

## 迁移学习研究进展\*

庄福振, 罗平, 何清, 史忠植

(中国科学院智能信息处理重点实验室(中国科学院 计算技术研究所), 北京 100190)

通讯作者: 庄福振, E-mail: zhuangfz@ics.ict.ac.cn

**摘要:** 近年来, 迁移学习已经引起了广泛的关注和研究. 迁移学习是运用已存有的知识对不同但相关领域问题进行求解的一种新的机器学习方法. 它放宽了传统机器学习中的两个基本假设: (1) 用于学习的训练样本与新的测试样本满足独立同分布的条件; (2) 必须有足够可利用的训练样本才能学习得到一个好的分类模型. 目的是迁移已有的知识来解决目标领域中仅有少量有标签样本数据甚至没有的学习问题. 对迁移学习算法的研究以及相关理论研究的进展进行了综述, 并介绍了在该领域所做的研究工作, 特别是利用生成模型在概念层面建立迁移学习模型. 最后介绍了迁移学习在文本分类、协同过滤等方面的应用工作, 并指出了迁移学习下一步可能的研究方向.

**关键词:** 迁移学习; 相关领域; 独立同分布; 生成模型; 概念学习

**中图法分类号:** TP181

中文引用格式: 庄福振, 罗平, 何清, 史忠植. 迁移学习研究进展. 软件学报, 2015, 26(1): 26-39. <http://www.jos.org.cn/1000-9825/4631.htm>

英文引用格式: Zhuang FZ, Luo P, He Q, Shi ZZ. Survey on transfer learning research. Ruan Jian Xue Bao/Journal of Software, 2015, 26(1): 26-39 (in Chinese). <http://www.jos.org.cn/1000-9825/4631.htm>

### Survey on Transfer Learning Research

ZHUANG Fu-Zhen, LUO Ping, HE Qing, SHI Zhong-Zhi

(Key Laboratory of Intelligent Information Processing of Chinese Academy of Sciences, Institute of Computing Technology, The Chinese Academy of Sciences, Beijing 100190, China)

**Abstract:** In recent years, transfer learning has provoked vast amount of attention and research. Transfer learning is a new machine learning method that applies the knowledge from related but different domains to target domains. It relaxes the two basic assumptions in traditional machine learning: (1) the training (also referred as source domain) and test data (also referred target domain) follow the independent and identically distributed (i.i.d.) condition; (2) there are enough labeled samples to learn a good classification model, aiming to solve the problems that there are few or even not any labeled data in target domains. This paper surveys the research progress of transfer learning and introduces its own works, especially the ones in building transfer learning models by applying generative model on the concept level. Finally, the paper introduces the applications of transfer learning, such as text classification and collaborative filtering, and further suggests the future research direction of transfer learning.

**Key words:** transfer learning; related domain; independent and identical distribution; generative model; concept learning

随着社会发展的信息化和网络化, 人们在日常生活和工作中无时无刻不在获取信息, 分析信息, 并以此作为决策的依据. 在一定程度上, 信息的拥有量已经成为决定和制约人类社会发展的的重要因素. 想要高效、准确地寻找到所需的信息, 信息分类是必不可少的第一步. 通过分类, 信息可以得到有效的组织管理, 有利于快速、准确地

\* 基金项目: 国家自然科学基金(61473273, 61473274, 61175052, 61203297); 国家高技术研究发展计划(863)(2014AA015105, 2013AA01A606, 2012AA011003)

收稿时间: 2014-02-21; 修改时间: 2014-03-31; 定稿时间: 2014-05-22; jos 在线出版时间: 2014-07-14

CNKI 网络优先出版: 2014-07-14 16:04, <http://www.cnki.net/kcms/doi/10.13328/j.cnki.jos.004631.html>

定位信息.分类学习问题,是机器学习中一种重要的学习方法,目前已经得到广泛的研究与发展.

在传统分类学习中,为了保证训练得到的分类模型具有准确性和高可靠性,都有两个基本的假设:(1) 用于学习的训练样本与新的测试样本满足独立同分布的条件;(2) 必须有足够可利用的训练样本才能学习得到一个好的分类模型.但是,在实际应用中我们发现,这两个条件往往无法满足.首先,随着时间的推移,原先可利用的有标签的样本数据可能变得不可用,与新来的测试样本的分布产生语义、分布上的缺口.比如,股票数据就是很有时效性的数据,利用上月份的训练样本学习得到的模型并不能很好地预测本月份的新样本.另外,有标签的样本数据往往很匮乏,而且很难获得.在 Web 数据挖掘领域,新数据不断涌现,已有的训练样本已经不足以训练得到一个可靠的分类模型,而标注大量的样本又非常费时费力,而且由于人的主观因素容易出错,这就引起了机器学习中另外一个重要问题,如何利用少量的有标签训练样本或者源领域数据,建立一个可靠的模型对目标领域数据进行预测(源领域数据和目标领域数据可以不具有相同的数据分布).He 等人<sup>[1]</sup>指出数据分类首先要解决训练集样本抽样问题,如何抽到具有代表性的样本集作为训练集是一个值得研究的重要问题.文献[1]提出极小样本集抽样方法用于基于超曲面分类算法,该方法可感知非结构化数据的分布,并以极小样本集作为代表子集.该文还指出了极小样本集有多少种表达方式,给出了样本缺失情况下准确率的精确估计.文献[1]表明,在实际中保证训练得到的分类模型具有高准确性和可靠性的两个基本假设并不是每种算法都能做到的,因此研究迁移学习变得非常重要.

近年来,迁移学习已经引起了广泛的关注和研究<sup>[2-18]</sup>.根据维基百科的定义<sup>\*\*</sup>,迁移学习是运用已存有的知识对不同但相关领域问题进行求解的一种新的机器学习方法.它放宽了传统机器学习中的两个基本假设,目的是迁移已有的知识来解决目标领域中仅有少量有标签样本数据甚至没有的学习问题.迁移学习广泛存在于人类的活动中,两个不同的领域共享的因素越多,迁移学习就越容易,否则就越困难,甚至出现“负迁移”<sup>[19,20]</sup>的情况,产生副作用.比如:一个人要是学会了骑自行车,那他就很容易学会开摩托车;一个人要是熟悉五子棋,也可以轻松地将知识迁移到学习围棋中.但是有时候看起来很相似的事情,却有可能产生“负迁移”现象.比如,学会骑自行车的人来学习三轮车反而不适应,因为这两种车型的重心位置不同<sup>[21,22]</sup>.近几年来,已经有相当多的研究者投入到迁移学习领域中,每年在机器学习和数据挖掘的顶级会议中都有关于迁移学习的文章发表,比如,ICML, SIGKDD, NIPS, ICDM 以及 CIKM 等.下面针对迁移学习研究展开系统的综述,特别是从迁移学习采用的技术角度进行介绍.

## 1 迁移学习算法研究进展

针对源领域和目标领域样本是否标注以及任务是否相同<sup>[13]</sup>,可以把以往迁移学习工作划分为归纳迁移学习、直推式迁移学习以及无监督迁移学习等.而按照迁移学习方法采用的技术划分,又可以把迁移学习方法大体上分为:i) 基于特征选择的迁移学习算法研究;ii) 基于特征映射的迁移学习算法研究;iii) 基于权重的迁移学习算法研究.本文主要从这两条线对迁移学习的工作进行总结和综述.

### 1.1 按源领域和目标领域样本是否标注以及任务是否相同划分

Pan 和 Yang<sup>[13]</sup>针对源领域和目标领域样本是否标注及任务是否相同或是否单一对迁移学习进行了划分.

根据源领域和目标领域中是否有标签样本可将迁移学习划分为 3 类:目标领域中有少量标注样本的归纳迁移学习(inductive transfer learning)<sup>[23]</sup>,只有源领域中有标签样本的直推式迁移学习(transductive transfer learning)<sup>[4]</sup>,以及源领域和目标领域都没有标签样本的无监督迁移学习<sup>[24,25]</sup>.另外,还根据源领域中是否有标签样本把归纳迁移学习划分成多任务学习、自学习.Pan 和 Yang<sup>[13]</sup>还给出了传统机器学习与各种迁移学习情形之间的关系,以及各种情形下,源领域与目标领域是否相同、源领域与目标领域的任务是否相同.迁移学习是和传统学习相对应的一大类学习方式,传统学习处理源领域和目标领域相同且源领域和目标领域的任务是相同的学习,迁移学习处理除此情形之外的学习,包括:源领域和目标领域的任务相关但不同的归纳迁移学习<sup>[6,12,26-31]</sup>,

\*\* [http://en.wikipedia.org/wiki/Transfer\\_learning](http://en.wikipedia.org/wiki/Transfer_learning)

源领域和目标领域相关但不相同而源领域和目标领域的任务相同的直推式迁移学习(transductive transfer learning)<sup>[3,7,32-36]</sup>.无监督迁移学习与归纳迁移学习类似,不过,无监督迁移学习主要处理源领域和目标领域都没有标签数据的问题<sup>[24,37]</sup>.Pan 和 Yang 还根据训练样本和测试样本是否来自于同一个领域,把直推式迁移学习划分为样本选择偏差、协方差偏移和领域自适应学习这些相关的子领域.

