

## 域自适应学习研究进展

刘建伟<sup>1</sup> 孙正康<sup>1</sup> 罗雄麟<sup>1</sup>

**摘要** 传统的机器学习假设测试样本和训练样本来自同一概率分布. 但当前很多学习场景下训练样本和测试样本可能来自不同的概率分布. 域自适应学习能够有效地解决训练样本和测试样本概率分布不一致的学习问题, 作为机器学习新出现的研究领域在近几年受到了广泛的关注. 鉴于域自适应学习技术的重要性, 综述了域自适应学习的研究进展. 首先概述了域自适应学习的基本问题, 并总结了近几年出现的重要的域自适应学习方法. 接着介绍了近几年提出的较为经典的域自适应学习理论和当下域自适应学习的热门研究方向, 包括样例加权域自适应学习、特征表示域自适应学习、参数和特征分解域自适应学习和多源域自适应学习. 然后对域自适应学习进行了相关的理论分析, 讨论了高效的度量判据, 并给出了相应的误差界. 接着对当前域自适应学习在算法、模型结构和实际应用这三个方面的研究新进展进行了综述. 最后分别探讨了域自适应学习在特征变换和假设、训练优化、模型和数据表示、NLP 研究中存在的问题这四个方面的有待进一步解决的问题.

**关键词** 域自适应学习, 迁移假设, 单源域-单目标域, 多源域-多目标域, 域自适应学习理论

**引用格式** 刘建伟, 孙正康, 罗雄麟. 域自适应学习研究进展. 自动化学报, 2014, 40(8): 1576-1600

**DOI** 10.3724/SP.J.1004.2014.01576

## Review and Research Development on Domain Adaptation Learning

LIU Jian-Wei<sup>1</sup> SUN Zheng-Kang<sup>1</sup> LUO Xiong-Lin<sup>1</sup>

**Abstract** Traditional supervised learning algorithms assume that the training data and the test data are drawn from the same probability distribution. But in many cases, this assumption is too simplified, and too harsh in light of modern applications of machine learning. Domain adaptation approaches are used to solve the problem that arises when the data distribution in the test domain is different from that in the training domain. Although the domain adaptation problem is a fundamental problem in machine learning, it only started gaining much attention very recently. In view of the theoretical and practical significance of domain adaptation methods, this paper summarizes the learning algorithm for domain adaptation. Firstly, the basic issues of domain adaptation and several important methods on domain adaptation are summarized. Next, learning theory and hot research direction on domain adaptation are described, including instance weighting based method, feature representation based method, parameter and feature decomposition based method, domain adaptation with multiple sources. Thirdly, the theoretical analysis for domain adaptation and the effective distribution metric learning are illustrated. At the same time, the error bounds of those algorithms are also presented. Fourthly, new research and development in three aspects on domain adaptation in recent years are reviewed, including learning algorithm, model structure and practical application. Finally, the problems to be solved in aspects of feature transform and assumption, optimization algorithm, data representation and model, and the problem to be solved in NLP are discussed.

**Key words** Domain adaptation learning, transfer hypothesis, single source-domain single target-domain, multiple sources-domain multiple targets-domain, domain adaptation learning theory

**Citation** Liu Jian-Wei, Sun Zheng-Kang, Luo Xiong-Lin. Review and research development on domain adaptation learning. *Acta Automatica Sinica*, 2014, 40(8): 1576-1600

收稿日期 2013-09-22 录用日期 2013-12-31  
Manuscript received September 22, 2013; accepted December 31, 2013

国家重点基础研究发展计划 (973 计划) (2012CB720500), 国家自然科学基金 (21006127), 中国石油大学 (北京) 基础学科研究基金 (JCXK-2011-07) 资助

Supported by National Basic Research Program of China (973 Program) (2012CB720500), National Natural Science Foundation of China (21006127), and Basic Subject Research Fund of China University of Petroleum (JCXK-2011-07)

本文责任编辑 刘成林

Recommended by Associate Editor LIU Cheng-Lin

1. 中国石油大学 (北京) 自动化研究所 北京 102249

1. Research Institute of Automation, China University of Petroleum, Beijing 102249

传统的机器学习算法中, 通常假设训练样本和测试样本来自同一概率分布, 然后设计相应的模型和判别准则对待测试的样例的输出进行预测. 但是实际上当前很多学习场景下训练样本的概率分布和测试样本的概率分布是不同的, 例如如何根据老用户的邮件信息设计一种面向新用户的垃圾邮件过滤系统、自然语言处理方面的情感分析、依存句法分析、跨语言处理等都属于域自适应学习问题. 如何在源域和目标域概率分布不一致的情况下进行学习即为域自适应学习问题. 因此, 域自适应学习的重

点在于如何克服源域分布和目标域分布不同, 实现目标域上的学习任务.

尽管域自适应学习有着广泛的应用, 但是直到最近几年才得到足够的重视. 目前关于域自适应学习产生了大量的理论研究成果, 提出了新的学习算法, 但是这些理论研究所涉及的领域庞杂, 如统计分类、自然语言处理、情感分析、机器翻译、气象分析等领域, 研究内容往往涉及域自适应学习的某一方面, 存在着概念及定义不统一的问题. 虽然目前国外涌现出大量关于域自适应学习的研究工作, 例如文献 [1] 对统计分类中的域自适应学习进行了综述; 文献 [2-4] 对域自适应学习的各种误差界理论进行了讨论; 文献 [5-6] 围绕域自适应学习方法进行了研究和改进; 文献 [7-8] 对多源域自适应学习问题进行了分析和讨论. 但这些文献对当前域自适应学习所涵盖的概念、学习方法分类、学习理论的论述并不全面, 不能对域自适应学习进行完整清晰地论述. 随着域自适应学习的快速发展, 有必要对域自适应学习方法和理论及近几年最新的相关研究成果进行全面的综述.

本文系统综述了域自适应学习理论、学习方法、研究进展和当下亟待解决的问题, 为进一步深入研究域自适应学习理论和拓展其应用领域奠定了一定的基础. 具体内容安排为: 第 1 节概述域自适应学习的基本概念和假设、分类和不同应用场景的学习问题; 第 2~5 节系统地阐述了当前域自适应学习的理论基础和研究进展. 分别从样例、特征表示和学习模型参数的角度对不同域自适应学习场景下的学习理论及方法进行了详细的归纳总结. 第 6 节对域自适应学习相关理论进行了分析, 讨论了高效的度量判据, 给出了不同场景下的学习误差界. 第 7 节对当前域自适应学习在学习算法、模型结构和实际应用这三个方面的研究新进展进行了综述; 第 8 节分别探讨了域自适应在特征变换和假设、训练优化、模型和数据表示、NLP 研究中存在的问题这四个方面的有待进一步研究解决的问题, 并指出未来发展方向; 第 9 节对全文进行总结.

## 1 域自适应学习概述

### 1.1 域自适应学习的定义及假设

域自适应学习又称为跨域学习. 常用于情感分析<sup>[9-10]</sup>、自然语言处理<sup>[11-12]</sup>、文本分类<sup>[13-15]</sup>、信息提取 (Information extraction)<sup>[16]</sup>、WIFI 定位<sup>[17]</sup>和视觉概念分类 (Visual concept classification)<sup>[18]</sup>等学习场景. 在域自适应学习中, 通常能够获得来自源域 (Source domain) 的已标识样例. 但是目标

域 (Target domain) 的样例无类标签. 或者只有少量的目标域已标识样例. 在实际学习场景中, 要获得目标域的样例的类标签需要耗费大量时间和昂贵的成本. 有时可能根本得不到目标域的样例的类标签. 简单地说, 域自适应学习就是根据源域的已学习到的知识对目标域的输出进行学习.

当前虽然出现了大量的域自适应学习方法和理论. 但对“域自适应学习问题”需要给出一个明确且具体的定义. 一般意义上的域自适应学习就是根据源域的已标识样例及目标域的先验知识对目标域模型进行学习. 分类 (Classification) 和回归 (Regression) 是域自适应学习中经常遇到的两类问题: 假设样例所在的输入空间为  $X \in \mathbf{R}^n$ , 则当样例对应的输出所在空间为  $Y \in \{c_1, c_2, \dots, c_k\}$ ,  $c_j \in \mathbf{N}$  时为域自适应分类学习问题, 当  $Y \in \mathbf{R}$  时为域自适应回归学习问题. 目前, 域自适应回归学习问题的研究还处于起步阶段, 基本还未开展研究. 为了更清晰地介绍域自适应学习问题, 本文对单源域-单目标域域自适应分类问题定义如下:

**定义 1 (域自适应学习).** 假设样例所在的输入空间为  $X \in \mathbf{R}^n$ , 样例对应的输出所在空间为  $Y$ ,  $D_S$  表示  $(X, Y)$  上的源域概率分布, 源域输出函数为  $f_S : X \rightarrow Y$ , 两者构成源域  $(D_S, f_S)$ ; 目标域记为  $(D_T, f_T)$ , 其中  $f_T$  为目标域输出函数,  $D_T$  为目标域概率分布. 域自适应学习是指在源域概率分布  $D_S$  和目标域概率分布  $D_T$  不同的情况下, 通过源域的输入-输出样本对集合  $\{(\mathbf{x}_1^S, \mathbf{y}_1^S), (\mathbf{x}_2^S, \mathbf{y}_2^S), \dots, (\mathbf{x}_{N_S}^S, \mathbf{y}_{N_S}^S)\} \in D_S$ 、目标域未标识样例集  $\{\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_{N_T}^T\} \in \chi$  和目标域先验知识对目标域的输出函数  $f_T$  进行学习.

单源域-单目标域域自适应回归学习问题和多源域-多目标域分类和回归学习问题可类似定义.

另外, 可以将域自适应学习的定义进行推广, 学习任务不仅仅局限于对输出函数  $f_T$  的学习, 还可以包含对目标域概率分布  $D_T$  的学习. 本文提出的广义域自适应学习强调的是对目标域  $(D_T, f_T)$  的学习, 定义如下:

**定义 2 (广义域自适应学习).** 在源域  $(D_S, f_S)$  和目标域  $(D_T, f_T)$  不同的情况下, 通过已学习的源域  $(D_S, f_S)$  的概率分布, 预测目标域  $(D_T, f_T)$  的概率分布.

可将定义 2 的广义域自适应学习分为三种情况: 1) 源域和目标域概率分布相同  $D_S = D_T$ , 输出函数不同  $f_S \neq f_T$ ; 2) 源域和目标域概率分布不同  $D_S \neq D_T$ , 输出函数相同  $f_S = f_T$ ; 3) 源域和目标域概率分布不同  $D_S \neq D_T$ , 输出函数也不同  $D_S \neq D_T$ . 情

况 1 和多任务学习相似, 但是多任务学习旨在提高所有单个任务的学习性能, 而域自适应学习则关注于对目标任务的学习. 情况 2 与重要性抽样方法的应用场景相似, 因此基于最小化经验风险的样例加权方法是处理该域自适应学习场景的最为有效的方法. 情况 3 并没有限定域的概率分布相同  $D_S = D_T$  或者输出函数相同  $f_S = f_T$ , 所以相对于前两种情况而言更具有一般性, 也更为复杂. 因此, 凡是满足上述 3 种情况中的任意一种情况, 都属于广义的域自适应学习问题. 反之, 任意一类域自适应学习必定属于上述 3 种情况中的一种. 需要注意的是为了简化问题, 域自适应学习往往是对目标域的输出函数  $f_T$  进行学习, 即定义 1 所指的域自适应学习中经常用到一些假设, 假设的成立与否, 对域自适应学习的性能有很大的影响. 但是在不同的应用场景中指的范畴不相同, 常常会产生歧义. 为了避免产生混淆, 下面对域自适应学习中的协变量迁移假设、概念迁移 (Concept shift) 假设和先验概率迁移 (Prior probability shift) 假设作如下定义:

**定义 3 (协变量迁移假设).** 对  $\forall \mathbf{x} \in X$ , 输出函数满足  $f_S(\mathbf{x}) = f_T(\mathbf{x})$ , 则称源域和目标域的分满足协变量迁移特性.

**定义 4 (概念迁移假设).** 对  $\forall \mathbf{x} \in X, \forall \mathbf{y} \in Y$ , 样例分布满足  $P_S(\mathbf{x}) = P_T(\mathbf{x})$ , 或者类标签分布满足  $P_S(\mathbf{y}) = P_T(\mathbf{y})$ , 则称源域和目标域的分满足概念迁移特性.

**定义 5 (先验概率迁移假设).** 对  $\forall \mathbf{x} \in X, \forall \mathbf{y} \in Y$ , 样例分布满足  $P_S(\mathbf{x}|\mathbf{y}) = P_T(\mathbf{x}|\mathbf{y})$ , 则称源域和目标域的分满足先验概率迁移特性.

以上三种假设, 特别是协变量迁移假设, 经常用于域自适应学习中. 当然在域自适应学习中这些假设并不一定满足, 例如输出函数可能存在  $f_S(\mathbf{x}) \neq f_T(\mathbf{x})$  的情况.

### 1.2 域自适应学习分类

下面分别从概率统计角度、应用场景角度和使用的学习方法角度对域自适应学习进行全面的分类和讨论. 首先根据源域和目标域概率分布情况可以对域自适应学习进行如下分类:

假设  $k$  个源域  $S_1, S_2, \dots, S_k$  上的样例-输出对的联合概率分布分别为  $P_{S_1}(X, Y), P_{S_2}(X, Y), \dots, P_{S_k}(X, Y)$ ,  $p$  个目标域  $T_1, T_2, \dots, T_p$  上的样例-输出对联合概率分布分别为  $P_{T_1}(X, Y), P_{T_2}(X, Y), \dots, P_{T_p}(X, Y)$ .

样本集的组成为:

1) 源域的样例-输出对:  $\{(\mathbf{x}_i^S \in X, \mathbf{y}_i^S \in Y) \in D_{S_1}\}_{i=1}^{N_{S_1}}, \{(\mathbf{x}_i^S \in X, \mathbf{y}_i^S \in Y) \in D_{S_2}\}_{i=N_{S_1}+1}^{N_{S_2}}, \dots,$

$\{(\mathbf{x}_i^S \in X, \mathbf{y}_i^S \in Y) \in D_{S_k}\}_{i=N_{S_{k-1}}+1}^{N_{S_k}}$ .  
 2) 目标域的样例-输出对:  $\{(\mathbf{x}_i^T \in X, \mathbf{y}_i^T \in Y) \in D_{T_1}\}_{i=1}^{N_{T_1}}, \{(\mathbf{x}_i^T \in X, \mathbf{y}_i^T \in Y) \in D_{T_2}\}_{i=N_{T_1}+1}^{N_{T_2}}, \dots, \{(\mathbf{x}_i^T \in X, \mathbf{y}_i^T \in Y) \in D_{T_p}\}_{i=N_{T_{p-1}}+1}^{N_{T_p}}$ .

在样本集中,  $X \in \mathbf{R}^n$  为样例集, 分类时, 输出  $Y \in \{1, 2, \dots, k\}$ ; 回归时,  $Y \in \mathbf{R}$ . 聚类和降维问题源域和目标域只含有输入  $X$ , 没有输出  $Y$ . 图 1 中, 根据源域和目标域概率分布不一致的各种情况, 将域自适应学习分为 12 类, 每类的严格数学定义如下:

1) 多源域多目标域联合概率分布一致域自适应学习

$$P_{S_1}(X, Y) = P_{S_2}(X, Y) = \dots = P_{S_k}(X, Y) = P_{T_1}(X, Y) = P_{T_2}(X, Y) = \dots = P_{T_p}(X, Y)$$

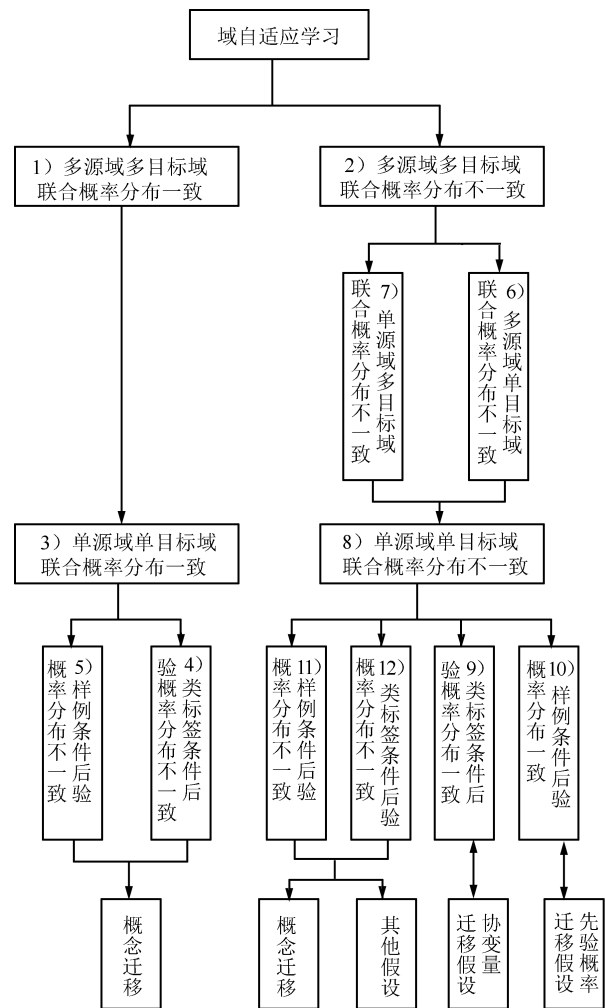


图 1 域自适应学习分类框图

Fig. 1 The classification chart for domain adaptation learning

2) 多源域多目标域联合概率分布不一致域自适应学习

$$P_{S_1}(X, Y) \neq P_{S_2}(X, Y) \neq \dots \neq P_{S_k}(X, Y) \neq P_{T_1}(X, Y) \neq P_{T_2}(X, Y) \neq \dots \neq P_{T_p}(X, Y)$$

3) 单源域单目标域联合概率分布一致域自适应学习

$$P_S(X, Y) = P_T(X, Y)$$

4) 类标签条件后验概率分布不一致域自适应学习

$$P_S(X, Y) = P_T(X, Y) \text{ 且 } P_S(Y|X) \neq P_T(Y|X)$$

5) 样例条件后验概率分布不一致域自适应学习

$$P_S(X, Y) = P_T(X, Y) \text{ 且 } P_S(X|Y) \neq P_T(X|Y)$$

6) 多源域单目标域联合概率分布不一致域自适应学习

$$P_{S_1}(X, Y) \neq P_{S_2}(X, Y) \neq \dots \neq P_{S_k}(X, Y) \neq P_T(X, Y)$$

7) 单源域多目标域联合概率分布不一致域自适应学习

$$P_S(X, Y) \neq P_{T_1}(X, Y) \neq P_{T_2}(X, Y) \neq \dots \neq P_{T_p}(X, Y)$$

8) 单源域单目标域联合概率分布不一致域自适应学习

$$P_S(X, Y) \neq P_T(X, Y)$$

9) 类标签条件后验概率分布一致域自适应学习

$$P_S(X, Y) \neq P_T(X, Y) \text{ 且 } P_S(Y|X) = P_T(Y|X)$$

10) 样例条件后验概率分布一致域自适应学习

$$P_S(X, Y) \neq P_T(X, Y) \text{ 且 } P_S(X|Y) = P_T(X|Y)$$

11) 样例签条件后验概率分布不一致域自适应学习

$$P_S(X, Y) \neq P_T(X, Y) \text{ 且 } P_S(X|Y) \neq P_T(X|Y)$$

12) 类标签条件后验概率分布不一致域自适应学习

$$P_S(X, Y) \neq P_T(X, Y) \text{ 且 } P_S(Y|X) \neq P_T(Y|X)$$

另一方面, 可定义域自适应学习中的正向和反向学习问题:

1) 正向学习问题:  $f : X \rightarrow Y$ , 由已知样例得到类标签的值;

2) 反向学习问题:  $f^{-1} : Y \rightarrow X$ , 由已知类标签得到样例的值.

