

doi:10.1631/FITEE.1601679

题目：基于关联规则进行词对齐的跨领域情感分类算法

概要：文本情感分类被应用于企业用户管理系统，通过自动对诸如评论、博客等带有情感倾向性文字进行分析，帮助商家更好地了解用户对商品或者服务的态度。然而，评论和博客等内容常源于不同应用领域，为每个领域训练一个能准确预测情感倾向的分类器非常困难。主要原因是，在不同领域，人们通常会用不同特征词表达相同情感，并且难以找到一个直接的映射函数，以建立不同领域特征词间的映射关系，从而消除领域间差异。因此，将某个领域训练好的分类器直接应用到另一个领域时，会因为领域间差异使得分类器准确率急速下降。本文提出一个新的基于关联规则进行特征词对齐的跨领域情感分类算法，该算法通过在同一领域中挖掘具有强关联关系的领域共享词和领域专有词词对，建立直接映射关系，并以领域共享词为桥梁，在不同领域的特征专有词之间建立间接映射关系，从而在一定程度上消除了源领域和目标领域之间的差异，有效提升了跨领域情感分类准确率。在亚马逊数据库上的实验结果证明该算法提高了跨领域情感分类性能。

关键词：情感分类；跨领域；关联规则