## 1.2 按采用的技术划分

近 10 年来,很多学者对迁移学习展开了广泛的研究,而且很多集中在算法研究上,即采用不同的技术对迁移学习算法展开研究.因此,下面首先介绍与迁移学习极其相关的半监督学习方法,然后再对采用各种技术的迁移学习工作进行介绍.

### 1.2.1 半监督学习方法

在传统的监督学习中,学习算法通过对大量有标签的训练样本进行学习,从而建立模型用于预测标记新来的没有标签的测试样本.但是随着信息技术、互联网以及存储技术的快速发展,数据量随之呈指数级增长.人们能够比较容易地收集大量的没有标签的数据,但要获取大量有标签的数据则较为困难,因为这可能需要耗费大量的人力物力.例如,在生物学中进行数据分类,得到一个训练样本的标签往往需要大量的、长时间的、昂贵的实验;在进行 Web 网页推荐时,用户也不愿意花费大量的时间来标记哪些网页是他感兴趣的,因此有标签的网页很少.实际上,在真实世界中通常存在大量的无标签的数据,而有标签的数据则较少.这就需要一种机器学习技术能够利用大量的无标签样本数据以及少量有标签的训练样本进行学习,以提高分类任务的准确率.

按照 Zhou 等人在文献[38]中的阐述,目前能够利用少量有标签数据和大量没有标签样本数据的技术有 3 类:半监督学习(semi-supervised learning)、直推式学习(transductive learning)和主动学习(active learning).这些学习方法都通过大量的无标签样本来辅助少量有标签样本的学习,但它们在思想上又有些不同.半监督学习指的是学习算法在学习过程中无需人工干预,基于自身对无标签数据加以利用.而直推式学习与半监督学习一样也无需人工干预,所不同的是,直推式学习假设无标签的数据就是最终要用来测试的数据,学习的目的就是在这些数据上取得最佳泛化能力.相对应地,半监督学习在学习时并不知道最终的测试用例是什么.因此,半监督学习考虑的是一个“开放的世界”,即在学习中不知道测试样本是什么,而直推式学习考虑的则是一个“封闭世界”,要测试的样本数据已参与到学习过程中.如果抛开是否对未知样本进行预测,直推式学习可以归结为半监督学习的一种特例.主动学习与半监督学习、直推式学习最大的区别在于它的学习过程需要人工干预,就是在学习过程中通过反馈尽可能地找到那些包含信息量大的样本来辅助少量有标签样本的学习.在传统机器学习中,这 3 种方法已经得到了广泛应用<sup>[39-44]</sup>.多视角学习(multi-view learning)也是半监督学习一个很重要的学习任务.Yarowsky<sup>[45]</sup>和 Blum 等人<sup>[46]</sup>认为数据的多视角表示方式可以提高半监督分类学习算法的性能.更进一步地,文献[47-49]用 PAC(probability approximately correct)理论分析了联合训练(co-training)在无标签数据上错误率的上界.

近年来也有很多研究者把这些技术应用到迁移学习领域.文献[22]对主动迁移学习模型进行了研究. Shi 等人<sup>[50]</sup>提出了一种跨领域的主动迁移学习方法,通过似然偏置的大小来选择领域外(out-of domain)有标签的样本.那些能够正确预测领域内(in-domain)数据且高似然偏置的有标签样本被利用,而那些低偏置的样本则通过主动学习进行选择.Liao 等人<sup>[6]</sup>提出了一种方法,即估计源领域中的每个样本与目标领域中少量标签数据之间的不匹配程度,并把该信息应用到逻辑回归中.Zhuang 等人<sup>[17]</sup>综合半监督学习的 3 种正则化技术,流形正则化<sup>[51]</sup>、熵正则化<sup>[52]</sup>以及期望正则化<sup>[53]</sup>,提出基于混合正则化的迁移学习方法.该方法首先从源领域训练得到一个分类器,然后通过混合正则化在目标领域数据上进行优化.

自学习(self-taught learning)<sup>[24,54]</sup>也是一种利用大量无标签数据来提高给定分类聚类任务性能的方法,自学习被应用于迁移学习中,因为它不要求无标签数据的分布与目标领域中的数据分布相同.Raina 等人<sup>[54]</sup>提出了一种自学习的方法,它利用稀疏编码技术对无标签的样本数据构造高层特征,然后少量有标签的数据以及目标领域无标签的样本数据都由这些简洁的高层特征表示.实验结果表明,这种方法可以极大地提高分类任务的准确率.

### 1.2.2 基于特征选择方法

基于特征选择的迁移学习方法是识别出源领域与目标领域中共有的特征表示,然后利用这些特征进行知识迁移<sup>[4,55,56]</sup>。Jiang 等人<sup>[55]</sup>认为,与样本类别高度相关的那些特征应该在训练得到的模型中被赋予更高的权重,因此他们在领域适应问题中提出了一种两阶段的特征选择框架。第 1 阶段首先选出所有领域(包括源领域和目标领域)共有的特征来训练一个通用的分类器;然后从目标领域无标签样本中选择特有特征来对通用分类器进行精化从而得到适合于目标领域数据的分类器。Dai 等人<sup>[4]</sup>提出了一种基于联合聚类(co-clustering)的预测领域外文档的分类方法 CoCC,该方法通过对类别和特征进行同步聚类,实现知识与类别标签的迁移。CoCC 算法的关键思想是识别出领域内(也称为目标领域)与领域外(也称为源领域)数据共有的部分,即共有的词特征。然后类别信息以及知识通过这些共有的词特征从源领域传到目标领域。Fang 等人<sup>[57]</sup>利用迁移学习对跨网络中的协作分类进行研究,试图从源网络将共同的隐性结构特征迁移到目标网络。该算法通过构造源网络和目标网络的标签传播矩阵来发现这些隐性特征。Wei 等人<sup>[23]</sup>提出一种借用长文本帮助短文本分类的迁移学习方法。他们首先把目标领域数据中的 Tag 作为关键词从搜索引擎检索最相关的网页,然后利用隐性语义分析方法抽取语义关键词。第三,建立无向图,Tag 作为节点,通过拉普拉斯特征映射,每个节点被表示到低维空间。这样,短文本可以转化为新的特征表示。最后通过最小化样本与特征表示之间的互信息,可以得到正确的短文本分类器。

### 1.2.3 基于特征映射方法

基于特征映射的迁移学习方法是把各个领域的数据从原始高维特征空间映射到低维特征空间,在该低维空间下,源领域数据与目标领域数据拥有相同的分布<sup>[3,58-61]</sup>。这样就可以利用低维空间表示的有标签的源领域样本数据训练分类器,对目标测试数据进行预测。该方法与特征选择的区别在于,这些映射得到的特征不在原始的特征当中,是全新的特征。

Pan 等人<sup>[58]</sup>提出了一种新的维度降低迁移学习方法,他通过最小化源领域数据与目标领域数据在隐性语义空间上的最大均值偏差(maximum mean discrepancy),从而求解得到降维后的特征空间。在该隐性空间上,不同的领域具有相同或者非常接近的数据分布,因此就可以直接利用监督学习算法训练模型对目标领域数据进行预测。Gu 等人<sup>[62]</sup>探讨了多个聚类任务的学习(这些聚类任务是相关的),提出了一种寻找共享特征子空间的框架。在该子空间中,各个领域的数据共享聚类中心,而且他们还把这个框架推广到直推式迁移分类学习。Blitzer 等人<sup>[3]</sup>提出了一种结构对应学习算法(structural corresponding learning,简称 SCL),该算法把领域特有的特征映射到所有领域共享的“轴”特征,然后就在这个“轴”特征下进行训练学习。SCL 算法已被用到词性标注<sup>[31]</sup>以及情感分析<sup>[63]</sup>中。类似的工作还有文献<sup>[64]</sup>等。

Kan 等人<sup>[65]</sup>提出一种新的目标化源领域数据的领域适应性方法,用于人脸识别。该方法首先将目标领域数据和源领域数据映射到一个共享的子空间,在该子空间中,源领域数据由目标领域数据线性表示,而且保持稀疏重构特性以及领域本身的结构。当求出线性表示系数以后,源领域数据可以由目标领域数据重新线性表示,最后利用监督模型进行学习分类。Shao 等人<sup>[66]</sup>讨论一种迁移学习方法用于视觉分类,该方法映射源领域和目标领域数据到一个泛化子空间,其中目标领域数据可以被表示为一些源数据的组合。通过在迁移过程中加入低秩约束,来保持源领域和目标领域的结构。Yeh 等人<sup>[67]</sup>提出一种新的领域适应性方法以解决跨领域模式识别问题。他们使用典型相关分析方法(CCA)得到相关子空间作为所有领域数据的联合表示,并提出核典型相关分析方法(KCCA)以处理非线性相关子空间的情况。特别地,他们提出一种新的带有相关性正则化的支持向量机方法,可以在分类器设计中加入领域适应性能力,从而进行领域适应性模式分类。Wang 等人<sup>[68]</sup>挖掘词特征上的概念进行知识迁移,用于跨语言网页分类。他们的工作基于以下观测:不同领域可能采用不同的词特征来表示同一个概念,那么就可以利用独立于领域的概念作为知识迁移的桥梁。Long 等人<sup>[69]</sup>进一步提出双重迁移学习方法,进一步考虑词特征概念的分类,分成两种不同的概念,即不同领域采用不同词特征的概念和不同领域也采用相同词特征的概念。