在域自适应学习中, 所涉及的问题通常都属于  $f : X \rightarrow Y$  正向学习问题而很少属于  $f^{-1} : Y \rightarrow X$  反向学习问题, 因此协变量迁移假设在域自适应学习中应用广泛且更为常见, 而只在很少的域自适应学习场景中会涉及到先验概率迁移假设. 我们可以得到域自适应学习中的三种假设与这两类问题的关系, 具体如下:

1) 协变量迁移假设

属于正向学习问题:  $f : X \rightarrow Y$ , 此时,  $P_T(Y|X) = P_S(Y|X)$  且  $P_T(X) \neq P_S(X)$ .

2) 先验概率迁移假设

属于反向学习问题:  $f^{-1} : Y \rightarrow X$ , 此时,  $P_T(X|Y) = P_S(X|Y)$   $P_T(Y) \neq P_S(Y)$ .

3) 概念迁移假设

包含正向学习问题:  $f : X \rightarrow Y$ , 此时,  $P_T(Y|X) \neq P_S(Y|X)$  且  $P_T(X) = P_S(X)$ .

包含反向学习问题:  $f^{-1} : Y \rightarrow X$ , 此时,  $P_T(X|Y) \neq P_S(X|Y)$   $P_T(Y) = P_S(Y)$ .

4) 其他假设

包含正向学习问题:  $f : X \rightarrow Y$ , 此时,  $P_T(Y|X) \neq P_S(Y|X)$  且  $P_T(X) \neq P_S(X)$ .

包含反向学习问题:  $f^{-1} : Y \rightarrow X$ , 此时,  $P_T(X|Y) \neq P_S(X|Y)$   $P_T(Y) \neq P_S(Y)$ .

从不同的学习场景看, 域自适应学习可分为: 分类问题、回归问题和聚类问题. 降维是处理域自适应学习问题的一种重要手段, 因此有必要将域自适应降维作为单一独立的主题, 与域自适应分类、域自适应回归和域自适应聚类一起进行讨论. 表 1 对这四方面重要内容进行了列举. 关于域自适应分类学习问题的研究比较多, 也有部分关于域自适应降维问题<sup>[19]</sup> 和域自适应回归问题<sup>[20]</sup> 的研究, 但是关于域自适应聚类的研究还很不充分.

对于多源域和多目标域的学习, 不同类的域自适应学习所侧重的应用场景也不同. 多源域自适应学习多用于多示例学习场景和多视图学习场景. 而多目标域自适应学习多用于多标签学习场景、结构化输出、偏序学习和排序学习等场景.

根据源域和目标域的观点, 自适应学习理论和方法由两部分构成: 单源域单目标域自适应学习和多源域多目标域自适应学习. 如果能够对单源域-单

表 1 不同场景下的域自适应学习

Table 1 Different settings of domain adaption learning

学习场景	学习内容
分类问题	通过源域有类标签的样本和目标域无类标签的样例, 来获得性能优良的目标域分类器, 实现跨域的学习任务, 其中类标签 $y_i$ 取有限离散值 $y_i \in \{c_1, c_2, \dots, c_C\}$ , $c_j \in \mathbf{N}$
回归问题	根据有类标签的源域样本和无类标签的目标域样例, 通过训练获得性能良好的目标域回归器, 其中输出 $y_i$ 取连续值, $y_i \in \mathbf{R}$
聚类问题	根据最大化类内样例的相似性, 最小化类间的相似性准则进行聚类, 得到目标域样例的聚类簇划分高维
降维问题	通过对样例的特征空间维数进行约减, 找到合适的高维输入数据的低维结构, 在保持原始数据结构和成对约束 (Pair-wise constraints) 不变的情况下, 减少源域和目标域之间的概率分布不一致影响

目标域实现域自适应学习, 那么通过组合等方法可以扩展为多源域-多目标域的域自适应学习. 即单源域-单目标域学习理论和方法可视为域自适应学习的“基石”, 多源域-多目标域学习可以看作“扩展”.

单源域-单目标域的域自适应学习分为样例加权域自适应学习、特征表示域自适应学习、基于特征和参数分解的域自适应学习. 样例加权域自适应学习对训练样本附加权值, 使加权后的训练样本的概率分布逼近目标域的概率分布, 根据域自适应学习权值的计算方法的不同, 将样例加权域自适应学习分为: 重要性抽样方法、核均值匹配方法和 Adaboost 方法; 特征表示域自适应学习是将样例映射到新的特征空间中, 通过选择合适的特征表示方式, 使源域和目标域在新的特征空间下的分布相同或尽可能相同. 基于特征表示的域自适应学习算法多种多样, 最重要的两种方法是基于核映射函数的方法和结构对应学习, 此外还包含协同聚类方法、维数约简方法、迁移分量分析方法. 其中基于核映射函数的方法又包含最小化均值偏差和最小化离散度偏差两种映射方法; 基于特征和参数分解的方法的基本思路是对特征和模型参数分解, 这种基于特征和模型参数分解的方法是处理域自适应学习问题的另一类有效方法. 分解方法包括增广特征和模型参数分解两种形式. 通过适当的处理方法 (例如组合加权等方法), 可以将单源域-单目标域学习理论有效的扩展为多源域-多目标域的情况. 目前解决多源域自适应学习问题的方法可分为三大类: 基于源域判别函数学习的方法、基于正则化项加权组合的方法和基于流形平滑性假设的方法.

通过图 2 的域自适应学习方法的分类框图, 能够更全面清晰地了解域自适应学习理论和方法的构成, 这也是第 2~5 节的结构框架. 图 1 中域自适应学习分类框图的概念意义上的分类, 图 2 中域自

适应学习方法为目前机器学习学者研究过的自适应学习方法, 图 1 中“其他假设”中包含的反向学习问题:  $f^{-1}: Y \rightarrow X, P_T(X|Y)$  且  $\neq P_S(X|Y)P_T(Y) \neq P_S(Y)$  和概念迁移中的反向学习问题, 都是当前域自适应学习尚未尝试的方面, 也是未来有待研究的内容.

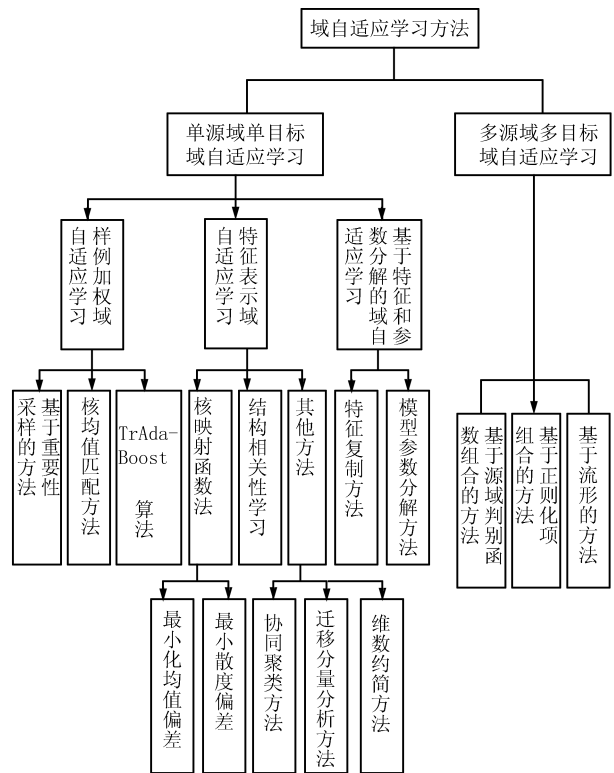


图 2 域自适应学习方法分类框图

Fig. 2 The classification chart for domain adaptation learning methods

## 2 样例加权域自适应学习

样例加权域自适应学习是一类应用广泛的域自适应学习方法, 由于学习算法和重要性抽样相似,

也被称为重要性加权域自适应学习 (Importance weighting). 样例加权域自适应学习通过对源域的样本附加权, 训练出目标域的误差最小参考模型  $h$ , 实现源域概率分布到目标域概率分布的转换. 样例加权域自适应学习的核心问题是如何计算“附加权”, 目前的方法有: 重要性抽样方法、核均值匹配方法和 AdaBoost 方法. 重要性抽样的方法主要用于计算先验概率迁移假设下的权值问题; 核均值匹配法适用于计算协变量迁移假设下的权值问题; AdaBoost 方法常用于已知目标域的部分已标识样例的权值学习问题.

## 2.1 基于重要性抽样的方法

令  $X \in \mathbf{R}^n$  为样例集, 分类问题: 输出  $Y \in \{1, 2, \dots, k\}$ ; 回归问题:  $Y \in \mathbf{R}$ . 样例-类标签对为  $\mathbf{z} = (\mathbf{x}, \mathbf{y})$ . 源域已标识样例-类标签集为  $S = \{(\mathbf{x}_i^S, \mathbf{y}_i^S)\}_{i=1}^{N_S}$ , 目标域样例-类标签集为  $T = \{(\mathbf{x}_i^T, \mathbf{y}_i^T)\}_{i=1}^{N_T}$ .  $P_S(\mathbf{x}, \mathbf{y})$  表示源域中  $X$  和  $Y$  的联合概率分布,  $P_T(X, Y)$  表示目标域中  $X$  和  $Y$  的联合概率分布,  $P_S(X)$ ,  $P_S(Y)$  表示源域中  $X$  和  $Y$  的边缘分布,  $P_T(X)$ ,  $P_T(Y)$  表示目标域中  $X$  和  $Y$  的边缘分布. 重要性加权域自适应学习的基本观点是在训练过程中对每一个样本附加权值, 经过加权后, 使参考模型  $h$  下的误差  $\hat{R}_w(h)$  最小:

$$\hat{R}_w(h) = \sum_{i=1}^m w_i c(\mathbf{z}_i, h) \quad (1)$$

其中,  $\hat{R}_w(h)$  是经过加权后的目标域训练误差,  $\mathbf{z} = (\mathbf{x}_i, \mathbf{y}_i)$ ,  $c(\mathbf{z}_i, h)$  是参考模型  $h$  下样本  $\mathbf{z}_i$  的损失函数. 可以证明, 当权  $w_i = P_T(\mathbf{z}_i)/P_S(\mathbf{z}_i)$  时, 源域的加权训练经验误差  $\hat{R}_w(h)$  是真实的目标域分布误差  $R(h) = \mathbb{E}_{\mathbf{z} \sim D_T} [c(\mathbf{z}, h)]$  的无偏估计, 即:

$$\mathbb{E}_{\mathbf{z} \sim D_S} [\hat{R}_w(h)] = \mathbb{E}_{\mathbf{z} \sim D_T} [c(\mathbf{z}, h)] = R(h) \quad (2)$$

假定  $H$  表示参考模型假设函数类, 我们从假设函数类  $H$  中选择目标域分布  $P_T(X, Y)$  中最优的模型  $h^* \in H$ :

$$h^* = \arg \min_{h \in H} \sum_{\mathbf{z}_i \in X \times Y} P_T(\mathbf{x}, \mathbf{y}) c(\mathbf{z}_i, h) \quad (3)$$

其中,  $c(\mathbf{z}_i, h)$  为样本  $\mathbf{z}_i$  的损失函数. 由式 (2) 可以将上述问题转化为求源域加权分布的最优模型:

$$h^* = \arg \min_{h \in H} \sum_{\mathbf{z}_i \in X \times Y} P_S(\mathbf{x}, \mathbf{y}) c(\mathbf{z}_i, h) \approx \arg \min_{h \in H} \sum_{i=1}^{N_S} \frac{P_T(\mathbf{x}_i^S, \mathbf{y}_i^S)}{P_S(\mathbf{x}_i^S, \mathbf{y}_i^S)} c(\mathbf{x}_i^S, \mathbf{y}_i^S, h) \quad (4)$$

通过对源域的样例进行加权, 可以从假设类  $H$  中得到最优的参考模型  $h$ , 使源域概率分布尽量接近目标域概率分布. 因此, 根据每一对样本  $(\mathbf{x}_i^S, \mathbf{y}_i^S)$  和  $P_T(\mathbf{x}_i^S, \mathbf{y}_i^S)/P_S(\mathbf{x}_i^S, \mathbf{y}_i^S)$ , 可以对目标域进行准确的学习, 域自适应学习的关键问题转变为计算每一对样例  $(\mathbf{x}, \mathbf{y})$  的  $P_T(\mathbf{x}, \mathbf{y})/P_S(\mathbf{x}, \mathbf{y})$  的值, 即“权”的计算问题.

在满足先验概率迁移假设下,  $P_S(X|Y = \mathbf{y}) = P_T(X|Y = \mathbf{y})$ ,  $P_S(Y) \neq P_T(Y)$ , 权值公式变为  $P_T(\mathbf{x}, \mathbf{y})/P_S(\mathbf{x}, \mathbf{y}) = P_T(\mathbf{y})/P_S(\mathbf{y})$ . 该假设下可以对源域进行重采样来近似目标域的标签分布. 虽然可以通过重采样的方法得到相应的权值, 但是在实际的域自适应学习中, 很少涉及到先验概率迁移假设问题, 而是讨论更为常见的协变量迁移假设域自适应学习问题.

## 2.2 核均值匹配法

在满足协变量迁移假设  $P_S(Y|X = \mathbf{x}) = P_T(Y|X = \mathbf{x})$ ,  $P_S(X) \neq P_T(X)$  条件下, 可以采用核均值匹配 (Kernel-mean matching, KMM) 方法来计算权值. KMM 方法通过在再生核希尔伯特空间 (Reproducing-kernel Hilbert space, RKHS) 中对源域和目标域样例进行均值匹配, 来直接获得权值  $P_S(\mathbf{x})/P_T(\mathbf{x})$ , KMM 可以归结为如下的二次规划问题:

$$\begin{aligned} \min_{\boldsymbol{\beta}} & \frac{1}{2} \boldsymbol{\beta}^T \mathbf{K} \boldsymbol{\beta} - \boldsymbol{\kappa}^T \boldsymbol{\beta} \\ \text{s.t.} & \beta_i \in [0, B] \\ & \left| \sum_{i=1}^{n_S} \beta_i - n_S \right| \leq n_S \varepsilon \end{aligned} \quad (5)$$

其中,  $\mathbf{K} = \begin{bmatrix} K_{SS} & K_{ST} \\ K_{TS} & K_{TT} \end{bmatrix}$ ,  $K_{SS}$  和  $K_{TT}$  为源域和目标域的核矩阵,  $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$ ,  $\boldsymbol{\kappa}_i = \frac{N_S}{N_T} \times \sum_{j=1}^{N_T} k(\mathbf{x}_i, \mathbf{x}_j^T)$ ,  $\mathbf{x}_i \in X_S \cup X_T$ ,  $\mathbf{x}_j^T \in X_T$ .

上述二次规划问题中,  $\boldsymbol{\beta}$  即为协变量迁移假设下的权值向量, 其中  $\beta_i = P_T(\mathbf{x}_i^S)/P_S(\mathbf{x}_i^S)$ . 使用内点法或其他二次最优化算法求解该问题即可直接得到域自适应学习的权值. KMM 算法避免了估计  $P_T(X)$  和  $P_S(X)$  的概率分布函数.

核均值匹配法能够有效地求解协变量迁移假设下的权值学习问题, 除此之外, 还有许多文献对协变量假设的权值学习问题提出了相应的解决方法. Zadrozny<sup>[21]</sup> 将密度估计问题转化为预测样例是来自源域还是目标域的问题, 通过构建简单的分类问题, 分别估计  $P_T(X)$  和  $P_S(X)$ , 得到相应的概率密

度比(即权值). 但实际应用中, 当处理样本的维数过高时, 算法的精度和效率并不理想, 且当样本数目较少时无法对  $P_T(X)$  和  $P_S(X)$  进行较为精确的估计, 权值估计误差较大; Sugiyama 等<sup>[22]</sup> 提出基于 K-L 散度最小化方法计算  $P_T(X)/P_S(X)$  的值, 其基本思路为: 假设权值的估计为  $\hat{w}(\mathbf{x})$ , 在该权值下的目标域样例的分布估计为  $\hat{P}_T(\mathbf{x}) = \hat{w}(\mathbf{x})P_S(\mathbf{x})$ , 通过最小化  $\hat{P}_T(\mathbf{x})$  与样例真实分布  $P_T(\mathbf{x})$  之间的 K-L 散度, 可以得到权值的最优估计. 该方法避免了直接对密度进行估计, 且优化问题为凸规划问题, 计算成本也更小. KMM 算法作为协变量迁移假设下计算权值的一种典型算法, 相对于密度估计这类权值计算方法<sup>[23]</sup> 具有算法精度高、计算成本低等优点, 且与全新的度量学习理论相结合可以得到多种多样的扩展形式<sup>[24]</sup>, 具有较大的研究价值.

