因此,我们将这种方法作为目前跨领域情感分类的最好的方法之一进行比较。

基于评价对象类别的协同学习(Co-training with opinion target categories, COTC):将评论中的评价对象分为整体、硬件、软件和服务这4大类,并将每一类别作为一个视图,利用协同训练的方法将每一类别中置信度高的目标领域的样本加入训练集,通过迭代得到最佳性能。

下面通过图4—图11,给出了4种跨领域情感分类方法的性能比较。

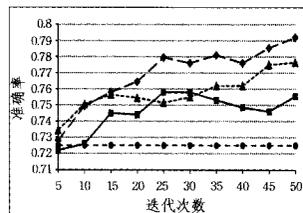


图4 酒店—>箱包领域各种跨领域方法性能比较

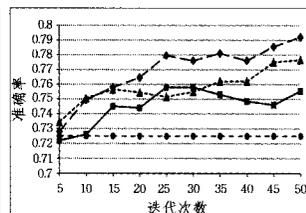


图5 酒店—>家具领域各种跨领域方法性能比较

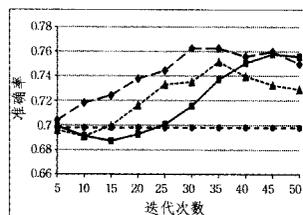


图6 酒店—>数码相机领域各种跨领域方法性能比较

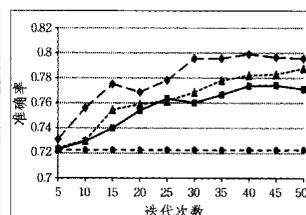


图7 酒店—>笔记本领域各种跨领域方法性能比较

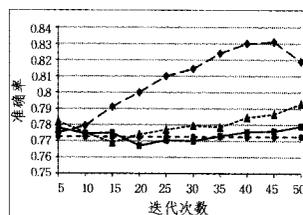


图8 笔记本—>箱包领域各种跨领域方法性能比较

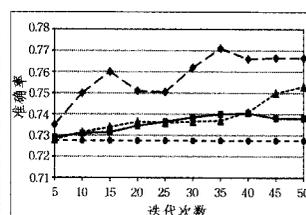


图9 笔记本—>家具领域各种跨领域方法性能比较

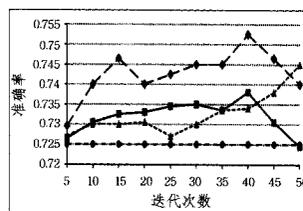


图10 笔记本—>数码相机领域各种跨领域方法性能比较

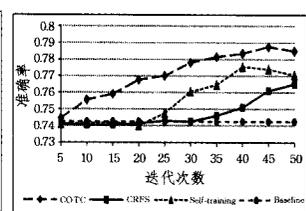


图11 笔记本—>酒店领域各种跨领域方法性能比较

通过图4—图11可以看出,在源领域是酒店,目标领域分别是箱包、家具、数码相机和笔记本这一组实验中,除了目标领域是数码相机这一组实验以外,其他组中的基于评价对象类别的方法明显好于随机子空间和自学习方法。例如,在

参考文献

- [1] Shafer G. A mathematical theory of evidence[M]. Princeton, Princeton University Press, 1976
- [2] Zadeh L A. Rview of Shafer's a mathematical theory of evidence [J]. AI Magaine, 1984, 5(3): 81-83
- [3] Smarandache F, Dezert J. Applications and Advances of DS_mT for Information Fusion [M]. Rehoboth: American Research Press, 2009
- [4] Yager R R. Comparing approximate reasoning and probabilistic reasoning using the Dempster-shafer framework [J]. International Journal of Approximate Reasoning, 2009, 50(5): 812-821
- [5] 孙全, 叶秀清, 顾伟康. 一种新的基于证据理论的合成公式[J]. 电子学报, 2000, 28(08): 117-119
- [6] 张山鹰, 潘泉, 张洪才. 证据推理冲突问题研究[J]. 航空学报, 2001, 22(4): 369-372
- [7] 邓勇, 施文康. 一种改进的证据推理组合规则[J]. 上海交通大学学报, 2003, 37(8): 1275-1278
- [8] 韩德强, 韩崇昭, 邓勇, 等. 基于证据方差的加权证据融合[J]. 电子学报, 2011, 34(3): 153-157
- [9] 胡丽芳, 关欣, 邓勇, 等. 广义幂集空间中证据冲突的原因分析