### 1.2.4 基于权重方法

在迁移学习中,有标签的源领域数据的分布与无标签的目标领域数据的分布是不一样的,因此那些有标签

的样本数据并不一定是全部有用的.如何侧重选择那些对目标领域分类有利的训练样本?这就是基于实例的迁移学习所要解决的问题.基于实例的迁移学习通过度量有标签的训练样本与无标签的测试样本之间的相似度来重新分配源领域中样本的采样权重.相似度大的,即对训练目标模型有利的训练样本被加大权重,否则权重被削弱.Jiang 等人<sup>[26]</sup>提出了一种实例权重框架来解决自然语言处理任务下的领域适应问题.他们首先从分布的角度分析了产生领域适应问题的原因,主要有两方面:实例的不同分布以及分类函数的不同分布.因此他们提出了一个最小化分布差异性的风险函数,来解决领域适应性问题.Dai 等人<sup>[12]</sup>将 Boosting 学习算法扩展到迁移学习中,提出了 TrAdaBoost 算法.在每次迭代中改变样本被采样的权重,即在迭代中源领域中的样本权重被减弱,而有利于模型训练的目标领域中的样本权重被加强.他们还利用 PAC 理论分析证明了该算法的有效性.下面简要介绍 TrAdaBoost 算法.

用于迁移学习任务中的源领域数据与目标领域数据虽然分布不同,但却是相关的.也就是说,辅助的源领域中的训练样本存在一部分比较适合用来学习一个有效的分类模型,并且对目标测试样本是适用的.于是 TrAdaBoost 算法的目标就是从辅助的源数据中找出那些适合测试数据的实例,并把这些实例迁移到目标领域中少量有标签样本的学习中去.该算法的关键思想是利用 Boosting 的技术过滤掉源领域数据中那些与目标领域中少量有标签样本最不像的样本数据.其中,Boosting 技术用来建立一种自动调整权重机制,于是重要的源领域样本数据权重增加,不重要的源领域样本数据权重减小.在 TrAdaBoost 中,AdaBoost<sup>[70]</sup>被用在目标领域中少量有标签的样本中,以保证分类模型在目标领域数据上的准确性;而 Hedge( $\beta$ )<sup>[70]</sup>被用在源领域数据上,用于自动调节源领域数据的重要度.一个直观 TrAdaBoost 的例子如图 1 所示.另外,对参数加权组合的工作,如文献[71].

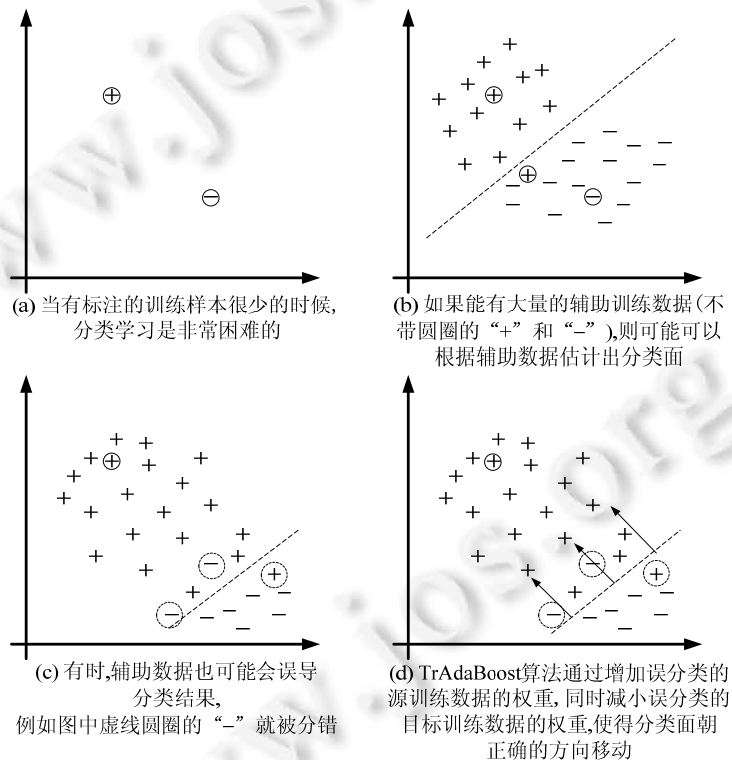


Fig.1 An intuitive example about the idea of TrAdaBoost<sup>[12]</sup>

图1 关于 TrAdaBoost 算法思想的一个直观示例<sup>[12]</sup>

根据是否从多个源领域数据学习,迁移学习算法又可以分为单个源领域以及多个源领域的迁移学习. Ben-David 等人<sup>[2]</sup>分析了领域数据的表示,并提出了一个很好的模型,该模型不仅最小化分类模型在训练数据上的泛化误差,而且最小化源领域与目标领域之间的不同性.Ling 等人<sup>[72]</sup>提出了一种新的光谱分类算法,该算法通

过优化一个目标函数来寻找源领域中的监督信息与目标领域的本质结构之间的最大一致性.Zhuang 等人<sup>[17]</sup>综合半监督学习中的几种正则化准则,提出了基于混合正则化准则的迁移学习框架.Mahmud 等人<sup>[53,73]</sup>从算法信息论的角度来研究迁移学习,该方法度量了不同任务之间的相关性,然后决定多少信息可以做迁移以及如何迁移这些信息.Xing 等人<sup>[7]</sup>提出了一种直推式迁移学习方法,该方法首先开发利用所有数据集(包括源领域数据和目标领域数据)上的几何分布结构,然后再利用目标领域上的流形结构.针对多源领域学习问题,Gao 等人<sup>[74]</sup>提出了一种多模型局部结构映射方案,实际上是对不同源领域训练得到的模型赋予不同的投票权重,而该权重是由预测样本本身的局部分布结构决定的.Gao 等人<sup>[75]</sup>解决了不同模型的一致性问题.这两个多源领域学习的工作很好地处理了多个模型的集成问题.为了更加深入地挖掘、开发各个源领域数据的内部结构或者数据分布,Luo 和 Zhuang 等人<sup>[14,18]</sup>提出了一致性正则化框架,在这个框架下,局部的子分类器不仅考虑了在源领域上的可利用的局部数据,而且考虑了这些由源领域知识得到的子分类器在目标领域上的预测的一致性.Duan 等人<sup>[15]</sup>将由源领域训练得到的模型作为辅助分类器.

## 2 迁移学习相关理论研究

从理论层面上看,迁移学习问题研究以下问题:第一,什么条件下从源领域数据训练出的分类器能够在目标领域表现出优异的分类性能,即什么条件下可进行迁移?第二,给定无标注目标领域,或者有少量的标记数据,如何在训练过程中与大量有标记的源数据结合使得测试时的误差最小,即迁移学习算法的研究.目前对迁移学习理论研究比较多的主要是在领域适应性方面.

关于领域适应性问题的理论分析最早是在文献[2]中提出的.该文基于 VC 维对领域适应性问题给出了推广性的界.该文最有价值的贡献在于定义了分布之间的距离,此距离与领域适应性有关.在此基础上,对有限 VC 维情况,可用文献[76]中提出的方法,从有限个样本估计适应推广能力.但当 VC 维不是有限的情况下会有什么结论该文并未给出研究,需要进一步探讨.另外,不同的领域分布之间的距离会得出不同精度的误差估计,由此可以通过研究各具特色的距离以解决领域适应性问题,从而适应不同应用场合的需要.Ben-David 还通过实验指出 SCL(structural correspondence learning)<sup>[3]</sup>方法确实能够达到在  $\mathcal{A}$  距离最小的同时间隔损失最低,从而提高目标领域上的学习性能.Ben-David 等人<sup>[2]</sup>分析了领域数据的表示,并提出了一个很好的模型,该模型不仅最小化分类模型在训练数据上的泛化误差,而且最小化源领域与目标领域之间的不同性.这项工作后续研究的阶段性成果发表在文献[36],该文从源数据加权组合获得模型,并给出在特定的经验风险最小化的情形下的误差率.最新的成果发表在 2010 年的 MACHINE LEARNING 杂志上<sup>[77]</sup>.该文研究了在什么条件下一个分类器能在目标领域很好地完成分类任务,还研究了给定目标领域少量的已标注的样本,如何在训练过程中把它们与大量的已标注的源数据相结合,以实现目标误差最小.