### 2.3 基于 Adaboost 算法的域自适应学习

TrAdaboost 算法是 Adaboost 算法的一种扩展, 由 Dai 等提出<sup>[25]</sup>. TrAdaboost 算法将源域中的样本分为两类: 1) 对目标域学习“有利”的样本; 2) 对目标域学习“无用甚至有害”的样本. 其思路和重采样算法相似, 通过不断地迭代加权来减少源域的“有害”样本对目标域学习的影响, 同时增加对目标域学习“有利”的样本的权, 使源域概率分布逐渐转变为目标域概率分布.

TrAdaboost 算法需要利用少量的目标域已标识样例来构建目标域分类器, 在每一次迭代中都会训练一个带权的基分类器 (Base classifier), 如果源域中某一训练样例产生错误预测, 那么必然会和目标域已标识样例的输出预测产生冲突, 这表明该训练样例是“有害”样例, 那么算法会将该样例的权值向量  $\mathbf{w}$  乘以修正参数  $\beta_i \in (0, 1]$  来减少该样例的权值. 在下一迭代中该“有害”样例对目标域学习的影响将会减小, 最终经过多次迭代, 就可以实现对目标域的学习. Dai 等在文献 [25] 中证明了 TrAdaboost 算法的预测误差有上界, 并在 20News-groups、SRAA 和 Reuters-21578 三个文本数据集及其他非文本数据集进行实验验证, 表明 TrAdaBoost 算法在处理域自适应学习问题时具有理想的学习效果, 同时能够和传统的 AdaBoost 算法一样保证算法的收敛性.

Xu 和 Sun 在文献 [26] 中提出了 MV-TLAdaboost 算法, 该算法将多视图学习与 Adaboost 算法相结合, 从不同视图对源域和目标域进行学习, 实现了良好的域自适应学习. 文献 [27] 通过引入多个源域, 将 MV-TLAdaboost 算法扩展为 MsTL-MvAdaboost 算法, 改进后的算法能够有效

地避免负迁移并提高学习精度.

## 3 特征表示域自适应学习

特征表示域自适应学习是将样例映射到新的特征空间中, 通过选择合适的特征表示方式, 使源域和目标域在新的特征空间下的分布相同或尽可能相同. 基于特征表示的域自适应学习算法多种多样, 最主要的处理手段有基于映射函数的方法和结构对应学习 (Structural correspondence learning, SCL). 另外, 协同聚类、维数约简、迁移分量分析这些方法也可以应用到某些特定的域自适应学习算法中. 特征表示域自适应学习的关键问题为: 如何有效地度量不同域的分布相似度和差异; 如何找到合适的特征变换方法将原始特征空间  $X$  映射为新的特征空间  $\mathcal{X}$ .

### 3.1 基于核映射的特征表示域自适应学习

基于核映射的特征表示域自适应学习主要观点是通过寻找合适的特征映射, 将源域和目标域的样本映射到高维特征空间中, 使源域和目标域在新的特征空间下的分布尽可能相同 (偏差最小), 然后再建立相应的学习模型. 这种基于核映射函数的特征表示域自适应学习在部分论文中也常称为域自适应度量学习 (Domain adaptation metric learning, DAML), 如何确定有效的分布偏差度量判据是主要的研究课题.

域自适应度量学习通常将数据依赖的正则化项引入到再生核希尔伯特核空间中, 通过在希尔伯特核空间中最小化最大均值偏差, 使源域和目标域在 RKHS 中具有相近的分布, 解决两域的概率分布不同的问题.

定义  $\mathbf{x}_i$  和  $\mathbf{x}_j$  在核空间的欧氏距离为

$$\text{Dist}(\mathbf{x}_i, \mathbf{x}_j) = \|\varphi(\mathbf{x}_i) - \varphi(\mathbf{x}_j)\| \quad (6)$$

其中,  $\varphi: X \in \mathbf{R}^n \rightarrow \mathbf{R}^d$  为把样例映射到 RKHS 特征空间的变换函数. 通常  $n \ll d$ . 假定源域的判别函数为  $f_S: X \rightarrow Y$ , 目标域分类判别函数记为  $f_T$ . 首先要保证样例经非线性映射后在核空间中满足: 1) 具有相同类标签的样例对具有最小的期望偏差距离 (Discrepancy distance); 2) 具有不同类标签的样例对具有最大的期望偏差距离, 例如目标域 DAML 目标函数可定义为

$$\min_{\varphi \in F} \mathbb{E}_{\mathbf{x}_1 \sim D_T, \mathbf{x}_2 \sim D_T, f_T(\mathbf{x}_1) = f_T(\mathbf{x}_2)} \|\varphi(\mathbf{x}_1) - \varphi(\mathbf{x}_2)\|^2 \quad (7)$$

$$\text{s.t.} \quad \mathbb{E}_{\mathbf{x}_1 \sim D_T, \mathbf{x}_2 \sim D_T, f_T(\mathbf{x}_1) \neq f_T(\mathbf{x}_2)} \|\varphi(\mathbf{x}_1) - \varphi(\mathbf{x}_2)\| \geq 1 \quad (8)$$

实际上, 目标域分类判别函数  $f_T$  是未知的, 但是如果仅仅使用源域分类判别函数  $f_S$  来定义最优判别准则显然是不合理的. 此外还应该保证样例经过非线性映射  $\varphi$  后, 源域和目标域的分布特性差异显著减小, 因此加入数据依赖正则化项  $\Omega(\varphi)$  来约束解空间, 通过  $\Omega(\varphi)$  的最小化使核空间下的源域分布和目标域分布差异达到最小. 引入数据依赖正则化项  $\Omega(\varphi)$  的 DAML 问题为

$$\begin{aligned} \min_{\varphi \in F} \quad & \mathbb{E}_{\mathbf{x} \sim Q, \mathbf{z} \sim Q, f_Q(\mathbf{x}) \neq f_Q(\mathbf{z})} \|\varphi(\mathbf{x}) - \varphi(\mathbf{z})\|^2 + \lambda \Omega(\varphi) \\ \text{s.t.} \quad & \mathbb{E}_{\mathbf{x} \sim Q, \mathbf{z} \sim Q, f_Q(\mathbf{x}) \neq f_Q(\mathbf{z})} \|\varphi(\mathbf{x}) - \varphi(\mathbf{z})\|^2 \geq 1 \end{aligned} \quad (9)$$

上述 DAML 问题的经验估计形式为

$$\min_{\varphi \in F} \sum_{i,j=1}^m I_{y_i=y_j} \|\varphi(x_i) - \varphi(x_j)\|^2 + \lambda \Omega(\varphi) \quad (10)$$

$$\text{s.t.} \quad \sum_{i,j=1}^m I_{y_i=y_j} \|\varphi(x_i) - \varphi(x_j)\| \geq 1 \quad (11)$$

DAML 的本质是通过利用  $X_S$  中的已标记样例和  $X_T$  的概率分布信息来获得最优的非线性映射函数  $\varphi$ . 由于样本个数和算法的时间复杂性成比例, RKHS 特征空间的维数通常较高, 导致学习算法复杂性过高. 使用核主分量分析 (Kernel principle component analysis, KPCA) 技术来选择主分量能够有效减少 DAML 的算法复杂性.

当前的 DAML 算法中, 通常使用分布的均值偏差作为正则化项  $\Omega(\varphi)$ . 随着域自适应学习的研究不断深入, 为了获得更为准确的域自适应学习效果, 提出引入分布散度偏差作为正则化项  $\Omega(\varphi)$ : 即同时将最小化源域和目标域的分布均值偏差和分布散度偏差作为度量判据. 对这种最小化分布均值偏差和分布散度偏差的域自适应问题, Tao 等<sup>[28]</sup> 提出了较为完整的学习框架.

和前面章节相一致,  $D_S$  表示源域的概率分布,  $X_S \in \mathbf{R}^n$  为源域的已标识样例集,  $D_T$  表示目标域的概率分布,  $X_T \in \mathbf{R}^n$  为源域的样例集. 核空间下的线性分类函数为  $f: f(\mathbf{x}) = \langle w, \varphi(\mathbf{x}) \rangle$ . 定义源域和目标域的均值分布偏差为

$$\gamma_{KM}(D_S, D_T)^2 = \left\| \int_{X_S} f_{\mathbf{x} \sim D_S}(\mathbf{x}) dD_S - \right.$$

$$\left. \int_{X_T} f_{\mathbf{x} \sim D_T}(\mathbf{x}) dD_T \right\|^2 \quad (12)$$

定义源域和目标域的分布散度偏差为

$$\gamma_{KS}(D_S, D_T) = \left| \int_{X_S} f_{\mathbf{x} \sim D_S}(\mathbf{x}) [f_{\mathbf{x} \sim D_S}(\mathbf{x})]^T dD_S - \int_{X_T} f_{\mathbf{x} \sim D_T}(\mathbf{x}) [f_{\mathbf{x} \sim D_T}(\mathbf{x})]^T dD_T \right| \quad (13)$$

分布度量判据为两者加权和:

$$\gamma_{KMS}(D_S, D_T) = (1 - \lambda) \gamma_{KM} + \lambda \gamma_{KS} \quad (14)$$

这里,  $\lambda \in [0, 1]$ . 学习的问题归结为在最小化分布度量偏差和损失函数  $R(\mathbf{x}_i^S, \mathbf{y}_i^S, f)$  情况下, 找到最优分类超平面  $f: f(\mathbf{x}) = \langle w, \varphi(\mathbf{x}) \rangle$ .

$$\begin{aligned} \arg \min_{k,f} \quad & \frac{1}{2} \|\mathbf{w}\|_k^2 + \gamma_{KMS}(D_S, D_T) + \\ & C \sum_{i=1}^n R(\mathbf{x}_i^S, \mathbf{y}_i^S, f) \end{aligned} \quad (15)$$

经过推导化简, 得到对应的域自适应核学习目标函数为

$$\arg \min_{k,f} \frac{1}{2} \boldsymbol{\beta}^T \Omega \boldsymbol{\beta} + C \sum_{i=1}^n R(\mathbf{x}_i^S, \mathbf{y}_i^S, f) \quad (16)$$

其中,  $\boldsymbol{\beta}$  是加权矢量,  $\mathbf{x}_i^S$  为源域已标识样例,  $\mathbf{x}_i^T$  为目标域未标识样例;  $\Omega$  为由关于  $\mathbf{x}_i^S$  和  $\mathbf{x}_i^T$  的核映射  $k(\mathbf{x}_i^S, \mathbf{x}_j^T) = \langle \varphi(\mathbf{x}_i^S), \varphi(\mathbf{x}_j^T) \rangle$  作为元素构成的半正定核矩阵.

基于核映射的域自适应学习具有广泛的适用性和研究价值, 通过合理改进可以得到形式多样、性能优良的域自适应学习算法. 例如文献 [29] 提出的 pbTL 算法, 将源域和目标域的特征集合划分为多个子集, 对每一特征子集进行核映射学习, 通过学习每一特征子集的最优核映射参数得到目标域中对应子集的分类器, 最后将这些子集分类器进行加权组合得到精确高效的目标域分类器. 再如文献 [30], 域自适应问题要求最小化不同域的分布偏差距离, 分类问题则需要最大化类间距离, 将两者进行统一即可得到性能优良的学习算法, 对这种跨域度量学习进行核化后, 能够有效地处理具有非线性模型的样本集上的域自适应学习问题.

### 3.2 结构对应学习

Blitzer 等<sup>[31]</sup> 提出了结构对应学习 (Structural correspondence learning, SCL), SCL 通过使用目



标域的未标识样例来提取一些对应特性, 从而减少域之间的差异性. 在结构对应学习中, 首先针对源域和目标域中未标识样例定义“关键特征”集合, 假设从样例中可学习  $m$  个关键特征, 这些关键特征作为模型向量, 构建  $m$  个分类器:

$$f_l(\mathbf{x}) = \text{sgn}(\mathbf{w}_l^T \cdot \mathbf{x}), \quad l = 1, \dots, m \quad (17)$$

记参数矩阵  $W = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m]$ . 对参数矩阵  $W$  进行奇异值分解  $W = \mathbf{U}\mathbf{D}\mathbf{V}^T$ , 得到投影矩阵  $\theta = U_{[1:h,:]}^T$ , 这里,  $h$  是用户选择的共有特征的数目, 其中  $\theta$  的列是矩阵  $W$  的左奇异向量. 将原始特征矩阵  $X$  映射到低维特征空间  $\theta \cdot X$ , 等价于将其他的特征用这些独立的关键特征表示, 然后利用原始特征  $X$  和新特征  $\theta \cdot X$ , 得到增广向量, 最后对增广向量学习分类器.

SCL 能够解决源域和目标域某些特性差异较大的情况下的域自适应学习问题. SCL 的优点在于可以对不同域的特征进行统一建模和实现归一化, 即使这些特征并不相同. 另外它并不需要任何目标域的已标记样本数据, 使其具有很高的实际应用价值. 但是如何确定和定义关键特征是 SCL 方法应用的难点, 且我们并不知道关键特征对最终实际学习性能的影响. 另外, 在特征处理 (设计分类预测器时) 中, 所有的特征都是统一对待的, 如果能够对关键特征和非关键特征进行有针对性的选择, 可进一步提高学习性能.

### 3.3 其他特征表示域自适应学习方法

前面介绍了特征表示域自适应学习中最为主要的两种学习方法, 实际上, 协同聚类、维数约减和迁移分量分析都可以用于特征表示域自适应学习中, 得到其他类型的域自适应学习算法. 例如在文本分类问题中, Dai 等<sup>[32]</sup> 基于协同聚类的域自适应方法对文本进行分类. 通过词在文本中出现的频率来得到相应的词聚类簇, 然后根据词聚类簇对文本进行文本特征聚类得到文本聚类簇. 尽管两个域之间的分布不同, 但是可以通过共有信息来进行协同聚类, 这里的共有信息为词聚类簇. 首先根据域内  $D_i$  样本通过传递类标签信息得到聚类簇的结构, 然后扩展到域外  $D_o$ , 得到相应的基于文本特征和基于词特征的聚类簇, 即两个域的共有信息, 然后通过协同聚类得到基于词的聚类簇. 过程如图 3 所示.

维数约减也是特征表示域自适应学习的方法之一, Pan 等<sup>[33]</sup> 在迁移学习中提出了维数约减的学习方法, 通过 MMDE (Maximum mean discrepancy embedding) 方法得到合适的低维空间, 减少不同域之间的分布差异. 但是 MMDE 方法的计算过程往

往较为复杂, 为解决这一问题, Pan 等<sup>[34]</sup> 随后又提出了迁移分量分析 (Transfer component analysis, TCA), TCA 能够在不同域中学习某些迁移分量, 然后由这些迁移分量张成子空间, 在子空间中不同域的样本具有相同或相近的分布. 然后对这些样本进行学习. 其本质仍是 MMDE 的思路, 但是具有更高的特征提取效率.

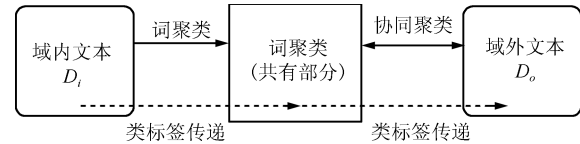


图 3 基于联合聚类的自适应学习

Fig. 3 Domain adaptation learning based on co-clustering

## 4 基于特征和参数分解的域自适应学习

基于特征和参数分解的域自适应学习与上一节的核映射域自适应学习有很多的共性, 但是基于特征和参数分解的域自适应学习的不同之处在于对特征和模型参数的处理方式, 其基本思路是将特征和模型参数分解为两部分: 1) 域“特有”特征或参数; 2) 域“共享”特征或参数. 这种基于特征和模型参数分解的方法是处理域自适应学习问题的另一类有效方法. 基于特征和参数分解的域自适应学习主要有增广特征和模型参数分解两种形式, 增广特征主要对样例进行特征复制产生新的样例的特征表示, 用样例的特征表示训练域自适应分类器; 模型参数分解则是把模型参数分解为域“特有”参数和域“共享”参数, 同时学习这两个模型参数.

### 4.1 特征复制方法

特征复制方法 (Feature duplication) 由 Daumé 提出<sup>[35]</sup>, 主要观点是对特征进行复制增广, 增广后的样例由域普通特征和域特殊特征构成. 如果原始输入空间为  $X = \mathbf{R}^n$ , 那么经过特征复制后得到的增广的输入空间为  $\chi = \mathbf{R}^{3n}$ . 映射  $\phi_S : X \rightarrow \chi$  和  $\phi_T : X \rightarrow \chi$  分别为源域和目标域从原始输入空间到增广输入空间的映射函数, 定义如下:

$$\phi_S(\mathbf{x}) = (\mathbf{x}, \mathbf{x}, 0)$$

$$\phi_T(\mathbf{x}) = (\mathbf{x}, 0, \mathbf{x})$$

源域上的映射向量  $(\mathbf{x}, \mathbf{x}, 0)$  的第 1 项表示源域和目标域的共有特征, 第 2 项表示源域的特殊特征, 第 3 项为目标域的特殊特征. 如果  $\mathbf{x}_i$  来自于目标域, 那么最后一项为 0. 相应的, 每一个目标域中样例经  $\phi_T$  映射后的表示为  $(\mathbf{x}_i, 0, \mathbf{x}_i)$ .