- [J]. 控制理论与应用, 2011, 28(12): 1717-1722
- [10] Deng Yong, Shi Wen-kang, Zhu Zhen-fu, et al. Combining belief functions based on distance of evidence [J]. Decision Support Systems, 2004, 38(3): 489-493
- [11] Murphy C K. Combining belief functions when evidence conflicts [J]. Decision Support Systems, 2000, 29(1): 1-9
- [12] Haenni R. Are alternatives to Dempster's rule of combination real alternative? Comments on about the belief function combination and the conflict management problem [J]. Information Fusion, 2002, 3(4): 237-239
- [13] 张捍东, 王翠华, 强克坤. 基于焦元支持度的合成规则[J]. 控制理论与应用, 2011, 28(5): 741-744
- [14] 邓勇, 蒋雯, 韩德强. 广义证据理论的基本框架子[J]. 西安交通大学学报, 2010, 44(12): 119-124
- [15] Jousselme A L, Liu Chun-sheng, Grenier D, et al. Measuring ambiguity in the evidence theory [J]. IEEE Transactions on Systems, Man and Cybernetics, Part A, 2006, 36(5): 890-903
- [16] Klement E P, Mesiar R, Endre P, et al. Triangular norms, Position Paper I: Basic analytical and algebraic properties [J]. Fuzzy Sets Systems, 2004, 143: 5-26

(上接第 232 页)

目标领域是箱包时,本文的方法比自学习方法平均提高 2.34 个百分点,比随机子空间协同学习方法提高 1.12 个百分点,比基准系统高 4.4 个百分点。总体来看,在 8 组实验中,我们的方法表现出了较好的性能,同时验证了从一个新的角度-评价对象类别来进行跨领域情感分析是有效的。

为了进一步体现本文方法的优势,还将我们的方法同跨领域情感分类中比较流行的 SCL^[3]方法进行了比较研究。值得一提的是, SCL 在很大程度上依赖于枢轴的选择,合适的枢轴才能取得最佳效果。因此,通过不断地调整参数,使得 SCL 方法取得最佳性能。表 5 给出了本文提出的方法与 SCL 方法的性能对比结果。

表 5 各种跨领域情感分类方法比较

源领域→目标领域	COTC	SCL	Baseline
酒店→笔记本	0.795	0.732	0.722
酒店→家具	0.712	0.680	0.662
酒店→箱包	0.792	0.739	0.725
酒店→数码相机	0.749	0.732	0.697
笔记本→酒店	0.785	0.765	0.742
笔记本→家具	0.766	0.740	0.727
笔记本→箱包	0.819	0.797	0.782
笔记本→数码相机	0.740	0.740	0.725

根据表 5 可以看出,在源领域是酒店,目标领域是其他领域时,基于评价对象类别的方法平均比 SCL 方法高出 4.12 个百分点,在源领域是笔记本,目标领域是其他领域时,比 SCL 方法高出 1.68 个百分点。此实验结果再次验证了基于评价对象类别进行跨领域研究的有效性。

结束语 本文主要研究了从评价对象类别的角度进行跨领域情感分类。首先人工标注了各种评价对象类别,包括整体、硬件、软件和服务类别。然后,在源领域利用以上 4 类评价对象构建分类器,通过将不同评价对象类别当作不同的视图加入目标领域的样本来实现跨领域情感分类。实验结果表明,提出的基于评价对象类别的方法对跨领域情感分类的性

能有显著提高。

参考文献

- [1] Aue A, Gamon M. Customizing Sentiment Classifiers to New Domains: A Case Study [C] // Proceeding of RANLP. 2005
- [2] 吴琼, 谭松波. 跨领域倾向性分析相关技术研究 [J]. 中文信息学报, 2010, 24(1): 77-83
- [3] Blitzer J, Dredze M, Pereira F. Biographies, Bollywood, Boomboxes and Blenders: Domain Adaptation for Sentiment Classification [C] // Proceedings of ACL. 2007: 432-439
- [4] Pan S J, Ni X C, Sun J T, et al. Cross-domain Sentiment Classification via Spectral Feature Alignment [C] // Proceedings of WWW. 2010: 751-760
- [5] Li S, Huang C, Zong C. Multi-domain Sentiment Classification with Classifier Combination [J]. Journal of Computer Science and Technology (JCST), 2011, 26(1): 25-33
- [6] 候锋, 王传廷, 李国辉. 网络意见挖掘、摘要与检索研究综述 [J]. 计算机科学, 2009, 36(7): 15-19
- [7] Hu M, Liu B. Mining and Summarizing Customer Reviews [C] // Proceedings of KDD. 2004: 168-177
- [8] Zhuang L, Jing F, Zhu X. Movie Review Mining and Summarization [C] // Proceedings of CIKM. 2006: 43-50
- [9] Jakob N, Gurevych I. Extracting Opinion Targets in a Single and Cross-Domain Setting with Conditional Random Fields [C] // Proceedings of EMNLP. 2010: 1035-1045
- [10] 宗成庆. 统计自然语言处理 [M]. 北京: 清华大学出版社, 2008: 341
- [11] Jacob C. A Coefficient of Agreement for Nominal Scales [J]. Educational and Psychological Measurement, 1960, 20(1): 37-46
- [12] Ganchev K, Graça J, Blitzer J, et al. Multi-View Learning over Structured and Non-Identical Outputs [C] // Proceedings of UAI. 2008