Mansour 指出,对任意给定的目标函数,存在一个对源假设的领域加权分布组合使得损失至多为给定的值<sup>[78]</sup>.他还对于任意的目标分布,给出了基于源领域和目标领域之间的 Rényi 散度的领域推广误差<sup>[79]</sup>.更为精确的推广误差上界估计应用到回归和一般的损益函数,并提出通过加权实现经验分布能够更好地反映目标领域分布<sup>[80]</sup>.文献[81]提出一种新的框架来分析典型的领域适应性类型学习过程的理论性质,即将多个源领域和单个目标领域的学习问题结合考虑.该文使用积分概率度量来测量两个域的分布之间的差异,同时与 H-散度和差异距离进行了比较.并且,针对多领域分别开发了 Hoeffding 型、Bennett 和 McMiarmid 型偏差不等式,然后给出了对称不等式.接下来,又利用以上不等式分别获得基于统一熵数的 Hoeffding 型和 Bennett 型泛化边界.此外,文献[81]还提出了一种基于 Rademacher 复杂度的泛化边界.最后,分析了渐近收敛性和学习过程的收敛速度.尽管已经进行了一些理论尝试,但还远远不够,对迁移学习有效性的理论研究还有待进一步深入.

下面介绍我们利用生成模型在迁移学习方面所做的工作.

## 3 基于生成模型的迁移学习方法

目前很多迁移学习算法都是基于判别模型的学习算法<sup>[13,58-60]</sup>,判别算法是根据给定源领域数据  $X$ ,直接训

练得到判别模型  $P(Y|X)$ .由于源领域与目标领域数据分布不一致,判别模型未考虑联合概率  $P(X,Y)$ ,因此有时不能得到很好的预测结果.区别于判别模型,生成模型先计算得到联合概率  $P(X,Y)$ ,然后再计算  $P(Y|X)$ .这样,生成模型提供了一种很好的机制,可对源领域和目标领域数据不同分布进行建模,从而实现源领域与目标领域之间的知识迁移以提高算法的性能<sup>[68,82-84]</sup>.文献[82]还对为什么采用生成模型进行迁移学习算法研究进行了讨论.由于生成模型对联合概率进行建模,具有更强的领域间不同数据分布的建模能力,所以更适合于迁移学习.

在迁移学习文本分类中,源领域数据与目标领域数据在原始词特征上分布不一致,也就是说,它们可能会采用不同的词特征来表示同一个语义概念.但我们发现,不同的领域数据,其词特征聚类(又称词特征概念)与文档类别(又称文档聚类、文档概念)之间的关联关系可能是一样的.比如,表示词特征概念“Computer Science”的词有‘hardware’,‘software’,‘program’,‘programmer’,‘disks’以及‘rom’等,但是这些词在不同的领域中可能频率相差很大.在关于硬件公司的新闻网页中,‘hardware’,‘disks’以及‘rom’可能是高频词.相反地,在关于软件公司的新闻网页中,‘software’,‘program’,以及‘programmer’更可能是高频词.因此不同的领域表示同一个概念的词特征差异很大,这就会导致用原始特征训练得到的分类器可能是不可靠的.如果我们能够找出各个领域的词特征概念,并用它们来预测样本的类别,那么就会比直接用原始特征要更可靠和有效.

从上面的例子可以看出,一个网页无论来自于哪一个领域,只要其包含特征概念“Computer Science”,那么该网页就是属于计算机相关的文档类.我们把表示词特征概念的词,定义为词概念外延,把词特征概念与文档类别之间的关系定义为词概念内涵,文档类别中包含的具体文档定义为文档类别外延.Zhuang 等人<sup>[82]</sup>研究基于生成模型的挖掘多领域之间共性与特性的跨领域分类方法,对有效挖掘词特征聚类与文档类别关联关系进行了深入研究.其主要思想如图 2 所示,图中  $y$  表示词特征概念, $z$  表示文档类别(或文档聚类), $c$  表示领域标签, $w$  和  $d$  分别表示词和文档.

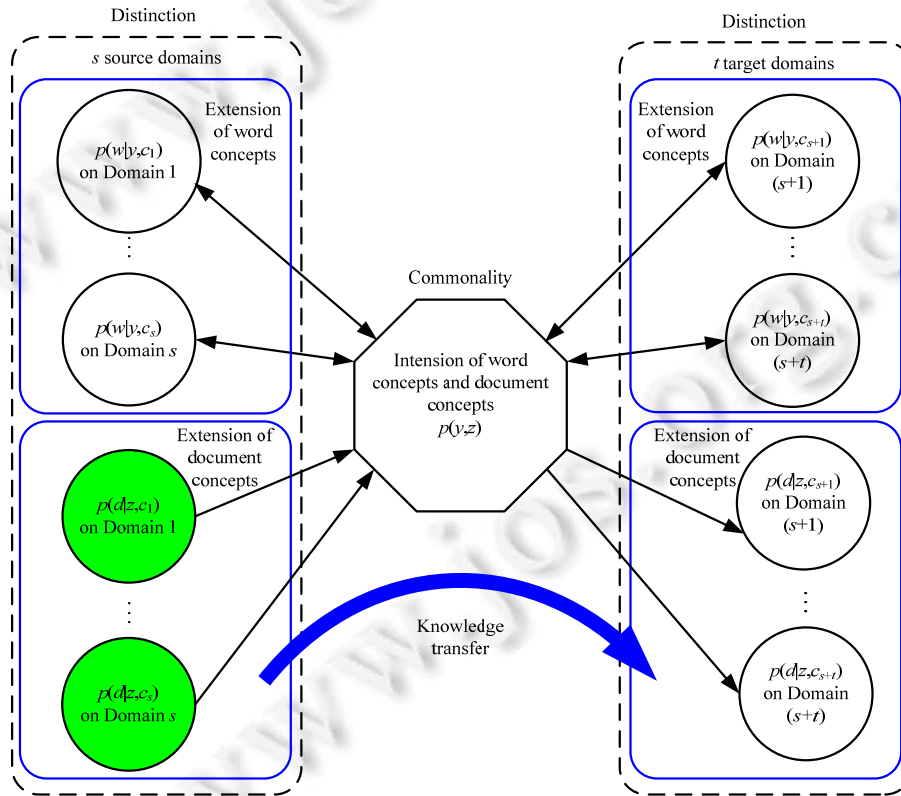


Fig.2 The commonality and distinction among different domains

图 2 不同领域之间的共性和特性

图 2 中,每个大矩形框中又包含两个小的矩形框,分别为各个领域词特征概念的外延和文档概念的外延.领域的特性包括所有的外延,而领域的共性则是它们共享的词特征概念与文档概念之间的联合概率分布,如图中的八边形所示.实际上,源领域中的数据是有标记的,即源领域中文档概念的外延已知,可以作为整个模型的监督信息,如图中的实心圆圈所示.这些监督信息通过领域之间的共性实现知识的迁移,领域的共性起到桥的作用,最后实现对目标领域数据的分类预测.实验结果表明,该算法具有较强的迁移学习能力,可以处理迁移学习比较难的分类问题.

表 1 给出了该算法挖掘出来的词特征概念的部分结果,可以看到,词特征概念“Space Science”在不同的领域中采用不同的关键词进行表示.他们还对基于判别模型和基于生成模型的迁移学习算法进行了初步的探讨,认为基于生成模型的方法可以有效地对源领域与目标领域之间的不同性进行建模,可能更适合做迁移学习.

以往的工作<sup>[68,82]</sup>假设源领域和目标领域共享相同的概念集,但是除了共享概念以外,不同领域可能还包含自己独特的概念.Zhuang 等人<sup>[84]</sup>对不同领域的概念进行了深入的分析,把概念分成 3 类:一致性概念、相似概念以及领域特有的概念.他们提出了一般的概率统计模型来挖掘这 3 种概念,并开发了一种 EM 算法进行求解.大量的实验结果表明,所提出的模型优于所比较的迁移学习算法.