假设在源域中有大量已标识训练样例且源域中的每一个特征  $\mathbf{x}_i$  都可以复制, 在目标域中有少量的已标识样例. 对增广的样本进行学习, 通过最小化  $l_2$  范数得到增广样本上的分类器的模型向量. 具体方法为: 令  $\mathbf{w}_S$  为共有特征和源域特征的分类模型向量之和,  $\mathbf{w}_T$  为共有特征和目标域特征的分类模型向量之和,  $\mathbf{w}_g$  为参考模型向量, 通过选择  $\mathbf{w}_g$  使得  $\|\mathbf{w}_g\|^2 + \|\mathbf{w}_S - \mathbf{w}_g\|^2 + \|\mathbf{w}_T - \mathbf{w}_g\|^2$  最小化, 得到相应的模型向量  $\mathbf{w}_S$  和  $\mathbf{w}_T$ . 由于是同时优化  $\mathbf{w}_S$  和  $\mathbf{w}_T$ , 因此这比优化  $\|\mathbf{w}_S - \mathbf{w}_T\|^2$  的方案的学习效果要好.

为了对特征复制方法做出更为直观的解释, 讨论特征复制方法在核空间下的表示形式. 将样例  $\mathbf{x}$  进行非线性映射映射到 RKHS 空间, 源域和目标域的样例在未增广之前的核函数为  $K: X \times X \rightarrow \mathbf{R}$ , 重新定义  $\phi_S(\mathbf{x}) = (\phi(\mathbf{x}), \phi(\mathbf{x}), 0)$ ,  $\phi_T(\mathbf{x}) = (\phi(\mathbf{x}), 0, \phi(\mathbf{x}))$ , 那么增广后的样本在 RKHS 中的核  $\tilde{K}$  与初始样本在 RKHS 中的核  $K$  的关系为

$$\tilde{K}(x_i, x_j) = \begin{cases} 2K(x_i, x_j), & \text{源域已标识样例} \\ K(x_i, x_j), & \text{目标域已标识样例} \end{cases} \quad (18)$$

这表明 RKHS 下来自相同域的样本的相似度是来自不同域样本的相似性的两倍. Daumé 提出的这种域自适应学习方法相当于将模型参数  $\theta$  分解为  $\theta_{\text{General}} + \theta_{\text{Special}}$ , 其中,  $\theta_{\text{General}}$  是目标域和源域共享的模型参数,  $\theta_{\text{Special}}$  是每个域所特有的模型参数.

该方法的优点是算法容易实现且增广特征可应用于任意分类器. 但是学习性能依赖于目标域训练样例的个数. 这要求必须有一定数量的目标域的训练样例参与到学习中. 特征复制方法不适用于源域和目标域特有特性不相同但相互关联的学习, 且不能拓展到多个域的自适应学习. Blitzer 等通过源域和目标域的经验风险的凸组合最小化, 对这种算法的错误界进行了分析.

#### 4.2 模型参数分解域自适应学习

和特征复制方法的基本观点一样, 多任务支持向量机 (Support vector machine, SVM) 学习也是将原始的 SVM 模型参数分解为“域特有模型参数 + 域共享模型参数”的形式, 然后对这些模型参数进行求解. 多任务学习与域自适应学习的区别在于域自适应学习强调对目标域任务进行最优的学习, 而多任务学习强调利用多个任务之间的相关信息改善所有单个任务的学习性能, 实际上, 如果给目标域任务赋以很大的损失权来确保目标域的最优学习, 这时多任务学习可以近似看作域自适应学习. 模型

参数分解域自适应学习正是由多任务学习改进得到的, 因此和多任务学习的算法结构相类似.

假设源域学习任务的分类超平面为  $f_S = \mathbf{w}_S \cdot \mathbf{x}$ , 其模型向量参数为  $\mathbf{w}_S$ , 相应的目标域模型向量参数为  $\mathbf{w}_T$ . 源域有  $n_S$  个已标识样例, 目标域有  $n_T$  个已标识样例 (少量).

将源域模型参数  $\mathbf{w}_S$  和目标域模型参数  $\mathbf{w}_T$  分解为两部分:  $\mathbf{w}_S = \mathbf{w}_0 + \mathbf{v}_S$ ,  $\mathbf{w}_T = \mathbf{w}_0 + \mathbf{v}_T$ .  $\mathbf{w}_0$  是域共享模型参数,  $\mathbf{v}_S$  和  $\mathbf{v}_T$  为域特有的模型参数. 多任务学习关注的是  $N$  个任务中每一个任务  $t \in \{1, 2, \dots, N\}$  的最优化, 而域自适应学习只关注原域任务和目标域任务, 即  $t \in \{S, T\}$ , 因此目标函数只需要对源域任务和目标域任务进行优化, 模型参数分解域自适应学习问题如下:

$$\begin{aligned} \min_{\mathbf{w}_0, \mathbf{v}_t, \xi_{it}} J(\mathbf{w}_0, \mathbf{v}_t, \xi_{it}) &= \sum_{t \in \{S, T\}} \sum_{i=1}^{n_t} \xi_{ti} + \\ &\frac{\lambda_1}{2} \sum_{t \in \{S, T\}} \|\mathbf{v}_t\|^2 + \lambda_2 \|\mathbf{w}_0\|^2 \quad (19) \\ \text{s.t. } &\mathbf{y}_{ti}(\mathbf{w}_0 + \mathbf{v}_i) \cdot \mathbf{x}_{ti} \geq 1 - \xi_{ti}, \\ &\xi_{ti} \geq 0, \quad t \in \{S, T\} \end{aligned}$$

通过上述优化问题求解可以同时得到参数  $\mathbf{w}_0$ ,  $\mathbf{v}_S$ ,  $\mathbf{v}_T$ .

此外, 许多多任务学习算法也可以从参数分解角度改进为相应的域自适应学习算法. 例如文献 [36] 中基于高斯过程提出的高效的多任务学习算法 MT-IVM、文献 [37] 提出的分层贝叶斯学习、文献 [38] 提出的局部加权相似性学习等, 这些方法稍作改进后都可以看作基于参数的域自适应学习算法.

### 5 多源域自适应学习

在实际应用中, 域自适应学习并不一定只是单源域-单目标域的情况, 常常会出现多源域-多目标域的自适应学习问题. 如果能够得到性能良好的单源域-单目标域学习算法, 只需要稍加改进即可扩展为多源域-多目标域的情况. 因此解决这种更为广泛的多源域-多目标域的自适应问题的关键是如何将单源域-单目标域学习理论有效地扩展为多源域-多目标域的情况. 目前解决多源域自适应学习问题的方法可分为三大类: 1) 基于源域判别函数 (Prelearned classifiers) 学习的方法; 2) 矩阵正则化方法; 3) 基于流形平滑性假设的方法. 如果能够实现多源域-单目标域的自适应学习, 则其可用于多源域-多目标域的学习. 因此这两种情况可以归为如何实现多源域-单目标域情况下的学习问题.

上述方法 1~3 的相似点在于都是通过对各个源域进行加权组合来对目标域进行学习, 不同点是: 方法 1 对各个源域的判别函数进行加权组合; 方法 2 是将各个源域关于目标域学习的正则化项进行加权组合, 得到相应的正则化矩阵; 方法 3 则是利用流形平滑性假设, 首先产生单个源域的目标域的特征子空间 (或多个特征子空间), 然后将每个源域的目标域的子空间进行组合 (例如取均值), 将单源域-单目标域扩展为多源域-多目标域学习的情况。

### 5.1 基于源域判别函数组合的方法

基于源域判别函数组合的域自适应学习方法需要部分目标域的已标识样例. 假设有  $p$  个源域,  $D_l^T = \{(\mathbf{x}_1^T, \mathbf{y}_1^T), \dots, (\mathbf{x}_{n_l}^T, \mathbf{y}_{n_l}^T)\}$  表示目标域的已标识样本,  $D_u^T = \{\mathbf{x}_{n_l+1}^T, \dots, \mathbf{x}_{n_T}^T\}$  表示目标域中的未标识样例,  $f^T$  表示目标域输出函数.  $D^{S_i} = \{(\mathbf{x}_1^{S_i}, \mathbf{y}_1^{S_i}), \dots, (\mathbf{x}_{n_{S_i}}^{S_i}, \mathbf{y}_{n_{S_i}}^{S_i})\}$  为第  $S_i \in \{1, 2, \dots, p\}$  个源域的已标识样例,  $f^{S_i}$  表示第  $S_i$  个源域的判别函数. 并且假定  $f^{S_i}$  可以由源域已标识样例得到.

基于源域已学习分类器的多源域自适应学习的核心思想是: 先由每个源域的已标识样例学习得到源域的分类判别函数  $f^{S_i}$ ,  $S_i \in \{1, \dots, p\}$ , 对每个源域的分类器分配一个权值, 最后加权组合得到目标域的判别函数  $f^T$ :

$$f^T(\mathbf{x}) = \sum_{i=1}^P \gamma_{S_i} f^{S_i}(\mathbf{x}) + \Delta f(\mathbf{x}) \quad (20)$$

其中,  $\Delta f(\mathbf{x})$  为扰动函数, 由目标域已标识样例  $D_l^T$  得到, 用于度量目标域已标识样例对目标域判别函数的影响,  $\gamma_{S_i} \in [0, 1]$  是源域辅助判别函数的权重, 且  $\sum_{i=1}^P \gamma_{S_i} = 1$ .

通过非线性特征映射  $\phi(\cdot)$ , 将样例映射到核函数为  $\mathbf{k}(\cdot, \cdot) = \langle \phi(\cdot), \phi(\cdot) \rangle$  的核空间下, 扰动函数  $\Delta f(\mathbf{x})$  可以表示为  $\Delta f(\mathbf{x}) = \sum_{i=1}^{n_l} \alpha_i^T \mathbf{y}_i^T \mathbf{k}(\mathbf{x}_i^T, \mathbf{x})$ . 假定源域判别函数为  $f^{S_i}(\mathbf{x}) = \sum_{i=1}^{n_{S_i}} \alpha_i^{S_i} \mathbf{y}_i^{S_i} \mathbf{k}(\mathbf{x}_i^{S_i}, \mathbf{x})$ , 因此目标域判别函数变为

$$f^T(\mathbf{x}) = \sum_{i=1}^P \gamma_{S_i} \sum_{j=1}^{n_{S_i}} \alpha_j^{S_i} \mathbf{y}_j^{S_i} \mathbf{k}(\mathbf{x}_j^{S_i}, \mathbf{x}) + \sum_{j=1}^{n_l} \alpha_j^T \mathbf{y}_j^T \mathbf{k}(\mathbf{x}_j^T, \mathbf{x}) \quad (21)$$

其中,  $\alpha_j^{S_i}$  为第  $S_i$  个源域的第  $j$  个已标识样例的系数,  $\alpha_j^T$  为目标域第  $j$  个已标识样例的系数.

这种加权组合域自适应学习存在的问题是: 当测试样例数量较多时算法的学习效率较低, 即不适用于测试样例规模较大的情况; 该方法并没利用目

标域中有价值的未标识样例信息, 不能充分地利用所有已知信息来估计目标域的判别函数. 随着域自适应学习的发展, 这种基于判别函数组合的域自适应学习算法得到了极大的丰富和完善, 并能够进一步利用目标域所包含的信息进行跨域学习. 例如引入多视图学习、对多个源域和目标域进行特征划分、对特征子集判别函数进行加权组合等. 除了参与组合的判别函数形式多样, 在权值的处理方面, 可以将动态加权的观点应用到判别函数组合域自适应学习中<sup>[39]</sup>.

### 5.2 基于正则化项组合的方法

基于正则化项的方法和 DAML 的出发点是相同的, 即都是通过添加正则化项来将域自适应学习问题转化为最优化问题. 但是本节所提到的正则化方法是将每个源域关于目标域学习的正则化项进行加权组合, 得到相应的正则化矩阵, 进而将单源域-单目标域的特征表示域自适应学习扩展为多源域-多目标域场景下的域自适应学习问题. 为了和单源域-单目标域 DAML 进行区别, 将多源域-多目标域学习场景下的正则化项方法称为组合正则化项方法.

加权正则化方法相对于基于源域判别函数方法, 除了处理方法上存在差异外, 矩阵正则化方法还能够充分利用目标域的未标识样例信息, 通过多个正则化项的组合加权, 可以得到非常精确的多源域自适应学习模型. 与上一节相同, 假设有  $p$  个源域,  $D_l^T = \{(\mathbf{x}_1^T, \mathbf{y}_1^T), \dots, (\mathbf{x}_{n_l}^T, \mathbf{y}_{n_l}^T)\}$  表示目标域的已标识样本,  $D_u^T = \{\mathbf{x}_{n_l+1}^T, \dots, \mathbf{x}_{n_T}^T\}$  表示目标域中的未标识样例,  $f^T$  表示目标域输出函数. 多源域自适应学习目标函数如下:

$$\min_{f^T} \Omega(f^T) + \lambda_L \Omega_L(f^T) + \lambda_D \Omega_D(f^T) \quad (22)$$

其中,  $\Omega(f^T)$  用于衡量目标域判别函数  $f^T$  复杂度的正则化项;  $\Omega_L(f^T)$  是目标域判别函数  $f^T$  关于目标域已标识样例  $D_l^T$  的损失函数的正则化项. 该目标函数中的  $\Omega(f^T)$  和  $\Omega_L(f^T)$  与单源域-单目标域 DAML 中的正则化项是一致的, 但是组合正则化方法与 DAML 的区别在于正则化项  $\Omega_D(f^T)$ .  $\Omega_D(f^T)$  表示多个源域的正则化项的组合加权, 通过添加该正则化项可以实现更为准确的多源域-多目标域域自适应学习, 其具体形式如下:

$$\Omega_D(f^T) = \frac{1}{2} \sum_{s=1}^p \gamma_s \sum_{i=n_l+1}^{n_T} (f_i^T - f_i^S)^2 =$$

$$\frac{1}{2} \sum_{s=1}^p \gamma_s \left\| \mathbf{f}_u^T - \mathbf{f}_u^S \right\|^2 \quad (23)$$

其中,  $f_i^T = f^T(\mathbf{x}_i^T)$ ,  $f_i^S = f^S(\mathbf{x}_i^S)$ ,  $i \in \{n_l + 1, \dots, n_T\}$ ; 向量  $\mathbf{f}_u^T = [f_{n_l+1}^T, \dots, f_{n_T}^T]^T$  表示目标域未标识样例的判别函数值, 向量  $\mathbf{f}_u^S = [f_{n_l+1}^S, \dots, f_{n_T}^S]^T$  表示目标域未标识样例在源域判别函数  $f^S$  下的输出。

与 DAML 一样, 可以根据不同的损失函数选择不同的正则化项  $\Omega_L(f^T)$ , 如最小平方损失函数和铰链损失函数. 例如为了度量目标域判别函数  $f^T$  关于目标域已标识样例  $D_l^T$  的损失, 正则化项  $\Omega_L(f^T)$  可选为  $\Omega_L(f^T) = \sum_{i=1}^{n_l} (f_i^T - y_i^T)^2$ .

举例来讲, 假定目标域判别函数的形式为  $f^T(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x})$ , 正则化项选为  $\Omega(f^T) = \|\mathbf{w}\|^2 / 2$ , 取域独立正则化项为  $\Omega_A(f^T)$ , 则多源域-单目标域自适应学习对应的非稀疏学习问题为

$$\min_{f^T} \frac{\|\mathbf{w}\|^2}{2} + \frac{\lambda_L}{2} \sum_{i=1}^{n_l} (f_i^T - y_i^T)^2 + \frac{\lambda_D}{2} \sum_{S=1}^P \gamma_S \sum_{i=n_l+1}^{n_T} (f_i^T - f_i^S)^2 \quad (24)$$

同样的, 多源域-单目标域自适应学习对应的稀疏学习问题为

$$\min_{f^T} \frac{\|\mathbf{w}\|^2}{2} + C \sum_{i=1}^{n_T} (\xi_i + \xi_i^*) + \frac{\lambda_L}{2} \sum_{i=1}^{n_l} (f_i^T - y_i^T)^2 + \frac{\lambda_D}{2} \sum_{S=1}^P \gamma_S \sum_{i=n_l+1}^{n_T} (f_i^T - f_i^S)^2 \quad (25)$$

其中,  $\xi_i$  和  $\xi_i^*$  是松弛变量。

### 5.3 基于流形平滑性假设的学习方法

基于流形平滑性假设的学习方法能够实现更为复杂的域自适应学习, 例如目标域中没有任何已标识样例, 或者源域样例和目标域样例的维数、特征选择等性质不同的情况下的学习. 首先介绍基于流形平滑性假设的单源域单目标域学习, 然后扩展为多源域自适应学习。

基于流形的域自适应学习方法为: 将目标域和源域投影到 Grassmann 流形中的两个点, 然后建立测地线, 通过在测地线上进行抽样得到相应的中间子空间. 最后将获得的源域样例投影到这些子空间上并对目标域样例进行分类训练. 图 4 为学习过程示意图。

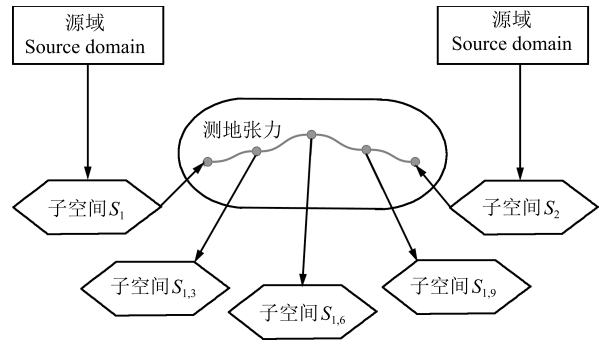


图 4 Grassmann 流形学习过程示意图

Fig. 4 Grassmann manifold learning sketch

令  $X_S = \{\mathbf{x}_i^S\}_{i=1}^{N_S}$  为源域已标识样例,  $X_T = \{\mathbf{x}_i^T\}_{i=1}^{N_T}$  为目标域未标识样例,  $S_1$  和  $S_2$  为  $X_S$  和  $X_T$  的投影子空间. 测地线为  $\psi(t)$ ,  $t \in [0, 1]$  是单变量的函数, 通过改变  $t$  的值可以获得测地线上不同的中间子空间. 将源域已标识样例  $\mathbf{x}_i^S$  在这些子空间上的投影  $\mathbf{x}_i^{S'}$  组成相应的向量  $\mathbf{x}_i^S$ , 所有已标识样例的投影向量构成投影矩阵  $X_l^S = [\mathbf{x}_1^S, \mathbf{x}_2^S, \dots, \mathbf{x}_{N_S}^S]$ , 相应地可以得到源域输出矩阵  $Y_l^S$ , 设计相应的分类器  $D(X_l^S, Y_l^S)$ . 最后用  $D(X_l^S, Y_l^S)$  对目标域的样例投影向量进行分类测试。

扩展到多源域的情形, 假设有  $K_1$  个源域和  $K_2$  个目标域, 分别生成源域子空间  $S_{11}, S_{12}, \dots, S_{1K_1}$  以及目标域子空间  $S_{21}, S_{22}, \dots, S_{2K_2}$ , 然后计算源域子空间的均值  $\bar{S}_1$  和目标域子空间的均值  $\bar{S}_2$ , 然后用上述方法处理即可。

## 6 域自适应学习理论分析

域自适应学习理论研究主要集中于: 1) 学习的可能性, 即满足什么条件才有可能实现域自适应学习; 2) 学习的鲁棒性, 域自适应学习的鲁棒性与普通意义上的鲁棒性是不同的: 在域自适应学习的鲁棒性中, 将训练样本分布和测试样本分布之间存在的“差异”视为“扰动”, 通过克服“扰动”, 实现对目标域的学习; 3) 学习的统计估计一致性, 域自适应学习算法应使用尽可能少的样本数, 具有尽可能好的泛化能力, 实现最小的误差上界。

### 6.1 域自适应学习的可能性

“满足哪些假设才能够对目标域进行有效学习”一直是域自适应学习需要讨论的问题之一. 其中, 学习机的学习能力、源域和目标域的相关性、域自适应学习的误差界分析和学习任务的先验知识这四点直接影响到域自适应学习的可能性, 如果不能很好地处理这些问题, 则不能实现有效的域自适应学习。

许多文献对域自适应学习的有效性进行了分析: 文献 [40] 提出了域自适应学习的不可能性定理, 并证明了在满足协变量迁移假设下, 对于假设 1, 源域和目标域的概率分布之间具有一定程度的相似性; 对于假设 2, 存在某一参考假设函数类, 其关于源域和目标域都有较小的估计误差, 如果不满足假设 1 或假设 2 中的任一条, 或者同时不满足假设 1 和假设 2, 那么任何的域自适应学习算法都是不可实现的. Ben-David 等在文献 [41] 中讨论了域自适应学习的硬度 (Hardness) 问题, 在协变量迁移假设下, 如果对目标域没有足够的先验知识, 并不能保证域自适应学习的有效性, 这主要是由于在协变量迁移假设下若想实现有效的目标域自适应学习需要的样本个数趋于无限大. 文献 [42] 详细讨论了域自适应学习的误差界, 并指出对于不同的域自适应学习算法, 源域的样本数量的增加并不一定能够提高学习质量.

下面从学习机学习能力、源域和目标域的相关性、域自适应学习误差和学习任务的先验知识这四个方面对域自适应的可学习性进行理论分析.

1) 学习机的学习能力: 学习机的学习能力是判断能否对目标域进行学习的影响因素之一. 对于域自适应学习, 如果训练得到的学习机的学习能力较差, 那么必然无法对域自适应问题进行学习. 域自适应学习机的学习能力的定义如下:

**定义 6 (学习机的学习能力).** 令  $m$  为源域已标识样本个数,  $n$  为目标域未标识样例个数,  $\varepsilon$  和  $\delta$  为任意正数,  $h: X \rightarrow Y$  为分类器模型参考假设函数,  $W$  为三元组  $(D_S, D_T, f_T)$  的集合, 其中  $D_S$  和  $D_T$  分别为源域概率分布和目标域概率分布,  $f_T: X \rightarrow Y$  为目标域判别函数,  $A(\varepsilon, \delta, m, n)$  为域自适应学习机. 对于  $\forall (D_S, D_T, f_T) \in W$ , 学习机由给定的样本学习得到的分类器假设函数  $h(\mathbf{x})$  以不小于  $1 - \delta$  的概率满足:  $P_{\mathbf{x} \sim D_T}(h(\mathbf{x}) \neq f_T(\mathbf{x})) \leq \varepsilon$ , 则称学习机  $A(\varepsilon, \delta, m, n)$  能够解决  $W$  类型的域自适应问题.

2) 源域和目标域的相关性: 源域和目标域的相关性是域自适应学习能否实现的另一衡量准则. 显而易见的是, 如果源域和目标域之间没有任何关联关系, 那么通过源域是不能实现对目标域的学习的, 即无法实现域自适应学习. 判断域自适应学习可实现性时必须对源域和目标域的相关性进行有效的度量, 如果度量准则选择不合适, 将无法对域相似性进行正确估计, 导致域自适应学习的失败. 常用的分布相似度量判据有变分距离、 $A$  距离等, 这些度量判据对域自适应学习的影响将在第 6.3 节中做更为详细的讨论.

3) 域自适应学习的误差: 域自适应学习必须具有理想的学习误差界, 算法的误差通常可以分为三种: 贝叶斯误差 (Bayes error)、近似误差 (Approximation error)、估计误差 (Estimation error), 对于域自适应学习而言, 还需要考虑另外一种误差: 分布偏离误差 (Distribution discrepancy error). 正是由于分布偏离误差的存在, 使得域自适应学习与普通的机器学习具有很大的差异, 域自适应学习中的假设函数不能仅仅满足对目标域具有较小的估计误差, 必须同时满足对源域和目标域都具有最小近似估计误差, 否则并不能保证域自适应学习的采样复杂度是一有界函数, 由于所需的采样尺寸趋于无穷, 将不能够实现域自适应学习.

4) 学习任务的先验知识: 域自适应学习中, 如果对目标域没有足够的先验知识, 即使源域和目标域具有很强的相关性 (例如满足协变量迁移假设), 仍不能保证域自适应学习的有效性. 实现不同学习场景下的域自适应学习需要的先验知识是不同的, 以 3) 为例, 域自适应学习需要一类假设函数  $H$ , 必须保证该类假设函数对源域和目标域同时具有最小估计误差, 因此必须满足这些先验假设才能实现域自适应学习.

## 6.2 域自适应学习的鲁棒性问题

假设样本空间为  $Z = X \times Y$ , 将输入空间  $X$  划分为  $K_x$  部分, 即  $X = \cup_{i=1}^{K_x} X_i$ ; 将输出空间  $Y$  划分为  $K_y$  部分, 即  $Y = \cup_{j=1}^{K_y} Y_j$ ; 将总样本空间  $Z$  划分为  $K$  部分, 即  $K = K_x \times K_y$ . 分布  $D$  的判别函数为  $f: X \rightarrow Y$ , 容许假设函数类  $H$  中的参考假设函数记为  $h: X \rightarrow Y$ . 参考假设函数  $h$  在样本点  $\mathbf{z} = (\mathbf{x}, \mathbf{y})$  的损失函数记为  $l(h, \mathbf{z}) \equiv l(h(\mathbf{x}), \mathbf{y})$ ; 参考假设函数  $h$  在分布  $D$  上关于判别函数  $f$  的损失函数记为  $L_D(h, f) = E_{\mathbf{x} \sim D}[l(h(\mathbf{x}), f(\mathbf{x}))]$ ; 两者的关系为  $L_D(h) = E_{(\mathbf{x}, \mathbf{y}) \sim D}[l(h(\mathbf{x}), \mathbf{y})]$ .

算法的鲁棒性首先由 Xu 和 Mannor<sup>[43]</sup> 提出, 算法鲁棒性定义为:

**定义 7.** 样本空间  $Z$  可以划分为  $K$  个不相交子集  $\{C_i\}_{i=1}^K$ ,  $S_1$  为从分布  $D$  中获得的已标识样本集合,  $\mathbf{s} \in S_1$  为  $S_1$  中任意的样本点, 如果  $\mathbf{s} \in C_k$ ,  $k \in \{1, 2, \dots, K\}$ , 对于任意  $\mathbf{z} \in C_k$  都满足:  $|l(h, \mathbf{s}) - l(h, \mathbf{z})| \leq \varepsilon$ , 则称算法  $A$  是满足  $(K, \varepsilon)$  的鲁棒性算法.

接下来给出  $\lambda$  迁移的定义.

**定义 8.** 令  $\sigma$  为类标签  $Y = \{y_1, \dots, y_r\}$  上的概率分布,  $y_v$  在分布  $\sigma$  上的概率记为  $\sigma^v$ , 剩余类标签的概率和记为  $\sigma^{-v}$ . 如果  $\forall y_v \in Y$ , 分布  $\rho$  和  $\sigma$  满足  $\rho^v \leq \sigma^v + \lambda \sigma^{-v}$  且  $\rho^v \geq \sigma^v (1 - \lambda)$ , 那么分布

$\rho$  为分布  $\sigma$  的  $\lambda$  迁移, 记为  $\rho \in \lambda(\sigma)$ .

在域自适应学习中, 由于源域和目标域的概率分布不同, 导致训练样例和测试样例分布不一致, 因此在源域上训练的模型往往不适用于目标域的学习. 域自适应学习的鲁棒性能够度量学习算法对训练样本改变的敏感程度, 故可以视为估计学习模型的泛化界的另外一种方法. 域自适应学习算法的鲁棒性评价方法把输入域 (Input domain)  $X$  划分为若干区域, 域自适应学习的鲁棒性假设认为在每一划分区域中学习算法的损失函数都具有有界损失方差, 且源域和目标域分布非常接近, 体现在每一划分区域上的源域分布和目标域分布满足  $\lambda$  迁移 ( $\lambda$ -shift) 关系.  $\lambda$  迁移是一种度量源域和目标域的相似性的方法, 通过  $\lambda$  迁移来解决目标域类别标签学习问题, 进而得到域自适应学习的目标域泛化误差界. 需要注意的是, 域自适应学习中的鲁棒性是将源域和目标域的概率分布距离视为“扰动”, 这与一般意义上的“扰动”是有区别的.

Mansour 等在文献 [44] 中提出了域自适应学习的鲁棒性问题. 通常情况下如果源域和目标域的概率分布非常接近, 例如具有很小的  $L_1$  范数距离, 通过  $\lambda$  迁移可以得到目标域学习机的泛化界, 其实质是将源域的经验风险误差与目标域的期望风险误差联系在一起. 此外  $A$  距离和相对偏差距离都可作为源域和目标域的概率分布相似度量函数.

对于域自适应学习问题, 源域分布  $D_S$  上的样本记为  $S = \{(x_i, y_i)\}_{i=1, \dots, m}$ , 目标域分布  $D_T$  上的未标识样例记为  $T = \{x_j\}_{j=1, \dots, n}$ . 由于可以得到源域的已标识样本集合  $S$ , 因此如果域自适应学习算法满足  $(K, \varepsilon)$  的鲁棒性, 则参考假设函数  $h$  关于源域分布的损失  $|L_{D_S}(h, f_S)|$  (即期望损失) 和参考假设函数  $h$  关于采样集合  $S$  的损失  $|L_S(h, f_S)|$  (即经验损失) 有如下关系:

$$|L_{D_S}(h, f_S) - L_S(h, f_S)| \leq \varepsilon + M \sqrt{\frac{2K \ln 2 + 2 \ln \frac{1}{\delta}}{|S|}} \quad (26)$$

其中,  $|S|$  为样本集  $S$  中样本的个数,  $f_S$  为源域判别函数, 通常损失函数  $L$  为非负有界函数,  $M$  为损失函数  $L$  的上界. 实际上, 损失  $|L_{D_S}(h, f_S)|$  和  $|L_S(h, f_S)|$  分别对应分布的期望损失和经验损失, 在后面一节中将具体讨论这两类损失的关系.

域自适应学习关注的是对目标域的学习, 需要得到目标域的泛化界. 通常域自适应学习中源域分布  $D_S$  和目标域分布  $D_T$  是不同但相似的, 因此可

假设  $D_S$  和  $D_T$  满足  $\lambda$  迁移特性, 那么根据式 (26) 可以推导出目标域的泛化误差界. 将  $Z = X \times Y$  划分为  $K = K_x \times K_y$  个不相交子集, 对于每一部分  $X_i, i \in \{1, \dots, K_x\}$ , 如果域自适应学习算法满足  $(K, \varepsilon)$  鲁棒性, 则参考假设函数下源域的期望损失为

$$L_{D_T}(h) \leq \varepsilon + M \sqrt{\frac{2K \ln 2 + 2 \ln \frac{1}{\delta}}{n}} + \sum_{i=1}^{K_x} P_T(X_i) l_S^\lambda(h, X_i) \quad (27)$$

其中,  $n$  为目标域未标识样例的个数,  $l_S^\lambda(h, X_i)$  为参考假设函数为  $h$  的情况下区域  $X_i$  上的最大平均损失,  $P_T(X_i)$  为区域  $X_i$  在目标域上的概率. 最后一项  $\sum_{i=1}^{K_x} P_T(X_i) l_S^\lambda(h, X_i)$  可以由源域已标识样本  $S = \{(x_i, y_i)\}_{i=1, \dots, m}$  和目标域未标识样例  $T = \{x_j\}_{j=1, \dots, n}$  估计得到.

本节所介绍的域自适应的鲁棒性是将分布之间的距离作为“扰动”, 通过参考假设函数  $h$  下源域的期望损失和经验损失的关系, 得到目标域的期望损失的泛化界. 在下一节中, 本文将结合度量判据, 详细地讨论由样例得到的经验损失函数和目标域期望损失函数之间的关系, 并分别给出相应的误差界.

### 6.3 域自适应学习的统计估计一致性

在能够保证对目标域进行学习的前提后, 另一个重要问题是如何进行有效的域自适应学习, 这往往依赖于有效的源域和目标域的概率分布度量判据, 必须保证域自适应学习算法具有合理的误差界, 本节将结合度量准则和算法误差界对域自适应学习的统计估计一致性进行理论分析.

第 6.3.1 节和第 6.3.2 节分别介绍域自适应学习的两类度量准则: 变分距离 ( $L^1$  散度) 和  $A$  距离 ( $H$  散度), 并给出对应准则下的目标域期望误差函数的误差界. 对于大多数域自适应学习, 虽然可以获得源域的已标识样例, 但是并不知道源域的真实分布, 这时最小化目标域期望误差并不是最优的选择, 而是需要对源域的期望分类误差和目标域的期望分类误差进行权衡. 针对这一问题, 第 6.3.3 节和第 6.3.4 节讨论了  $A$  距离度量准则下的单源域单目标域和多源域多目标域学习的组合学习的误差界. 最后第 6.3.5 节对重要性加权域自适应学习的误差界进行了分析, 讨论了权值有界和权值无界情况下的目标域期望误差和加权经验误差的关系.

假设样例空间为  $X \in \mathbf{R}^n$ , 源域记为  $(D_S, f_S)$ , 目标域记为  $(D_T, f_T)$ , 输出函数 (分类判别函数) 为

$f: X \rightarrow \{0, 1\}$ , 并且假设  $f_S, f_T$  分别为源域和目标域输出函数, 而容许假设函数类  $H$  中的参考假设函数  $h: X \rightarrow \{0, 1\}$ , 源域期望误差函数  $\epsilon_S(h, f_S) = E_{\mathbf{x} \sim D_S} [|h(\mathbf{x}) - f_S(\mathbf{x})|]$ .

### 6.3.1 变分距离 ( $L^1$ 散度) 下的误差界

假定  $D$  和  $D'$  为两个概率分布,  $B$  为分布  $D$  和  $D'$  上的可测子集, 如果使用变分散度距离  $d_1(D, D') = 2 \sup_{B \in \mathcal{B}} |p_D[B] - p_{D'}[B]|$  ( $L^1$  散度) 来度量不同概率分布的差异程度, 则在假设函数  $h$  下的目标域期望误差函数  $\epsilon_T(h)$  上界为

$$\begin{aligned} \epsilon_T(h) \leq & \epsilon_S(h) + d_1(D_S, D_T) + \\ & \min\{E_{D_S} [|f_S(\mathbf{x}) - f_T(\mathbf{x})|], \\ & E_{D_T} [|f_S(\mathbf{x}) - f_T(\mathbf{x})|]\} \quad (28) \end{aligned}$$

该误差界中, 第 1 项是假设函数  $h$  下源域的期望误差函数, 第 3 项为两个域的判别函数的差异产生的误差. 采用  $L^1$  散度估计误差界有两个缺点: 首先在有限样本的情况下不能得到  $L^1$  散度的准确估计值, 因而也就无法准确地估计目标域误差界; 其次由于  $L^1$  散度是所有可测子集的最小上确界 (Supremum) 距离, 因而会增大目标域误差上界, 所以给出的误差界估计比较保守.