**Table 1** The word concepts output by the proposed method in Ref.[82]

**表 1** 文献[82]提出方法挖掘的词特征概念

Associated with concept: space science	Domain 1	rocket, esa, assist, frank, af, thu, helsinki, ron, atlantic, jet, observer, satellite, venus, sei, min, ir, russia, stars, star, ray
	Domain 2	relay, km, rat, pixel, command, elements, arc, acceleration, nasa, earth, fuse, ground, bulletin, pub, anonymous, faq, unix, cit, ir, amplifier
	Domain 3	from, earth, science, word, pictures, years, center, data, national, dale, nasa, gif, reports, mil, planet, field, jpl, ron, smith, unix
	Domain 4	service, archive, unit, magnetic, thousands, technology, information, arc, keys, faq, probes, ir, available, gov, embedded, tens, data, system, unix, mil
Associated with concept: computer science	Domain 1	support, astronomer, near, thousands, million, you, vnet, copy, ad, bright, lab, idea, data, hardware, engines, ibm, project, soviet, software, program
	Domain 2	legally, schemes, protected, bytes, mq, disks, patch, registers, machine, pirates, install, card, rom, screen, protection, disk, ram, tape, mb, copy
	Domain 3	discomfort, friend, normal, self, tests, programmer, steve, state, program, lab, you, your, jon, my, headache, trial, she, pain, page, trials
	Domain 4	wcs, cipher, scheme, brute, user, file, encryption, message, serial, decryption, crypto, keys, cryptosystems, skipjack, plaintext, secure, key, encrypted, nsa, des
Associated with concept: car	Domain 1	saves, power, was, at, disappointment, al, europeans, will, ny, north, their, they, deal, best, year, sports, cs, new, series, gm
	Domain 2	crash, price, vehicle, insurance, handling, gas, xs, dealer, cruiser, leather, buy, latech, fj, paint, ride, buying, bmw, engine, car, Honda
	Domain 3	or, value, they, wade, good, car, better, best, three, performance, more, runner, than, average, dl, extra, base, cs, al, year
	Domain 4	dealer, camry, saab, engine, eliot, requests, mazda, liter, mustang, diesel, wagon, nissan, mileage, byte, saturn, toyota, si, cars, car, db

\*Domain 1: rec.sport.hockey vs. sci.space, Domain 2: rec.motorcycles vs. sci.electronics  
Domain 3: rec.sport.baseball vs. sci.med, Domain 4: rec.autos vs. sci.crypt

## 4 迁移学习应用研究

目前,迁移学习典型的应用方面的研究主要包含有文本分类、文本聚类、情感分类、图像分类、协同过滤、基于传感器的定位估计、人工智能规划等.

在文本处理领域已有大量的迁移学习工作<sup>[4,5,18,24,56,62,69,83,85]</sup>.Dai 等人<sup>[4]</sup>提出联合聚类方法,同时对文档以及词特征进行聚类,通过不同领域共享相同的词特征进行知识迁移.他们还提出迁移贝叶斯分类器<sup>[5]</sup>,首先估计源领域数据的数据分布,然后不断修正使其适应于目标领域数据.Zhuang 等人<sup>[56]</sup>在概念层面对文本进行处理,提出挖掘文档概念与词特征概念的迁移学习方法.在此基础上 Long 等人<sup>[69]</sup>提出了双重迁移模型,进一步对概念进行划分,提高算法分类准确率.Gu 等人<sup>[62]</sup>提出共享子空间的多任务聚类方法,并应用于迁移分类中.文献[63,86,87]等对情感分类进行了研究, Li 等人<sup>[86]</sup>利用独立于领域的带有情感倾向的词先验以及领域相关的无标签数据,来提高情感分类的准确率.Pan 等人<sup>[87]</sup>提出一种新的谱特征对齐方法,以独立于领域的词特征为桥梁,把



来自不同领域的领域特有词特征对齐到一个聚类中.这样,这些聚类可以降低领域特有的词的缺口,从而提高目标领域上的准确率.

在图像处理方面,Dai 等人<sup>[16]</sup>提出一种翻译迁移学习方法,借助文本数据来辅助图像聚类.Raina 等人<sup>[54]</sup>提出一种新的从无标签数据进行自学习的方法,该方法利用系数编码技术从大量的无标签数据上构造高层特征,以提高图像分类性能.Zhu 等人<sup>[88]</sup>研究了一种异构迁移学习方法,利用图像上的 Tag 标签信息作为文本与图像之间知识迁移的桥梁,从而提高图像数据上的分类效果.在图像分类方面的工作还有文献<sup>[59]</sup>等.

在协同过滤方面的工作有文献<sup>[75,89-92]</sup>等,Wang 等人提出特征子空间的迁移学习方法来克服协同过滤中的稀疏问题,即从辅助数据中学习得到的用户特征子空间被迁移到目标领域中.Pan 等人<sup>[90]</sup>研究了协同过滤中带有不确定评分的迁移学习算法,即在优化目标矩阵分解中考虑不确定评分的辅助数据作为限制.Cao 等人<sup>[92]</sup>提出基于项目潜在特征共享策略的链接预测模型,性能上比单个任务的学习有所提升.

香港科技大学 Qiang Yang 实验室还做了一系列的室内定位的迁移学习方面的工作<sup>[13,58,93]</sup>.Wang 等人<sup>[93]</sup>的工作基于这样的观测,即使不同楼层的信号差别非常大,但是不同楼层的设计是相似的,因此可以把不同楼层的数据嵌入到一些共同的低维流形,这样可以通过这些流形把标签从标记样本传播到无标记样本.Zheng 等人<sup>[94]</sup>提出一种新的多任务学习,对室内定位系统的多个设备进行同时学习,并假设各个设备在隐性空间上的假设是相似的.另外,在智能规划方面,Zhou 等人<sup>[95]</sup>提出一种新的迁移学习框架 TRAMP,将迁移学习用于人工智能规划中的动作模型获取.该方法首先建立源领域与目标领域之间的结构映射来迁移知识,然后从 Web 搜索中开发额外的知识来从源领域中连接和迁移知识.在排序以及度量学习方面的工作有文献<sup>[61,96,97]</sup>等.传统的视觉重排序已经不能很好地提高基于文本的视频搜索,因为低层次的视觉特征与高层次的语义概念有很大的缺口.Tian 等人<sup>[61]</sup>提出一种新的维度降低工具,可以有效地编码用户的标记信息(标记信息由交互得到)来提高检索结果.Geng 等人<sup>[97]</sup>介绍了一种领域适应性度量学习方法.在可再生核希尔伯特空间中,该方法在常规度量学习中引入依赖于数据的正则化项,来解决源领域与目标领域之间的分布不同性.

总之,越来越多的迁移学习工作被应用到各种应用中,也在研究上取得了非常显著的效果.但在目前大数据背景下,已有的算法还不能满足实际的应用需求,处理的数据量还比较小,而且算法复杂度也比较高.下一步的研究应更关注于高效算法的设计上,以做到确实满足实际需要.

## 5 未来研究方向

本文系统地给出了迁移学习算法以及相关理论的研究进展状况.迁移学习作为一个新兴的研究领域,还很年轻,主要研究仍集中在算法方面,因此值得我们进一步展开和深入.

迁移学习最早来源于教育心理学,这里借用美国心理学家 Judd<sup>\*\*\*</sup>提出的“类化说”学习迁移理论来讨论目前机器学习领域迁移学习研究还存在的 3 个问题.首先,Judd 认为在先期学习  $A$  中获得的東西,之所以能迁移到后期学习  $B$  中,是因为在学习  $A$  时获得了一般原理,这种原理可以部分或全部运用于  $A, B$  之中.根据这一理论,两个学习活动之间存在的共同要素,是产生迁移的必要前提.这也就是说,想从源领域中学习知识并运用到目标领域中,必须保证源领域与目标领域有共同的知识,那么如何度量这两个领域的相似性与共同性,是问题之一.第二,Judd 的研究表明,知识的迁移是存在的,只要一个人对他的经验、知识进行了概括,那么从一种情境到另一种情境的迁移是可能的.知识概括化的水平越高,迁移的范围和可能性越大.把该原则运用到课堂上,同样的教材采用不同的教学方法,产生的迁移效果是不一样的,即可能产生积极迁移也可能产生相反的作用.即同样的教材内容,由于教学方法不同,而使教学效果大为悬殊,迁移的效应也大不相同.所以针对不同的学习问题,研究有效的迁移学习算法也是另一个重要问题.第三,根据 Judd 的泛化理论,重要的是在讲授教材时要鼓励学生对本质的基本概念进行抽象或概括.抽象与概括的学习方法是最重要的方法,在学习时对知识进行思维加工,区别本质的和非本质的属性,偶然的和必然的联系,舍弃那些偶然的、非本质的东西,牢牢把握那些必然的本质的东西.这

\*\*\* <http://baike.baidu.com/subview/226455/11107073.htm>

种学习方法能使学生的认识从低级的感性阶段上升到高级的理性阶段,从而实现更广泛、更成功的正向迁移。也就是说在迁移学习的过程中,应该避免把非本质的、偶然的知识,当成本质的(领域共享的)、必然的知识,实现正迁移。所以,如何实现正迁移,避免负迁移也是迁移学习的一个重要研究问题。

针对以上讨论分析,我们认为后续研究有以下几个可能的方向。第一,针对领域相似性、共同性的度量,目前还没有深入的研究成果,那么首要任务就是研究准确的度量方法。第二,在算法研究方面,不同的应用,迁移学习算法的需求有所不同。目前很多研究工作主要集中在迁移学习分类算法方面,其他方面的应用算法有待进一步研究,比如情感分类、强化学习、排序学习、度量学习,人工智能规划等。第三,关于迁移学习算法有效性的理论研究还很缺乏,研究可迁移学习条件,获取实现正迁移的本质属性,避免负迁移,也是方向之一。最后,在大数据环境下,研究高效的迁移学习算法尤为重要。目前的研究主要仍集中在研究领域,数据量小而且测试数据非常标准,应把研究的算法瞄准实际应用数据,以顺应目前大数据挖掘研究浪潮。