### 6.3.2 $A$ 距离 ( $H$ 散度) 下的误差界

针对变分散度的不足, Ben-David 提出用  $A$  距离度量两个概率分布的距离, 并给出了  $A$  距离下的误差界, 首先定义  $A$  距离为  $d_A(D, D') = 2 \sup_{A \in \mathcal{A}} |P_{r_D}[A] - P_{r_{D'}}[A]|$ , 取  $A$  为参考假设函数类  $z$ , 则得到  $H$  散度距离为  $d_H(D, D') = 2 \sup_{h \in H} |P_{r_D}[I(h)] - P_{r_{D'}}[I(h)]|$ , 其中  $I(h)$  为参考假设函数  $h \in H$  的集合, 即如果  $\mathbf{x} \in I(h)$ , 则  $h(\mathbf{x}) = 1$ . 对于  $\forall h, h' \in H$  假设函数, 定义对称差假设空间  $H\Delta H$  为  $g \in H\Delta H \Leftrightarrow g(\mathbf{x}) = h(\mathbf{x}) \oplus h'(\mathbf{x})$ , 这里  $\oplus$  是对称差运算. 每一个  $g \in H\Delta H$  表示  $H$  中的两个假设  $h$  和  $h'$  不一致的集合, 源域  $D_S$  和目标域  $D_T$  的对称差距离为  $d_{H\Delta H}(D_S, D_T) = 2 \sup_{h, h' \in H} |P_{\mathbf{x} \sim D_S}[h(\mathbf{x}) \neq h'(\mathbf{x})] - P_{\mathbf{x} \sim D_T}[h(\mathbf{x}) \neq h'(\mathbf{x})]|$ . 使用  $H$  散度来度量不同概率分布的距离, 则对于  $\forall \epsilon_T(h)$ ,  $\exists \delta$ , 以至少  $1 - \delta$  的概率对所有  $h \in H$  以下不等式成立:

$$\begin{aligned} \epsilon_T(h) \leq & \epsilon_S(h) + \frac{1}{2} \hat{d}_{H\Delta H}(U_S, U_T) + \\ & 4 \sqrt{\frac{2d \log(2m') + \log\left(\frac{2}{\delta}\right)}{m'}} + \lambda \quad (29) \end{aligned}$$

得到  $H$  散度下的目标域期望误差函数  $\epsilon_T(h)$  的上

界. 其中  $d$  是假设空间  $H$  的 VC 维维数,  $U_S, U_T$  是分别对应源域  $D_S$  和目标域  $D_T$  的大小为  $m'$  的未标识样例的样本集概率分布函数,  $\hat{d}_{H\Delta H}$  为对称差距离的经验估计,  $\lambda = \epsilon_S(h^*) + \epsilon_T(h^*)$ , 其中  $h^* = \arg \min_{h \in H} \epsilon_S(h) + \epsilon_T(h)$  为使组合误差最小的理想假设函数.  $H$  散度解决了  $L^1$  散度的不足之处, 首先  $H$  散度可以由有限的样本计算得到, 其次对于任何假设函数类  $H$ ,  $H$  散度的误差界都不会超过  $L^1$  散度误差界, 得到的误差界更紧.

通过  $A$  距离能够有效地度量源域和目标域的分布距离, 下面将  $A$  距离应用于单源域单目标域学习场景及多源域多目标域学习场景, 并给出相应的域自适应学习的组合误差界.

### 6.3.3 单源域 - 单目标域的组合学习误差界

假定  $H$  是 VC 维维数为  $d$  的假设空间,  $S = (S_T, S_S)$  为包含  $m$  个已标识样例的样本集, 其中  $S_T$  为来自目标域  $D_T$  的  $\beta m$  个已标识样例集 ( $\beta$  通常很小),  $S_S$  为来自源域  $D_S$  的  $(1 - \beta)m$  个已标识样例集,  $U_S, U_T$  为个数为  $m'$  的未标识样例的样本集, 学习任务是找到使目标域期望误差  $\epsilon_T(h)$  最小的假设函数. 直接最小化目标域期望误差并不是理想的选择, 而是将源域和目标域凸组合的期望误差  $\epsilon_\alpha(h) = \alpha \epsilon_T(h) + (1 - \alpha) \epsilon_S(h)$  作为最小化目标函数,  $\alpha \in [0, 1]$  记凸组合的经验误差为  $\hat{\epsilon}_\alpha(h) = \alpha \hat{\epsilon}_T(h) + (1 - \alpha) \hat{\epsilon}_S(h)$ ,  $\hat{h} \in H$  是使组合经验误差  $\hat{\epsilon}_\alpha(h)$  最小的假设函数,  $h_T^* = \min_{h \in H} \epsilon_T(h)$  是使目标域误差最小的假设函数. 下面给出假设函数  $\hat{h}$  关于目标域的期望误差  $\epsilon_T(\hat{h})$  与  $h_T^*$  关于目标域的期望误差  $\epsilon_T(h_T^*)$  的关系. 对于  $\forall \epsilon_T(\hat{h})$ ,  $\exists \delta$ , 以至少  $1 - \delta$  的概率对所有  $h \in H$  以下不等式成立:

$$\begin{aligned} \epsilon_T(\hat{h}) \leq & \epsilon_T(h_T^*) + 4 \frac{\alpha^2}{\beta} + \frac{(1 - \alpha)^2}{1 - \beta} \times \\ & \frac{2d \log(2(m + 1)) + 2 \log\left(\frac{8}{\delta}\right)}{m} + \\ & 2(1 - \alpha) \times \left\{ \frac{1}{2} \hat{d}_{H\Delta H}(U_S, U_T) + \right. \\ & \left. \frac{2d \log(2m') + \log\left(\frac{8}{\delta}\right)}{4m'} + \lambda \right\} \quad (30) \end{aligned}$$

通过选择不同的  $\alpha$ , 可以对源域的样例个数和目标域样例个数进行合理的权衡, 得到最优的目标域假设函数. 当  $\alpha = 0$  (即忽略目标域数据) 时, 式 (30) 等价于式 (29) 加上源域的经验误差估计, 当  $\alpha$

= 1 时, 式 (30) 则变为只使用目标域数据学习误差界估计. 当  $\lambda = 0$  时, 式 (30) 对样本大小  $m$  的依赖性将得到较大的改善.

忽略式 (30) 的附加常数项, 将上述误差界表示为关于  $\alpha$  的函数:

$$f(\alpha) = 2B \sqrt{\left(\frac{\alpha^2}{\beta} + \frac{(1-\alpha)^2}{1-\beta}\right)} + 2(1-\alpha)A \quad (31)$$

其中,  $A = \{\frac{1}{2}\hat{\mathbf{d}}_{H\Delta H}(U_S, U_T) + 4\sqrt{\frac{2d\log(2m') + \log(\frac{\xi}{8})}{m'}} + \lambda\}$  为源域和目标域的总相似性度量.  $B = 2\sqrt{\frac{\alpha^2}{\beta} + \frac{(1-\alpha)^2}{1-\beta}} \times \sqrt{\frac{2d\log(2(m+1)) + 2\log(\frac{\xi}{8})}{m}}$  为假设函数空间复杂性项, 可近似为  $\sqrt{d/m}$ .

令目标域已标识样例个数为  $m_T = \beta m$ , 源域已标识样例个数为  $m_S = (1-\beta)m$ ,  $D = \sqrt{d}/A$ . 可知函数  $f(\alpha)$  为  $m_T$ 、 $m_S$  和  $D$  的函数,  $\alpha$  的最优值为

$$\alpha^*(m_T, m_S; D) = \begin{cases} 1, & m_T \geq D^2 \\ \min\{1, v\}, & m_T \leq D^2 \end{cases} \quad (32)$$

其中,  $v = \frac{m_T}{m_T + m_S} \left(1 + \frac{m_S}{\sqrt{D^2(m_S + m_T) - m_S m_T}}\right)$ . 如果  $m_T = 0$ , 则  $\alpha^* = 0$ ; 如果  $m_S = 0$ , 则  $\alpha^* = 1$ . 这说明当只有源域已标识样本或只有目标域已标识样本时, 最优的组合是只使用当前可获得的已标识样例来进行学习. 如果源域和目标域概率分布相同, 即  $A = 0$  (或  $D \rightarrow \infty$ ), 那么  $\alpha^* = \beta$  是最优选择, 此时使用均匀加权 (Uniform weighting) 后的样本作为训练样本得到最优组合误差界.

### 6.3.4 多源域多目标域的组合学习误差界

假设有  $N$  个不同的源域, 这些源域总共有  $m$  个已标识样例. 第  $j$  个源域  $S_j$  记为  $(D_j, f_j)$ , 源域  $S_j$  有  $m_j = \beta_j m$  个已标识样例. 域的加权向量为  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N)$ ,  $\sum_{j=1}^N \alpha_j = 1$ . 根据这些不同源域的样本与目标域分布的相似度, 来衡量组合分布的误差估计上界, 多源域假设下目标域的组合经验误差可定义为  $\hat{\epsilon}_\alpha(h) = \sum_{j=1}^N \alpha_j \hat{\epsilon}_j(h) = \sum_{j=1}^N \frac{\alpha_j}{m_j} \times \sum_{x \in S_j} |h(x) - f_j(x)|$ . 以下分别对成对散度 (Pairwise divergence) 的误差界和组合散度的误差界两种情况进行讨论.

#### 1) 成对散度下的误差界

成对散度的目标域误差界着重考虑目标域和每个源域之间的  $H\Delta H$  距离, 对最小化目标域的平均散度和收敛速度进行权衡. 成对散度的目标域误差界没有充分利用源域之间的关联关系.

记  $\hat{h} \in H$  为加权向量  $\alpha$  下  $\hat{\epsilon}_\alpha(h)$  的经验最小假设函数,  $h_T^* = \min_{h \in H} \epsilon_T(h)$ , 下面给出假设函数  $\hat{h}$  关于目标域的期望误差  $\epsilon_T(\hat{h})$  与  $h_T^*$  关于目标域的期望误差  $\epsilon_T(h_T^*)$  的关系. 对于  $\forall \epsilon_T(\hat{h})$ ,  $\exists \delta$ , 以至少  $1 - \delta$  的概率对所有  $h \in H$  以下不等式成立:

$$\begin{aligned} \epsilon_T(\hat{h}) &\leq \epsilon_T(h_T^*) + \\ &2\sqrt{\left(\sum_{j=1}^N \frac{\alpha_j^2}{\beta_j}\right) \left(\frac{d\log(2m) - \log(\delta)}{2m}\right)} + \\ &\sum_{j=1}^N \alpha_j (2\lambda_j + \mathbf{d}_{H\Delta H}(D_j, D_T)) \end{aligned} \quad (33)$$

其中,  $\lambda_j = \min_{h \in H} \{\epsilon_T(h) + \epsilon_j(h)\}$ ,  $\delta \in (0, 1)$ .

#### 2) 组合散度下的误差界

组合散度的误差界讨论着重考虑目标域  $D_T$  和源域组合分布  $D_\alpha$  之间的  $H\Delta H$  距离, 充分考虑了不同源域之间的关联性. 下面给出组合散度下假设函数  $\hat{h}$  关于目标域的期望误差  $\epsilon_T(\hat{h})$  与  $h_T^*$  关于目标域的期望误差  $\epsilon_T(h_T^*)$  的关系. 对于  $\forall \epsilon_T(\hat{h})$ ,  $\exists \delta$ , 以至少  $1 - \delta$  的概率对所有  $h \in H$  以下不等式成立:

$$\begin{aligned} \epsilon_T(\hat{h}) &\leq \epsilon_T(h_T^*) + \\ &4\sqrt{\left(\frac{d\log(2m) - \log(\delta)}{2m}\right)} \times \\ &\sqrt{\left(\sum_{j=1}^N \frac{\alpha_j^2}{\beta_j}\right) + 2\gamma_\alpha + \mathbf{d}_{H\Delta H}(D_\alpha, D_T)} \end{aligned} \quad (34)$$

其中,  $\gamma_\alpha = \min_h \{\epsilon_T(h) + \epsilon_\alpha(h)\} = \min\{\epsilon_T(h) + \sum_{j=1}^N \alpha_j \epsilon_j(h)\}$ .

### 6.3.5 重要性加权域自适应学习误差界

对于  $\alpha \geq 0$ , 目标域  $D_T$  和源域  $D_S$  的 Renyi 散度  $D_\alpha(D_T \| D_S)$  定义为

$$D_\alpha(D_T \| D_S) = \frac{1}{\alpha - 1} \log_2 \sum_{\mathbf{x}} P(\mathbf{x}) \left(\frac{P_T(\mathbf{x})}{P_S(\mathbf{x})}\right)^{\alpha-1} \quad (35)$$

Renyi 散度为非负数, 且对于任意  $\alpha \geq 0$ , 当分布  $D_T$  和  $D_S$  相同时,  $D_\alpha(D_T \| D_S) = 0$ . 记:

$$d_\alpha(D_T \| D_S) = 2^{D_\alpha(D_T \| D_S)} = \left[\sum_{\mathbf{x}} \frac{P_T^\alpha(\mathbf{x})}{P_S^{\alpha-1}(\mathbf{x})}\right]^{\frac{1}{\alpha-1}} \quad (36)$$



重要性加权域自适应学习中, 取权值为  $\mathbf{w}(\mathbf{x}) = P_T(\mathbf{x})/P_S(\mathbf{x})$ , 该权的优点是其加权误差是期望生成误差的无偏估计. 记  $\sup_{\mathbf{x}} \mathbf{w}(\mathbf{x}) = \sup_{\mathbf{x}} P_T(\mathbf{x})/P_S(\mathbf{x}) = d_\infty(D_T \| D_S)$  为损失函数, 定义重要性加权域自适应学习在假设函数  $h \in H$  下的加权误差为

$$\hat{\varepsilon}_w(h) = \frac{1}{m} \sum_{i=1}^m \omega(\mathbf{x}_i) L(h(\mathbf{x}_i), f_T(\mathbf{x}_i)) \quad (37)$$

原目标域误差为  $\varepsilon_w(h) = E_{\mathbf{x} \sim D_T} [L[h, f_T]]$ .

根据权值  $\mathbf{w}(\mathbf{x})$  的特性, 可以分为权值有界和权值无界两种情况.

#### 1) 权值有界的情况

当权值有界时, 即  $d_\infty(D_T \| D_S) < +\infty$  时, 记  $M = d_\infty(D_T \| D_S)$ . 对于任意  $\alpha \geq 1, \delta > 0$ , 参考假设函数  $h \in H$  下的重要性加权误差为  $\hat{\varepsilon}_w(h)$ , 则对于  $\forall \varepsilon_T(\hat{h}), \exists \delta$ , 以至少  $1 - \delta$  的概率对所有  $h \in H$  以下不等式成立:

$$\varepsilon_T(h) \leq \hat{\varepsilon}_w(h) + \frac{2M \log \frac{1}{\delta}}{3m} + \sqrt{\frac{2 \left[ d_{\alpha+1}(D_T \| D_S) \varepsilon_T(h)^{1-\frac{1}{\alpha}} - \varepsilon_T(h)^2 \right] \log \frac{1}{\delta}}{m}} \quad (38)$$

当  $\alpha = 1$ , 上式简化为

$$\varepsilon_T(h) \leq \hat{\varepsilon}_w(h) + \frac{2M \log \frac{1}{\delta}}{3m} + \sqrt{\frac{2d_{\alpha+1}(D_T \| D_S) \log \frac{1}{\delta}}{m}} \quad (39)$$

#### 2) 权值无界的情况

当权值无界时, 即条件  $d_\infty(D_T \| D_S) < +\infty$  不成立, 可以证明  $d_2(D_T \| D_S) < +\infty$  是有界的. 定义实值判别函数  $f$  的伪维数为  $Pdim(f)$ , 假设函数  $h \in H$  下目标域误差  $\varepsilon_T(h) = \varepsilon_T(h(x), f_T(x))$  的伪维数  $Pdim(\{\varepsilon_T(h), h \in H\}) = p < \infty$ . 那么重要性加权域自适应学习误差至少以概率  $1 - \delta$  满足:

$$\varepsilon_T(h) \leq \hat{\varepsilon}_w(h) + 2^{\frac{5}{4}} \sqrt{d_2(D_T \| D_S)} \times \sqrt{\frac{p \log \frac{2m\varepsilon}{p} + \log \frac{4}{\delta}}{m}} \quad (40)$$

## 7 域自适应学习最新研究进展

近年来许多研究人员对域自适应学习进行了深

入研究, 并提出了许多新的思路. 本节分别从理论与算法、结构与模型和实际应用三个方面展开综述.

### 7.1 域自适应学习算法

多源域自适应学习是当前较为重要的研究内容之一, Mansour 等在文献 [45] 中提出了多个源域的加权组合域自适应学习问题, 证明了对于每一个源域, 这些假设模型的凸组合的性能并不理想. 如果对每个源域进行适当的加权组合可能会有良好的学习效果. 假定由凸组合得到的目标域分布的近似模型的误差为  $\varepsilon$ , 那么通过加权组合规则得到的目标域分布的模型具有不超过  $3\varepsilon$  的损失上界.

Duan 等<sup>[46]</sup> 根据平滑性假设, 使用域依赖正则化技术提出了多源域的域自适应 SVM 学习机设计方法, 进一步设计了快速域自适应学习 (Fast DAM) 和 Universum 域自适应学习机 (Universum DAM), 并在域自适应学习场景下实现了 SVM 分类、SVM 回归和 LS-SVM 的应用. 在文献 [47] 中, Bellet 等根据 Balcan 提出的相似性学习框架, 提出了一种稀疏化方法来处理图像伸缩和旋转情况下的图像分类问题, 这种自适应方法具有较高的借鉴价值. Morvant 等<sup>[48]</sup> 对 Balcan 等提出的域自适应学习框架进行了拓展, 基于相似函数得到投影空间下的线性分类函数, 且该分类函数并不一定要满足对称或半正定条件. 并针对无监督情况和半监督情况, 详细讨论了基于相似函数的无监督和半监督域自适应学习问题.

Chapelle 等<sup>[49]</sup> 提出的一种全新多任务学习算法对当前域自适应学习具有重要的借鉴作用. 对于给定的  $m$  个任务的样本集, 分别建立  $m$  个对应的学习模型以及一个全局模型, 通过梯度提升回归算法 (Gradient boosted regression) 学习  $m$  个任务的共性和特性. 使用到多任务网页搜索过程得到了较为准确的预测结果. 同时这种算法并不局限于只建立一个全局模型, 还可适用于建立多个全局模型的多任务学习情形.

在命名实体识别中, Sun 和 Grishman<sup>[50]</sup> 首次提出跨域自举算法 (Cross-domain bootstrapping algorithm). 该算法基于源域词团的特征选择和多判据样例选择, 并没有使用任何注释数据, 同时在算法中也并没有添加某些已知的目标域独有特性. 因此算法具有较好的普适性. 但是关于跨域自举程序的自动终止判据仍需要进一步的研究探讨.

Novotney 等<sup>[51]</sup> 探讨了在没有任何域内人工翻译 (In-domain manual transcription) 情况下如何使用自适应方法, 将标准阿拉伯语模型 (Modern standard Arabic (MSA) models) 转换为 Levantine

阿拉伯语模型, 指出 Self-training 方法并不适合将 MSA 模型转化为 Levantine 模型的学习, 并对将来如何减少 MSA 模型和 Levantine 模型偏差的方法进行了探讨.