#### References:

- [1] He Q, Zhao XR, Shi ZZ. Minimal consistent subset for hyper surface classification method. *Int'l Journal of Pattern Recognition and Artificial Intelligence*, 2008,22(1):95-108. [doi: 10.1142/S0218001408006132]
- [2] Ben-David S, Blitzer J, Crammer K, Pereira F. Analysis of representations for domain adaptation. In: Platt JC, Koller D, Singer Y, Roweis ST, eds. *Proc. of the Advances in Neural Information Processing Systems 19*. Cambridge: MIT Press, 2007. 137-144.
- [3] Blitzer J, McDonald R, Pereira F. Domain adaptation with structural correspondence learning. In: Jurafsky D, Gaussier E, eds. *Proc. of the Int'l Conf. on Empirical Methods in Natural Language Processing*. Stroudsburg PA: ACL, 2006. 120-128.
- [4] Dai WY, Xue GR, Yang Q, Yu Y. Co-Clustering based classification for out-of-domain documents. In: *Proc. of the 13th ACM Int'l Conf. on Knowledge Discovery and Data Mining*. New York: ACM Press, 2007. 210-219. [doi: 10.1145/1281192.1281218]
- [5] Dai WY, Xue GR, Yang Q, Yu Y. Transferring naive Bayes classifiers for text classification. In: *Proc. of the 22nd Conf. on Artificial Intelligence*. AAAI Press, 2007. 540-545.
- [6] Liao XJ, Xue Y, Carin L. Logistic regression with an auxiliary data source. In: *Proc. of the 22nd Int'l Conf. on Machine Learning*. San Francisco: Morgan Kaufmann Publishers, 2005. 505-512. [doi: 10.1145/1102351.1102415]
- [7] Xing DK, Dai WY, Xue GR, Yu Y. Bridged refinement for transfer learning. In: *Proc. of the 11th European Conf. on Practice of Knowledge Discovery in Databases*. Berlin: Springer-Verlag, 2007. 324-335. [doi: 10.1007/978-3-540-74976-9\_31]
- [8] Mahmud MMH. On universal transfer learning. In: *Proc. of the 18th Int'l Conf. on Algorithmic Learning Theory*. Sendai, 2007. 135-149. [doi: 10.1007/978-3-540-75225-7\_14]
- [9] Samarth S, Sylvian R. Cross domain knowledge transfer using structured representations. In: *Proc. of the 21st Conf. on Artificial Intelligence*. AAAI Press, 2006. 506-511.
- [10] Bel N, Koster CHA, Villegas M. Cross-Lingual text categorization. In: *Proc. of the European Conf. on Digital Libraries*. Berlin: Springer-Verlag, 2003. 126-139. [doi: 10.1007/978-3-540-45175-4\_13]
- [11] Zhai CX, Velivelli A, Yu B. A cross-collection mixture model for comparative text mining. In: *Proc. of the 10th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. New York: ACM, 2004. 743-748. [doi: 10.1145/1014052.1014150]
- [12] Dai WY, Yang Q, Xue GR, Yu Y. Boosting for transfer learning. In: *Proc. of the 24th Int'l Conf. on Machine Learning*. San Francisco: Morgan Kaufmann Publishers, 2007. 193-200. [doi: 10.1145/1273496.1273521]
- [13] Pan SJ, Yang Q. A survey on transfer learning. *IEEE Trans. on Data Engineering*, 2010,22(10):1345-1359. [doi: 10.1109/TKDE.2009.191]
- [14] Luo P, Zhuang FZ, Xiong H, Xiong YH, He Q. Transfer learning from multiple source domains via consensus regularization. In: *Proc. of the 17th ACM Conf. on Information and Knowledge Management*. New York: ACM Press, 2008. 103-112. [doi: 10.1145/1458082.1458099]
- [15] Duan LX, Tsang IW, Xu D, Chua TS. Domain adaptation from multiple sources via auxiliary classifiers. In: *Proc. of the 26th Annual Int'l Conf. on Machine Learning*. New York: ACM, 2009. 289-296. [doi: 10.1145/1553374.1553411]
- [16] Dai WY, Chen YQ, Xue GR, Yang Q, Yu Y. Translated learning: Transfer learning across different feature spaces. In: Koller D, Schuurmans D, Bengio Y, Bottou L, eds. *Proc. of the Advances in Neural Information Processing Systems 20*. Cambridge: MIT Press, 2008. 353-360.

- [17] Zhuang FZ, Luo P, He Q, Shi ZZ. Inductive transfer learning for unlabeled target-domain via hybrid regularization. *Chinese Science Bulletin*, 2009,54(14):2470–2478. [doi: 10.1007/s11434-009-0171-x]
- [18] Zhuang FZ, Luo P, Xiong H, Xiong YH, He Q, Shi ZZ. Cross-Domain learning from multiple sources: A consensus regularization perspective. *IEEE Trans. on Knowledge And Data Engineering*, 2010,22(12):1664–1678. [doi: 10.1109/TKDE.2009.205]
- [19] Rosenstein MT, Marx Z, Kaelbling LP. To transfer or not to transfer. In: *Proc. of the Neural Information Processing Systems 2005 Workshop on Inductive Transfer: 10 Years Later*. Cambridge: MIT Press, 2005.
- [20] Dai WY, Jin O, Xue GR, Yang Q, Yu Y. Eigen transfer: A unified framework for transfer learning. In: *Proc. of the 24th Int'l Conf. on Machine Learning*. San Francisco: Morgan Kaufmann Publishers, 2009. 193–200. [doi: 10.1145/1553374.1553399]
- [21] Dai WY. Instance-Based and feature-based transfer learning [MS. Thesis]. Shanghai: Shanghai Jiaotong University, 2008. 1–55 (in Chinese with English abstract).
- [22] Shi XX. Research and application of active transfer learning models [MS. Thesis]. Guangzhou: Sun Yat-Sen University, 2009. 1–69 (in Chinese with English abstract).
- [23] Wei FM, Zhang JP, Chu Y, Yang J. FSFP: Transfer learning from long texts to the short. *Applied Mathematics & Information Sciences*, 2014,8(4):2033–2044. [doi: 10.12785/amis/080462]
- [24] Dai WY, Yang Q, Xue GR, Yu Y. Self-Taught clustering. In: *Proc. of the 24th Int'l Conf. on Machine Learning*. San Francisco: Morgan Kaufmann Publishers, 2008. 200–207. [doi: 10.1145/1390156.1390182]
- [25] Samanta S, Tirumalai SA, Das S. Cross-Domain clustering performed by transfer of knowledge across domains. In: *Proc. of the 2013 IEEE 4th National Conf. on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*. 2013. 1–4. [doi: 10.1109/NCVPRIPG.2013.6776213]
- [26] Jiang J, Zhai CX. Instance weighting for domain adaptation in NLP. In: Carroll JA, van den Bosch A, Zaenen A, eds. *Proc. of the 45th Annual Meeting of the Association for Computational Linguistics*. Stroudsburg PA: Association for Computational Linguistics, 2007. 264–271.
- [27] Lee SI, Chatalbashev V, Vickrey D, Koller D. Learning a meta-level prior for feature relevance from multiple related tasks. In: *Proc. of the 24th Int'l Conf. on Machine Learning*. New York: ACM, 2007. 489–496. [doi: 10.1145/1273496.1273558]
- [28] Wang C, Mahadevan S. Manifold alignment using procrustes analysis. In: *Proc. of the 25th Int'l Conf. on Machine Learning*. New York: ACM, 2008. 1120–1127. [doi: 10.1145/1390156.1390297]
- [29] Lawrence ND, Platt JC. Learning to learn with the informative vector machine. In: *Proc. of the 21st Int'l Conf. on Machine Learning*. New York: ACM, 2004. 65–72. [doi: 10.1145/1015330.1015382]
- [30] Schwaighofer A, Tresp V, Yu K. Learning Gaussian process kernels via hierarchical Bayes. In: *Proc. of the Advances in Neural Information Processing Systems 17*. Cambridge: MIT Press, 2005. 1209–1216.
- [31] Evgeniou T, Pontil M. Regularized multi-task learning. In: *Proc. of the 10th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. New York: ACM, 2004. 109–117. [doi: 10.1145/1014052.1014067]
- [32] Zadrozny B. Learning and evaluating classifiers under sample selection bias. In: *Proc. of the 21st Int'l Conf. on Machine Learning*. New York: ACM, 2004. 114–121. [doi: 10.1145/1015330.1015425]
- [33] Huang JY, Smola AJ, Gretton A, Borgwardt KM. Correcting sample selection bias by unlabeled data. In: Platt JC, Koller D, Singer Y, Roweis ST, eds. *Proc. of the Advances in Neural Information Processing Systems 19*. Cambridge: MIT Press, 2007. 601–608.
- [34] Fan W, Davidson I, Zadrozny B, Yu PS. An improved categorization of classifier's sensitivity on sample selection bias. In: *Proc. of the 5th Int'l Conf. on Data Mining*. Los Vaqueros: IEEE Computer Society, 2005. 605–608. [doi: 10.1109/ICDM.2005.24]
- [35] Ando RK, Zhang T. A high-performance semi-supervised learning method for text chunking. In: *Proc. of the 43rd Annual Meeting on Association for Computational Linguistics*. Stroudsburg, 2005. 1–9. [doi: 10.3115/1219840.1219841]
- [36] Blitzer J, Crammer K, Kulesza A, Pereira F, Wortman J. Learning bounds for domain adaptation. In: Koller D, Schuurmans D, Bengio Y, Bottou L, eds. *Proc. of the Advances in Neural Information Processing Systems 20*. Cambridge: MIT Press, 2008. 129–136.
- [37] Wang Z, Song YQ, Zhang CS. Transferred dimensionality reduction. In: *Proc. of the European Conf. on Machine Learning and Knowledge Discovery in Databases*. Berlin, Heidelberg: Springer-Verlag, 2008. 550–565. [doi: 10.1007/978-3-540-87481-2\_36]
- [38] Zhou ZH. Learning with unlabeled data and its application to image retrieval. In: *Proc. of the 9th Pacific Rim Int'l Conf. on Artificial Intelligence*. Berlin: Springer-Verlag, 2006. 5–10. [doi: 10.1007/978-3-540-36668-3\_3]
- [39] Zhu XJ. Semi-Supervised learning literature survey [Ph.D. Thesis]. Madison: Department of Computer Sciences, University of Wisconsin, 2005.