Zhuang 等<sup>[52]</sup> 针对跨域文本分类问题, 提出了相应的域自适应解决方案, 通过使用矩阵三因子分解 (Matrix tri-factorization) 来建立从源域到目标域的变换, 弥补了传统机器学习在跨域文本分类中的不足.

结构对应学习 (SCL) 是当下域自适应的热门课题之一. Prettenhofer 和 Stein<sup>[53]</sup> 拓展了结构对应学习, 将跨语言域自适应学习和结构学习相结合, 较为系统地阐述了跨语言结构学习的原理及方法, 相对于传统的主题分类和情感分类方法, 这种自适应学习方法的相对误差有非常明显的改善.

域自适应学习中常出现源域和目标域的初始特征空间不同的情况. 通常可以采用核映射函数的观点对该学习场景下的域自适应问题进行学习, 针对这一问题, Blitzer 等<sup>[54]</sup> 创新性地提出了一种基于低维耦合子空间的域自适应学习算法, 该学习算法通过将目标域特有特征和源域特征进行相互关联, 可以根据源域已标记样例来预测目标域样例的类标签. 算法高效简洁且可以很容易地与当前的监督和半监督域自适应学习方法相结合.

Chen 等<sup>[55]</sup> 提出了一种协同训练域自适应 (Co-training for domain adaptation, CODA) 学习新方法. 该方法将获得的目标域特征和样例逐渐加入到源域训练集中来逐渐减小源域和目标域概率分布的差异. 并验证了 CODA 在 Blitzer 给出的 12 种域自适应学习基准数据集上具有优良的性能. 在当前大部分的目标域监督学习中, CODA 几乎都能达到最优的性能.

## 7.2 框架和模型

Tao 等<sup>[56]</sup> 提出一种统一的域自适应核学习 (Kernel learning for domain adaptation learning, KLDAL) 框架, 并扩展到多核学习领域. 同时提出了域自适应核学习支持向量机 (KLDAL-SVM)、 $\mu$ -KLDALSSVM、最小二乘域自适应核学习算法 KLDAL-LSSVM. 与许多基于核映射的支持向量机 (如  $\nu$ -SVM<sup>[57]</sup>、TSVM<sup>[58]</sup>、LS-SVM<sup>[59]</sup>、DA-SVM<sup>[60]</sup>) 和基于最小化均值偏差的学习算法 (如 DTSVM<sup>[61]</sup>、FastDAM<sup>[62]</sup>) 相比, 该学习框架具有更精确的学习性和更多样的扩展性.

Yang 等<sup>[63]</sup> 针对域自适应学习提出了一种全新的域自适应支持向量机 (Adaptive SVM, A-SVM), 该学习结构首先通过对辅助域进行学习得到相应的

分类判别函数, 然后由这些辅助分类判别函数得到需要学习的目标域的判别函数. Jiang 等<sup>[64]</sup> 根据目标域的  $k$  临近 ( $k$ -nearest neighbors) 来对每个辅助模型参数定义权重, 然后根据重新加权的模型参数训练相应的域自适应支持向量机, 相对于上述几种学习结构<sup>[57-59]</sup>, 该学习结构的优势在于广泛的适用性和高效的学习性.

Quanz 和 Huan<sup>[65]</sup> 提出了一种大间隔直推式学习结构, 该结构通过引入正则化项使分类间隔最大化, 同时使源域和目标分布距离最小化. 通过正则化项的引入来确定判别函数能够充分有效地利用源域已标识样本和目标域未标识样例的信息进行学习, 该结构的性能比当前大部分的域自适应分类模型都要精确.

在跨域学习中常常要引入大量的辅助域, 常常会遇到不同域的特征分布存在的显著变化的情况, 针对这一问题, Duan 等<sup>[66]</sup> 提出了域迁移多核学习框架 (Domain transfer multiple kernel learning, DTMKL), 且该 DTMKL 下, 特征选择和 DKMKL 最优参数的确定这两个问题具有较大的研究和借鉴价值.

在网络调查报告中通常包含用户对产品和服务的评价信息, 为了解决获得评估报告的主题以及评价情感问题, Jo 和 Oh<sup>[67]</sup> 提出了一种统一的在线报告分析模型: 主题情感联合模型 (Aspect and sentiment unification model, ASUM). 该模型并不需要报告中的任何情感标签, 有效地解决了在获得情感标签成本很大的情况下如何自动进行报告分析.

Malandrakis 等<sup>[68]</sup> 基于情感词汇构造核模型, 建立了域独立的情感文本模型 (Affective text modeling), 相对于当前自然语言处理方面的域自适应模型, 该模型在新闻标题的语义评估、文明礼貌评估 (Politeness detection) 和沮丧度检测 (Frustration detection) 三大情感分析任务中实现了最优的预测结果.

## 7.3 实际应用

本节介绍当前域自适应学习在各个领域的最新应用研究进展.

1) 自然语言处理: 自然语言处理是当前域自适应学习理论应用的最前沿领域之一. Li 等<sup>[69]</sup> 对当前域自适应学习在自然语言处理中的理论、技术和应用进行了详细的论述.

a) 不同语料库的域自适应学习方面: 针对不同语料库的域自适应学习问题, Finkel 和 Manning<sup>[70]</sup> 基于分层贝叶斯先验, 提出了分层贝叶斯结构的域自适应学习方法, 解决了 CoNLL 数据和 MUC 数据

的匹配学习问题. Chelba 和 Acero<sup>[71]</sup> 基于高斯先验分布, 利用最大熵马尔科夫模型, 实现了从华尔街日报到广播新闻的单词大写 (Capitalization) 自动识别这一学习任务的自适应学习. 相对于 CoNLL 和 MUC 数据匹配, CSPACE 个人邮件数据和 MUC 数据的差异性更大, 为了实现这一复杂的域自适应学习, Arnold 等<sup>[72]</sup> 提出了一种全新的无监督分层域自适应学习方法, 该方法具有良好的学习性能及鲁棒性.

b) 同一语料库的主题划分问题: 针对语料库主题划分问题, Jiang 和 Zhai<sup>[73]</sup> 根据域内共有特征, 提出了一种二步加权域自适应学习方法, 并通过生物属性成功地将 BioCreAtIvE 数据集划分为 3 个主题. Daumé 等<sup>[74]</sup> 基于核空间下的特征复制域自适应学习方法, 将 ACE2005 语料库划分为 6 类主题. 在用户产品调查报告的情感分析问题上, Glorot 等<sup>[75]</sup> 进行了大量的实验和研究, 并提出了一种深度学习方法来提取调查报告中的有效信息, 该特征表示下的域自适应学习在多文本学习场景下具有非常好的学习性能.

c) 跨语言自适应学习问题: 在跨语言自适应学习方面, Zeman 和 Resnik<sup>[76]</sup> 基于特征表示域自适应学习框架, 提出了一种较为简单的跨语言深度学习算法, 将源域和目标域的资料库的词性标注映射为一个普通的标签集, 实现了丹麦语和瑞典语的语法解析匹配. Hara 等<sup>[77]</sup> 利用树图资料库来进行训练任务, 实现了 HPSG 句法分析到生物医学领域的自适应学习. Xu 等<sup>[78]</sup> 通过情感分析技术, 根据源域语言的情感模型, 成功地实现了目标域语言的域自适应学习, 并以汉语和英语两种语言为基础进行了相关的实验探讨和验证.

d) 情感分析: 情感分类是当下的热点课题之一, 也是域自适应的重要应用领域之一. Glorot 等<sup>[79]</sup> 针对大规模情感分类问题, 提出了一种无监督的深度学习算法, 通过学习多视图的特征表示, 该域自适应学习方法能够对亚马逊数据中出现的 22 种数据集进行良好的学习. 针对语义结构的句法分析问题, Das 和 Smith<sup>[80]</sup> 提出了一种基于图的半监督自适应学习策略, 在对语义知识库数据的实验中, 其准确度比传统的句法结构分析方法高 10%~17%. Blitzer 等<sup>[81]</sup> 把结构对应学习应用于情感分析问题, 该方法与传统的 SCL 方法比较, 相对预测误差减少了 30%. 在跨域语言情感分类中, Bollegala 等<sup>[82]</sup> 通过建立情感词库对特征向量进行扩充, 高效地实现了亚马逊产品报告中的多种产品报告的情感分类问题.

e) 机器翻译问题: 在机器翻译领域, Pecina

等<sup>[83]</sup> 对统计机器翻译的当前进展做了较为详细的综述, 提出了一种从网络中获取 (Crawl) 域特殊信息的策略, 并探讨了基于短语的统计机器翻译的域自适应学习方法, 同时指出了未来研究的方向和问题. Axelrod 等<sup>[84]</sup> 提出了一种应用于自然语言处理的基于样例加权的域自适应算法. 从和目标语料库相关的大量平行语料库中提取相应的句子, 通过多模型组合和语料库样本选择, 设计了相应的机器翻译学习机, 实现了高效的汉译英自适应学习.

2) 字符识别: 字符识别需要对不同书写风格的书写体进行识别, 由于不同书写体的字符具有差异性, 因此使用各种书写体数据集作为训练样本训练出的字符分类器 (Writer independent recognizer) 对特定书写人特定风格的书写体样本进行识别分类时, 分类精度通常较低. 针对字符识别中存在的域自适应问题, Ding 和 Jin<sup>[85]</sup> 对汉字识别中常用分类器类型的修正二次判别函数 (Modified quadratic discriminant function, MQDF) 提出了增量化 MQDF 和判别增量 MQDF. Zhang 和 Liu<sup>[86]</sup> 提出了风格迁移映射 (Style transfer mapping, STM), STM 为二次规划问题, 将特定书写人特定风格的样本定义为“源点集”, 而分类器中样本真实类别所对应的部分参数定义为“目标点集”, STM 通过将源点集映射到目标点集, 从而去除测试样本风格的影响, 实现精确的分类. 此外 STM 可以与多种分类器相结合, 应用于有监督、无监督和半监督的域自适应学习问题. STM 只需要少量的自适应样本即可实现有效的自适应学习, 且 STM 只对测试样本进行仿射变换, 因而能够保持住原始分类器的结构和参数. 在对手写体汉字库 CASIA-OLHWDB 识别实验中, STM 能够对汉字库 CASIA-OLHWDB 实现高效准确的分类.

3) 视频分析: 视频分析中常常涉及到跨模型学习问题, 因此是域自适应理论的应用领域之一. Faraji-Davar 等<sup>[87]</sup> 提出了一种面向体育视频行为识别的域自适应学习方法, 以羽毛球视频为源域, 网球视频作为目标域, 通过对源域特征进行线性转化并最大化域特征相似性, 极大地改善了体育视频行为的识别分类精度. 针对视频概念检测中的域自适应学习问题, Duan 等<sup>[88]</sup> 通过最小化结构风险函数及域分布偏差, 提出了一种新的跨域核学习方法, 并对 TRECVID 视频数据集实现了良好的跨域学习.

4) 文本搜索排序问题: 排序函数是文本检索系统的重要组成部分, 如今机器学习理论在排序问题中具有广泛的应用, 但是大部分的算法都是基于有监督的学习, Duh 和 Kirchhoff<sup>[89]</sup> 从域自适应学

习的角度, 采用特征变换和样例加权两种方法, 提出了文本检索的半监督排序方法 (Semi-supervised ranking for document retrieval), 这对域自适应理论的应用有着重要的借鉴作用; 在域垂直搜索引擎应用领域, 不同域需要不同的排序模型以保证具有良好的排序性能, 如果对每个域都建立一个独特的排序模型是非常费时费力的. 针对特定域搜索的排序模型自适应学习问题, Geng 等<sup>[90]</sup> 提出了排序自适应支持向量机 (RA-SVM), 并以数据集 TD2003 到 TD2004 为实验样本进行了算法高效性的实验验证.

## 8 当前域自适应存在的问题及未来的发展趋势

最近几年, 国内的域自适应学习也取得了一定的研究进展. 例如基于局部加权的域自适应学习框架<sup>[91-92]</sup> 及最新的域自适应核 SVM 学习<sup>[93]</sup> 等, 但是还存在较多尚未尝试的研究领域, 如域自适应学习的费用敏感度分析、域自适应聚类学习等. 当前的域自适应学习虽然取得了一定的研究进展, 但是仍然存在很多亟待解决的问题. 下面将结合国内外域自适应学习的研究进展, 从特征变换和假设、训练优化算法、数据表示与模型构造、NLP 研究中存在的问题这四大方面, 对当前域自适应学习中所存在的问题进行了总结和讨论, 并结合域自适应学习的发展趋势, 对未来需要进一步研究探讨的内容给出了合理化的预测和建议.

### 8.1 特征变换和假设问题

在域自适应方法中常常需要将样例映射到特征空间, 但是在做特征变换时我们往往会做出某些假设. 当前绝大部分的域自适应学习算法都要满足一定的假设条件才能保证最后的学习性能, 但这些假设通常不能满足. 例如经常使用的协变量迁移假设  $P_S(Y|X) = P_T(Y|X)$ , 实际上由于学习之前并不知道概率分布  $P_S(Y|X)$ 、 $P_S(Y|X)$ 、 $P_T(Y|X)$  和  $P_T(X|Y)$ , 因此这种假设并不一定满足. 即使满足这些假设, 自适应学习问题仍可能是一病态问题. 因此在进行域自适应学习的相关研究中往往带来较大的局限性. 如何判断域自适应学习的假设是否成立, 如何直接利用源域和目标域样本给出有概率担保的和实际可操作的判别准则, 提出更加切实可行的新的域自适应学习假设条件, 如何给出克服对现有假设条件依赖和给出新的特征表示技术, 如反向学习问题等, 是特别值得研究的前沿课题. 直推和转导学习思想似乎值得借鉴, 转导学习并不需要学习整个

模型的概率分布, 直接学习估计点的值.

另外, 克服学习的病态问题依赖于正则化方法, 而正则化项对应于贝叶斯理论的先验信息, 因而基于全贝叶斯学习的理论的引入必将给域自适应学习带来新的学习范式.

### 8.2 训练优化方面

核方法在域自适应学习中具有重要的意义, 当前最新的域自适应核学习算法大都是基于最大均值偏差 (Maximum mean deviation, MMD) 度量不同分布的差异性. 但是在某些特定的领域, 仅仅使用 MMD 并不能有效地度量分布的差异性, 除均值外往往还需要考虑方差等多种分布特性, 最近提出的大量新模型通常只采用方差作为度量而忽略了域分布的其他特性, 这严重地限制了域自适应学习的适用性和预测的准确性. 虽然近来有人提出了这一问题并进行了相应的研究, 例如在第 3.1 节中 Tao 等提出使用均值和方差两种特性来度量域分布的差异, 但是在域自适应学习中仍然缺少类似的多特性度量判据. 如何更为精确高效地度量域分布偏差这一研究问题仍需要引起足够的重视, 这也是未来域自适应学习的研究热点之一.

当前最新的域自适应学习方法都是基于支持向量机或其他核方法, 这些核方法的性能严重依赖于核函数的选择, 而我们事先并不知道对所学习的问题最合适的核是什么. 在样本很大的情况下, 使用常用的核函数来进行域自适应学习, 往往需要大量的运算时间, 实际不可行. 虽然最近又提出了多核学习 (Multiple kernel learning, MKL) 域自适应学习方法, 但是这些 MKL 算法也不能找到最优的核函数, 只有在特定的假设条件和约束下才能实现高效的基于核的域自适应学习. 这也是域自适应学习领域急需解决的问题.

### 8.3 数据表示与模型方面

域自适应学习的过程往往伴随着对源域的样本的表示进行变换, 例如对源域样本的复制、降维、投影到子空间, 计算邻接矩阵构造流形表示等, 而这些变换和原来常用的无监督的特征变换手段不同点在于, 经过特征变换后表示的联合概率分布要尽可能地与目标域概率分布一致.

当前构造源域的特征表示过程还没有充分利用问题背后的先验信息, 如平滑性假设、模型稀疏性假设、空间和时间上的一致性关系和具有多个解释因素的层次结构先验等先验信息. 构造特征表示过程, 应该引入结构化稀疏先验信息, 实现结构层次化模型, 并利用稀疏化过程选择源域的特征表示, 使得选

择的特征表示与目标域的概率分布一致<sup>[94]</sup>.

域自适应学习假定源域与源域之间、源域与目标域之间、目标域与目标域之间的概率分布不一样,而这些概率分布是未知的和定义在源域和目标域样本上的,域自适应学习的过程往往又依赖于对源域与目标域样本的表示的变换,所以,能够对概率分布直接表示并且适用于表示学习的基于能量函数的模型特别适合域自适应学习场景<sup>[95]</sup>,能量函数来自于统计物理势能和动能组成的拉格朗日函数,对应于一般机器学习目标函数中的损失函数 + 正则化项的学习框架,因而基于能量函数的模型包括了先验信息,可以克服如何选择先验信息也就是正则化项的问题,学习的目标是使得能量函数最小.

#### 8.4 NLP 中存在的问题

在当前自然语言处理领域,虽然提出了大量的算法,也在特定的实验中取得了较好的效果,但是仍存在许多要改进的方面:

1) 在命名实体识别 (Part-of-speech (POS) tagging) 中一些命名实体可能有多种单词表示,且上下文的语境和词汇可能不同,这对自适应学习的效果有很重要的影响,如何利用命名实体多个表示和上下文语境,实现域自适应学习是值得研究的问题,这方面多视图的学习技术<sup>[96-97]</sup>可能会有帮助.

2) 词义消歧中单词的先验分布不同问题:自然语言处理中,不同域中的单词样本的先验概率分布可能是不同的(即同一个单词的不同词义所占比例不同),训练样本和目标域样本之间单词的词义先验分布不同会降低词义消歧的精确度.实现域自适应词义消歧算法值得研究.

3) 情感分类中不同主题的学习问题:自动情感分类是自然语言处理中重要的研究方向之一,不同域的情感表达是不同的(例如不同类型产品的在线评论是不同的),但对所有感兴趣的域的语料库进行一一人工注释是不可行的,因此有必要对情感分类的域自适应学习算法进行研究.