- [40] Joachims T. Transductive inference for text classification using support vector machines. In: Buford JF, Stevens SM, Bulterman DCA, Jeffay K, Zhang HJ, eds. Proc. of the 16th Int'l Conf. on Multimedia. Augsburg, New York: ACM Press, 1999. 200–209.
- [41] Joachims T. Transductive learning via spectral graph partitioning. In: Rowe LA, Vin HM, Plagemann T, Shenoy PJ, Smith JR, eds. Proc. of the 16th Int'l Conf. on Multimedia. Augsburg, New York: ACM Press, 2003. 290–297.
- [42] Tong S, Chang E. Support vector machine active learning for image retrieval. In: Proc. of the 9th ACM Int'l Conf. on Multimedia. New York: ACM Press, 2001. 107–118. [doi: 10.1145/500141.500159]
- [43] Cohn D, Atlas L, Ladner R. Improving generalization with active learning. *Machine Learning*, 1994,15(2):201–221.
- [44] Sindhwani V, Niyogi P. A co-regularized approach to semi-supervised learning with multiple views. In: De Raedt L, Wrobel S, eds. Proc. of the ICML Workshop on Learning with Multiple Views. San Francisco: Morgan Kaufmann Publishers, 2005. 74–79.
- [45] Yarowsky D. Unsupervised word sense disambiguation rivaling supervised methods. In: Proc. of the 33rd Annual Meeting on Association for Computational Linguistics. Stroudsburg: Association for Computational Linguistics, 1995. 189–196. [doi: 10.3115/981658.981684]
- [46] Blum MT. Combining labeled and unlabeled data with co-training. In: Bartlett PL, Mansour Y, eds. Proc. of the 11th Annual Conf. on Computational Learning Theory. New York, 1998. 92–100.
- [47] Dasgupta S, Littman ML, McAlleste D. PAC generalization bounds for co-training. In: Proc. of the Advances in Neural Information Processing Systems 13. Cambridge: MIT Press, 2001. 375–382.
- [48] Abney A. Bootstrapping. In: Proc. of the 40th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: Association for Computational Linguistics, 2002. 360–367.
- [49] Abney A. Understanding the Yarowsky algorithm. *Journal of Computational Linguistics*, 2004. 365–395. [doi: 10.1162/0891201041850876]
- [50] Shi XX, Fan W, Ren JT. Actively transfer domain knowledge. In: Proc. of the European Conf. on Machine Learning and Knowledge Discovery in DataBases. Berlin: Springer-Verlag, 2008. 342–357. [doi: 10.1007/978-3-540-87481-2\_23]
- [51] Belkin M, Niyogi P, Sindhwani V. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *Journal of Machine Learning Research*, 2007,(7):2399–2434.
- [52] Grandvalet Y, Bengio Y. Semi-Supervised learning by entropy minimization. In: Proc. of the Advances in Neural Information Processing Systems 17. Cambridge: MIT Press, 2005. 529–536.
- [53] Mann GS, McCallum A. Simple, robust, scalable semi-supervised learning via expectation regularization. In: Proc. of the 24th Int'l Conf. on Machine Learning. San Francisco: Morgan Kaufmann Publishers, 2007. 593–600. [doi: 10.1145/1273496.1273571]
- [54] Raina R, Battle A, Lee H, Packer B, Ng AY. Self-Taught learning: Transfer learning from unlabeled data. In: Proc. of the 24th Int'l Conf. on Machine Learning. San Francisco: Morgan Kaufmann Publishers, 2007. 759–766. [doi: 10.1145/1273496.1273592]
- [55] Jiang J, Zhai CX. A two-stage approach to domain adaptation for statistical classifiers. In: Proc. of the 16th ACM Conf. on Information and Knowledge Management. New York: ACM Press, 2007. 401–410. [doi: 10.1145/1321440.1321498]
- [56] Zhuang FZ, Luo P, Xiong H, He Q, Xiong YH. Exploiting associations between word clusters and document classes for cross-domain text categorization. In: Proc. of the 10th SIAM Conf. on Data Mining. Philadelphia: SIAM Press, 2010. 13–24. [doi: 10.1002/sam.10099]
- [57] Fang M, Yin J, Zhu XQ. Transfer learning across networks for collective classification. In: Proc. of the 2013 IEEE 13th Int'l Conf. on Data Mining. 2013. 161–170. [doi: 10.1109/ICDM.2013.116]
- [58] Pan SJ, Kwok JT, Yang Q. Transfer learning via dimensionality reduction. In: Fox D, Gomes CP, eds. Proc. of the 23rd Conf. on Artificial Intelligence. Chicago: AAAI Press, 2008. 677–682.
- [59] Si S, Tao DC, Chan KP. Evolutionary cross-domain discriminative hessian eigenmaps. *IEEE Trans. on Image Processing*, 2010, 19(4):1075–1086. [doi: 10.1109/TIP.2009.2035867]
- [60] Si S, Tao DC, Geng B. Bregman divergence-based regularization for transfer subspace learning. *IEEE Trans. on Knowledge and Data Engineering*, 2010,22(7):919–942. [doi: 10.1109/TKDE.2009.126]
- [61] Tian X, Tao D, Rui Y. Sparse transfer learning for interactive video search reranking. *ACM Trans. on Multimedia Computing, Communications, and Applications (TOMCCAP)*, 2012,8(3):26:1–19. [doi: 10.1145/2240136.2240139]
- [62] Gu QQ, Zhou J. Learning the shared subspace for multi-task clustering and transductive transfer classification. In: Proc. of the 9th Int'l Conf. on Data Mining. Los Vaqueros: IEEE Computer Society, 2009. 159–168. [doi: 10.1109/ICDM.2009.32]

- [63] Blitzer J, Dredze M, Pereira F. Biographies, Bollywood, Boom-boxes and Blenders: Domain adaptation for sentiment classification. In: Carroll JA, van den Bosch A, Zaenen A, eds. Proc. of the 45th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2007. 440–447.
- [64] Xie SH, Fan W, Peng J, Verscheure O, Ren JT. Latent space domain transfer between high dimensional overlapping distributions. In: Proc. of the ACM Conf. on World Wide Web. New York: ACM Press, 2009. 91–100. [doi: 10.1145/1526709.1526723]
- [65] Kan MN, Wu JT, Shan SG, Chen XL. Domain adaptation for face recognition: Targetize source domain bridged by common subspace. Int'l Journal of Computer Vision, 2013. 1–16. [doi: 10.1007/s11263-013-0693-1]
- [66] Shao M, Kit D, Fu Y. Generalized transfer subspace learning through low-rank constraint. Int'l Journal of Computer Vision, 2014. 1–20. [doi: 10.1007/s11263-014-0696-6]
- [67] Yeh Y, Huang C, Wang Y. Heterogeneous domain adaptation and classification by exploiting the correlation subspace. IEEE Trans. on Image Processing, 2013. 2009–2018. [doi: 10.1109/TIP.2014.2310992]
- [68] Wang H, Huang H, Nie FP, Ding C. Cross-Language Web page classification via dual knowledge transfer using nonnegative matrix tri-factorization. In: Proc. of the 34th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. ACM, 2011. 933–942. [doi: 10.1145/2009916.2010041]
- [69] Long MS, Wang JM, Ding GG, Cheng W, Zhang X, Wang W. Dual transfer learning. In: Proc. of the SDM. California: SIAM, 2012. 540–551.
- [70] Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences, 1997,55(1):119–139. [doi: 10.1006/jcss.1997.1504]
- [71] Dredze M, Kulesza A, Crammer K. Multi-Domain learning by confidence-weighted parameter combination. Journal of Machine Learning, 2010,79(1-2):123–149. [doi: 10.1007/s10994-010-5169-8]
- [72] Ling X, Dai WY, Xue GR, Yang Q, Yu Y. Spectral domain-transfer learning. In: Proc. of 14th ACM Int'l Conf. on Knowledge Discovery and Data Mining. New York: ACM Press, 2008. 488–496. [doi: 10.1145/1401890.1401951]
- [73] Mahmud MMH, Ray S. Transfer learning using Kolmogorov complexity: Basic theory and empirical evaluations. In: Platt JC, Koller D, Singer Y, Roweis ST, eds. Proc. of the Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2007. 985–992.
- [74] Gao J, Fan W, Jiang J, Han JW. Knowledge transfer via multiple model local structure mapping. In: Proc. of the 13th ACM Int'l Conf. on Knowledge Discovery and Data Mining. New York: ACM Press, 2008. 283–291. [doi: 10.1145/1401890.1401928]
- [75] Gao J, Fan W, Sun YZ, Han JW. Heterogeneous source consensus learning via decision propagation and negotiation. In: Proc. of the 13th ACM Int'l Conf. on Knowledge Discovery and Data Mining. New York: ACM Press, 2009. 339–348. [doi: 10.1145/1557019.1557061]
- [76] Kifer D, Ben-David S, Gehrke J. Detecting change in data streams. In: Nascimento MA, Özsu MT, Kossman D, Miller RJ, Blakeley JA, Schiefer KB, eds. Proc. of the 30th Int'l Conf. on Very Large Data Bases. Toronto: VLDB Endowment, 2004. 180–191.
- [77] Ben-David S, Blitzer J, Crammer K, Kulesza A, Pereira F, Vaughan JW. A theory of learning from different domains. Journal of Machine Learning, 2010,79(1-2):151–175. [doi: 10.1007/s10994-010-5169-8]
- [78] Mansour Y, Mohri M, Rostamizadeh A. Domain adaptation with multiple sources. In: Koller D, Schuurmans D, Bengio Y, Bottou L, eds. Proc. of the Advances in Neural Information Processing Systems 20. Cambridge: MIT Press, 2008. 1–8.
- [79] Mansour Y, Mohri M, Rostamizadeh A. Multiple source adaptation and the Rényi divergence. In: Bilmes J, Ng AY, eds. Proc. of the 25th Conf. on Uncertainty in Artificial Intelligence. Arlington: AUAI Press, 2009. 367–374.
- [80] Mansour Y, Mohri M, Rostamizadeh A. Domain adaptation: Learning bounds and algorithms. In: Proc. of the 22nd Annual Conf. on Learning Theory. San Francisco: Morgan Kaufmann Publishers, 2009.
- [81] Zhang C, Zhang L, Fan W, Ye JP. Generalization bounds for representative domain adaptation. arXiv preprint arXiv: 1401.0376, 2014. 1–40.
- [82] Zhuang FZ, Luo P, Shen ZY, He Q, Xiong YH, Shi ZZ, Xiong H. Collaborative DUAL-PLSA: Mining distinction and commonality across multiple domains for text classification. In: Proc. of the 19th ACM Int'l Conf. on Information and Knowledge Management. ACM, 2010. 359–368. [doi: 10.1145/1871437.1871486]
- [83] Xue GR, Dai WY, Yang Q, Yu Y. Topic-Bridged PLSA for cross-domain text classification. In: Proc. of the 31st Annual Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. New York: ACM, 2008. 627–634. [doi: 10.1145/1390334.1390441]

- [84] Zhuang FZ, Luo P, Yin PF, He Q, Shi ZZ. Concept learning for cross-domain text classification: A general probabilistic framework. In: Rossi F, ed. Proc. of the 23rd Int'l Joint Conf. on Artificial Intelligence. AAAI Press, 2013. 1960–1966.
- [85] Yang P, Gao W. Information-Theoretic multi-view domain adaptation: A theoretical and empirical study. Journal of Artificial Intelligence Research, 2014,49:501–525.
- [86] Li T, Zhang Y, Sindhwani V. A non-negative matrix tri-factorization approach to sentiment classification with lexical prior knowledge. In: Su KY, Su J, Wiebe J, eds. Proc. of the Joint Conf. of the 47th Annual Meeting of the ACL and the 4th Int'l Joint Conf. on Natural Language Processing of the AFNLP. Singapore: Association for Computational Linguistics, 2009. 244–252.
- [87] Pan SJ, Ni X, Sun JT, Yang Q, Chen Z. Cross-Domain sentiment classification via spectral feature alignment. In: ACM Proc. of the 19th Int'l Conf. on World Wide Web. 2010. 751–760. [doi: 10.1145/1772690.1772767]
- [88] Zhu Y, Chen Y, Lu Z, Pan SJ, Xue GR, Yu Y, Yang Q. Heterogeneous transfer learning for image classification. In: Burgard W, Roth D, eds. Proc. of the AAAI. AAAI Press, 2011. 1304–1309.
- [89] Wang J, Ke L. Feature subspace transfer for collaborative filtering. Neurocomputing, 2014,136:1–6. [doi: 10.1016/j.neucom.2014.01.035]
- [90] Pan W, Xiang EW, Yang Q. Transfer learning in collaborative filtering with uncertain ratings. In: Hoffmann J, Selman B, eds. Proc. of the AAAI. AAAI Press, 2012. 662–668.
- [91] Li B, Zhu XQ, Yang Q. Transfer learning problems in collaborative filtering and link prediction. In: Zhou ZH, Yang Q, eds. Machine Learning and Its Applications 2011. Beijing: Tsinghua University Press, 2011. 33–50 (in Chinese).
- [92] Cao B, Liu NN, Yang Q. Transfer learning for collective link prediction in multiple heterogeneous domains. In: Fürnkranz J, Joachims T, eds. Proc. of the 27th Int'l Conf. on Machine Learning. Omnipress, 2010. 159–166.
- [93] Wang HY, Zheng VW, Zhao J, Yang Q. Indoor localization in multi-floor environments with reduced effort. In: Proc. of the 2010 IEEE Int'l Conf. on Pervasive Computing and Communications (PerCom). 2010. 244–252. [doi: 10.1109/PERCOM.2010.5466971]
- [94] Zheng VW, Pan SJ, Yang Q, Pan JJ. Transferring multi-device localization models using latent multi-task learning. In: Fox D, Gomes CP, eds. Proc. of the AAAI. AAAI Press, 2008. 1427–1432.
- [95] Zhuo HH, Yang Q. Action-Model acquisition for planning via transfer learning. Artificial Intelligence, 2014,212:80–103. [doi: 10.1016/j.artint.2014.03.004]
- [96] Bai J, Zhou K, Xue GR, Zha HY, Sun G, Tseng B, Zheng ZH, Chang Y. Multi-Task learning for learning to rank in Web search. In: Cheung DWL, Song IY, Chu WW, Hu XH, Lin JJ, eds. Proc. of the 18th ACM Conf. on Information and Knowledge Management. New York: ACM Press, 2009. 1549–1552.
- [97] Geng B, Tao D, Xu C. DAML: Domain adaptation metric learning. IEEE Trans. on Image Processing, 2011,20(10):2980–2989. [doi: 10.1109/TIP.2011.2134107]

#### 附中文参考文献:

- [21] 戴文渊.基于实例和特征的迁移学习算法研究[硕士学位论文].上海:上海交通大学,2008.1–55.
- [22] 施潇潇.主动迁移学习模型的研究与应用[硕士学位论文].广州:中山大学,2009.1–69
- [91] 李斌,朱兴全,杨强.协同过滤与链接预测的迁移学习问题.见:周志华,杨强,编.机器学习及其应用 2011.北京:清华大学出版社, 2011.33–50.



庄福振(1983—),男,博士,副研究员,CCF 会员,主要研究领域为迁移学习,数据挖掘,机器学习.



何清(1965—),男,博士,研究员,博士生导师,CCF 高级会员,主要研究领域为机器学习,数据挖掘,基于云计算的海量数据挖掘.



罗平(1979—),男,博士,副研究员,CCF 会员,主要研究领域为迁移学习,推荐系统,数据挖掘,机器学习.



史忠植(1941—),男,研究员,博士生导师,CCF 会士,主要研究领域为人工智能.