4) 机器翻译中如何提高算法预测精度,减少模型的复杂度的问题:大量文献<sup>[98-102]</sup>对机器翻译中的语料库句子排序、符号(Token)等的学习问题进行了研究.但是这些算法通常存在计算复杂度较高和算法预测准确度不高的问题,如何利用域自适应学习技术改进排序学习和符号学习问题是当前 NLP 需要解决的问题之一.

5) 在主题建模中, LDA 是一种常用的模型<sup>[103]</sup>,虽然 LDA 模型可以用于域自适应学习问题<sup>[104]</sup>,但是仍然存在许多不足,例如在不同的文本中,一个单词可能有不同的主题,包括后续的改进 TPLSA 模

型存在参数设定和分布假设太苛刻等问题,关于主题建模的域自适应学习的研究仍需要继续拓展和补充.

这些都是当前自然语言处理存在的问题,需要继续展开相应的研究,也是未来域自适应学习在 NLP 应用中的重要研究方向.

#### 8.5 拓展应用范围和应用领域

在回归和聚类等学习问题上的域自适应学习理论和算法的研究还很不充分,域自适应学习目前主要集中于语言、文本和机器翻译领域的研究,在生物信息学、计算机视觉、机器人学、语音识别与合成等应用领域,有许多研究工作要做,如何把当前热点研究问题如主动学习、弱监督学习、在线学习、大规模并行学习、表示学习和深度学习与域自适应学习相结合也是大有可为的研究方向.

### 9 结论

域自适应学习的研究在近年来受到了越来越多的关注,许多学者对域自适应学习理论和应用进行了广泛的研究.本文对域自适应学习方法和理论进行了综述,对样例加权域自适应学习、特征表示域自适应学习、参数和特征分解域自适应学习和多源域自适应学习进行了详细的阐述,同时结合域自适应学习的误差界、高效的度量判据等问题对域自适应学习的可能性、鲁棒性、统计估计一致性进行了理论分析.接着对当前域自适应学习在学习算法、模型结构和实际应用这三个方面的研究进展进行了综述.最后分别探讨了域自适应在特征变换和假设、训练优化、模型与数据表示、NLP、拓展应用范围和领域这五个方面中有待进一步研究解决的问题.随着域自适应理论和方法研究的深入,域自适应学习必将更加广泛地应用于传统机器学习的各个领域.

### References

- 1 Jiang J. A literature survey on domain adaptation of statistical classifiers [Online], available: [http://sifaka.cs.uiuc.edu/jiang4/domain\\_adaptation/survey](http://sifaka.cs.uiuc.edu/jiang4/domain_adaptation/survey), March 6, 2008
- 2 Mansour Y, Mohri M, Rostamizadeh A. Domain adaptation: learning bounds and algorithms. In: Proceedings of the 22nd Annual Conference on Learning Theory. Montréal, Canada: Omnipress, 2009. 34-47
- 3 Blitzer J, Crammer K, Kulesza A. Learning bounds for domain adaptation. In: Proceedings of the 21st Annual Conference on Neural Information Processing Systems. Vancouver, British Columbia, Canada: Curran Associates, 2007. 129-136
- 4 Cortes C, Mansour Y, Mohri M. Learning bounds for importance weighting. In: Proceedings of the 24th Annual Confer-

- ence on Neural Information Processing Systems. Vancouver, Canada: Curran Associates, 2010. 442–450
- 5 Zhong E, Fan W, Peng J, Zhang K, Ren J T, Turaga D S, Verscheure O. Cross domain distribution adaptation via kernel mapping. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Paris, France: ACM, 2009. 1027–1036
  - 6 Kulis B, Saenko K, Darrell T. What you saw is not what you get: domain adaptation using asymmetric kernel transforms. In: Proceedings of the 24th IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, CO, USA: IEEE, 2011. 1785–1792
  - 7 Ben-David S, Blitzer J, Crammer K, Kulesza A, Pereira F, Vauqhan J W. A theory of learning from different domains. *Machine Learning*, 2010, **79**(1–2): 151–175
  - 8 Joshi M, Dredze M, Cohen W W. What’s in a domain? multi-domain learning for multi-attribute data. In: Proceedings of the 2013 NAACL-HLT. Westin Peachtree Plaza Hotel, Atlanta, Georgia, USA: The Association for Computational Linguistics, 2013. 685–690
  - 9 Wan X J. Co-training for cross-lingual sentiment classification. In: Proceedings of the 2009 Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP. Singapore: ACL, 2009. 235–243
  - 10 Gabrilovich E, Markovitch S. Feature generation for text categorization using world knowledge. In: Proceedings of the 19th International Joint Conference on Artificial Intelligence. Edinburgh, Scotland, UK: Professional Book Center, 2005. 1048–1053
  - 11 Cucerzan S. Large-scale named entity disambiguation based on Wikipedia data. In: Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. Prague, Czech Republic: ACL, 2007. 708–716
  - 12 Sarinnapakorn K, Kubat M. Combining subclassifiers in text categorization: a DST-based solution and a case study. *IEEE Transactions on Knowledge and Data Engineering*, 2007, **19**(12): 1638–1651
  - 13 Gabriel P C F, Jeffrey X Y, Lu H J, Philip S Y. Text classification without negative examples revisit. *IEEE Transactions on Knowledge and Data Engineering*, 2006, **18**(1): 6–20
  - 14 Al-Mubaid H, Umair S A. A new text categorization technique using distributional clustering and learning logic. *IEEE Transactions on Knowledge and Data Engineering*, 2006, **18**(9): 1156–1165
  - 15 Gabrilovich E, Markovitch S. Enhancing text categorization with encyclopedic knowledge. In: Proceedings of the 21st National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference. 2006, Boston, Massachusetts, USA: AAAI Press, 2006. 1301–1306
  - 16 Joshi M, Cohen W W, Dredze M, Rosé C P. Multi-domain learning: when do domains matter? In: Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. Jeju, Island, Korea: Association for Computational Linguistics, 2012. 1302–1312
  - 17 Pan S J, Kwok J T, Yang Q, Pan J J. Adaptive localization in a dynamic WiFi environment through multi-view learning. In: Proceedings of the 22nd AAAI Conference on Artificial Intelligence. Vancouver, British Columbia, Canada: AAAI Press, 2007. 1023–1027
  - 18 Chen X Y, Yuan X T, Chen Q, Yan S C, Chua T S. Multi-label visual classification with label exclusive context. *IEEE Transactions on Computer Vision*, 2011, **1**(1): 834–841
  - 19 Tu W T, Sun S L. Transferable discriminative dimensionality reduction. In: Proceedings of the 23rd IEEE International Conference on Tools with Artificial Intelligence. Boca Raton, FL, USA: IEEE, 2011. 865–868
  - 20 Cortes C, Mohri M. Domain adaptation in regression. In: Proceedings of the 22nd Algorithmic Learning Theory International Conference. Espoo, Finland: Springer, 2011. 308–323
  - 21 Zadrozny B. Learning and evaluating classifiers under sample selection bias. In: Proceedings of the 21st International Conference on Machine Learning. New York, NY, USA: ACM, 2004. 114–115
  - 22 Sugiyama M, Nakajima S, Kashima H. Direct importance estimation with model selection and its application to covariate shift adaptation. In: Proceedings of the 21st Annual Conference on Neural Information Processing Systems. Vancouver, British Columbia, Canada: Curran Associates, 2008. 1433–1440
  - 23 Fan W, Davidson I, Zadrozny B, Yu P S. An improved categorization of classifier’s sensitivity on sample selection bias data mining. In: Proceedings of the 5th IEEE International Conference on Data Mining. Houston, Texas, USA: IEEE, 2005. 27–30
  - 24 Quionero-Candela J, Sugiyama M, Schwaighofer A. *Dataset Shift in Machine Learning*. Boston: The MIT Press, 2009
  - 25 Dai W Y, Yang Q, Xue G R, Yu Y. Boosting for transfer learning. In: Proceedings of the 24th International Conference on Machine Learning. Beijing, China: ACM, 2007. 193–200
  - 26 Xu Z J, Sun S L. Multi-view transfer learning with Adaboost. In: Proceedings of the 23rd IEEE International Conference on Tools with Artificial Intelligence. Boca Raton, FL, USA: IEEE, 2011. 399–402
  - 27 Xu Z J, Sun S L. Multi-source transfer learning with multi-view Adaboost. In: Proceedings of the 19th International Conference on Neural Information Processing. Doha, Qatar: Springer, 2012. 332–339
  - 28 Tao J W, Chung K F L, Wang S. On minimum distribution discrepancy support vector machine for domain adaptation. *Pattern Recognition*, 2012, **45**(11): 3962–3984
  - 29 Sun S L, Xu Z J, Yang M. Transfer learning with part-based ensembles. *Multiple Classifier Systems*, 2013, **7872**: 271–282
  - 30 Tu W T, Sun S L. Cross-domain representation-learning framework with combination of class-separate and domain-merge objectives. In: Proceedings of the 1st International Workshop on Cross Domain Knowledge Discovery in Web and Social Network Mining. Beijing, China: ACM, 2012. 18–25

- 31 Blitzer J, McDonald R, Pereira F. Domain adaptation with structural correspondence learning. In: Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing. Sydney, Australia: ALC, 2006. 120–128
- 32 Dai W Y, Xue G R, Yang Q, Yu Y. Co-clustering based classification for out-of-domain documents. In: Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Jose, California, USA: ACM Press, 2007. 210–219
- 33 Pan S J, Kwok J T, Yang Q. Transfer learning via dimensionality reduction. In: Proceedings of the 23rd AAAI Conference on Artificial Intelligence. Chicago, Illinois, USA: AAAI Press, 2008. 677–682
- 34 Pan S J, Tsang I W, Kwok J T, Yang Q. Domain adaptation via transfer component analysis. *IEEE Transactions on Neural Networks*, 2011, **22**(2): 199–210
- 35 Daumé III H. Frustratingly easy domain adaptation. In: Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics. Prague, Czech Republic: ACL, 2007. 1785–1787
- 36 Bonilla E, Chai K M, Williams C. Multi-task Gaussian process prediction. In: Proceedings of the 21st Annual Conference on Neural Information Processing Systems. Vancouver, British Columbia, Canada: Curran Associates, 2008. 153–160
- 37 Schwaighofer A, Tresp V, Yu K. Learning Gaussian process kernels via hierarchical Bayes. In: Proceedings of the 17th Annual Conference on Neural Information Processing Systems. Vancouver, British Columbia, Canada: MIT Press, 2004. 1209–1216
- 38 Gao J, Fan W, Jiang J, Han J W. Knowledge transfer via multiple model local structure mapping. In: Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. Las Vegas, Nevada, USA: ACM Press, 2008. 283–291
- 39 Tu W T, Sun S L. Dynamical ensemble learning with model-friendly classifiers for domain adaptation. In: Proceedings of the 21st International Conference on Pattern Recognition. Tsukuba, Japan: IEEE, 2012. 1181–1184
- 40 Ben-David S, Lu T, Luu T. Impossibility theorems for domain adaptation. *Journal of Machine Learning Research*, 2010, **9**: 129–136
- 41 Ben-David S, Urner R. On the hardness of domain adaptation. In: Proceedings of the 23rd International Conference. Lyon, France: Springer, 2012. 139–153
- 42 Ben-David S, Shalev-Shwartz S, Urner R. Domain adaptation — can quantity compensate for quality? In: Proceedings of the 2012 International Symposium on Artificial Intelligence and Mathematics. Fort Lauderdale, Florida, USA: ISAIM, 2012. 641–648
- 43 Xu H, Mannor S. Robustness and generalization. *Machine Learning*, 2012, **86**(3): 391–423
- 44 Mansour Y, Schain M. Robust domain adaptation. In: Proceedings of the 12th International Symposium on Artificial Intelligence and Mathematics. Florida, USA: ISAIM, 2012. 27–36
- 45 Mansour Y, Mohri M, Rostamizadeh A. Domain adaptation with multiple sources. In Proceedings of the 22nd Annual Conference on Neural Information Processing Systems. Vancouver, British Columbia, Canada: Curran Associates, 2008. 1041–1048
- 46 Duan L X, Xu D, Tsang I W. Domain adaptation from multiple sources: a domain-dependent regularization approach. *IEEE Transactions on Neural Networks and Learning Systems*, 2012, **23**(3): 504–518
- 47 Bellet A, Habrard A, Sebban M. Similarity learning for provably accurate sparse linear classification. In: Proceedings of the 29th International Conference on Machine Learning. Edinburgh, Scotland, UK: Omnipress, 2012. 1871–1878
- 48 Morvant E, Habrard A, Ayache S. Parsimonious unsupervised and semi-supervised domain adaptation with good similarity functions. *Knowledge and Information Systems*, 2012, **33**(2): 309–349
- 49 Chapelle O, Shivaswamy P, Vadrevu S, Weinberger K, Zhang Y, Tseng B. Boosted multi-task learning. *Machine Learning*, 2011, **85**(1–2): 149–173
- 50 Sun A, Grishman R. Cross-domain bootstrapping for named entity recognition. In: Proceedings of the SIGIR 2011 Workshop on Entity-Oriented Search. Beijing, China: ACM Press, 2011
- 51 Novotney S, Schwartz R M, Khudanpur S. Unsupervised arabic dialect adaptation with self-training. *Interspeech*, 2011, 541–544
- 52 Zhuang F Z, Luo P, Xiong H, He Q, Xiong Y H, Shi Z Z. Exploiting associations between word clusters and document classes for cross-domain text categorization. *Statistical Analysis and Data Mining*, 2011, **4**(1): 100–114
- 53 Prettenhofer P, Stein B. Cross-language text classification using structural correspondence learning. In: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. Uppsala, Sweden: ALC, 2010. 1118–1127
- 54 Blitzer J, Kakade S, Foster D P. Domain adaptation with coupled subspaces. *Journal of Machine Learning Research*, 2011, **15**(11): 173–181
- 55 Chen M, Weinberger K Q, Blitzer J. Co-training for domain adaptation. In: Proceedings of the 25th Annual Conference on Neural Information Processing Systems. Granada, Spain: Springer, 2011. 2456–2464
- 56 Tao J W, Chung F L, Wang S T. A kernel learning framework for domain adaptation learning. *Science China Information Sciences*, 2012, **55**(9): 1983–2007
- 57 Scholkopf B, Smola A J, Williamson R C, Bartlett P L. New support vector algorithms. *Neural computation*, 2000, **12**(5): 1207–1245
- 58 Joachims T. Transductive inference for text classification using support vector machines. In: Proceedings of the 16th International Conference on Machine Learning. Bled, Slovenia: Morgan Kaufmann, 1999. 200–209
- 59 Suykens J A K, Vandewalle J. Least squares support vector machine classifiers. *Neural Processing Letters*, 1999, **9**(3): 293–300

- 60 Bruzzone L, Marconcini M. Domain adaptation problems: A DASVM classification technique and a circular validation strategy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(5): 770–787
- 61 Duan L X, Tsang I W, Xu D, Maybank S J. Domain transfer svm for video concept detection. In: Proceedings of the 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Miami, Florida, USA: IEEE, 2009. 1375–1381
- 62 Gretton A, Fukumizu K, Sriperumbudur B K. A fast consistent kernel two-sample test. In: Proceedings of the 23rd Annual Conference on Neural Information Processing Systems. Vancouver, British Columbia, Canada: Curran Associates, 2009. 673–681
- 63 Yang J, Yan R, Hauptmann A G. Cross-domain video concept detection using adaptive SVMs. In: Proceedings of the 15th international conference on Multimedia. Augsburg, Germany: ACM, 2007. 188–197
- 64 Jiang W, Zavesky E, Chang S F, Loui A. Cross-domain learning methods for high-level visual concept classification. In: Proceedings of the 15th IEEE International Conference. Adelaide, Australia: IEEE, 2008. 161–164
- 65 Quanz B, Huan J. Large margin transductive transfer learning. In: Proceedings of the 18th ACM Conference on Information and Knowledge Management. New York, USA: ACM Press, 2009. 1327–1336
- 66 Duan L X, Tsang I W, Xu D. Domain transfer multiple kernel learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(3): 465–479
- 67 Jo Y, Oh A H. Aspect and sentiment unification model for online review analysis. In: Proceedings of the 4th International Conference on Web Search and Web Data Mining. Hong Kong, China: ACM, 2011. 815–824
- 68 Malandrakis N, Potamianos A, Iosif E. Kernel models for affective lexicon creation. In: Proceedings of the 12th Annual Conference of the International Speech Communication Association. Florence, Italy: ISCA press, 2011. 2977–2980
- 69 Li Q, Li H B, Ji H, Wang W, Zheng J, Huang F. Joint bilingual name tagging for parallel corpora. In: Proceedings of the 21st ACM International Conference on Information and Knowledge Management. Maui, USA: ACM, 2012. 1727–1731
- 70 Finkel J R, Manning C D. Hierarchical Bayesian domain adaptation. In: Proceedings of the 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Singapore: ACL, 2009. 602–610
- 71 Chelba C, Acero A. Adaptation of maximum entropy capitalizer: little data can help a lot. *Computer Speech and Language*, 2006, **20**(4): 382–399
- 72 Arnold A, Nallapati R, Cohen W W. Exploiting feature hierarchy for transfer learning in named entity recognition. In: Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics. Ohio, USA: ACL, 2008. 245–253
- 73 Jiang J, Zhai C X. Instance weighting for domain adaptation in NLP. In: Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics. Prague, Czech Republic: ACL, 2007. 22–23
- 74 Daumé H, Marcu D. Domain adaptation for statistical classifiers. *Artificial Intelligence Research*, 2011, **26**: 101–126
- 75 Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks. In: Proceedings of the 14th International Conference on Artificial Intelligence and Statistics. Fort Lauderdale, USA: JMLR, 2011. 315–323
- 76 Zeman D, Resnik P. Cross-language parser adaptation between related languages. In: Proceedings of the 3rd International Joint Conference on Natural Language Processing. Hyderabad, India: ACL, 2008. 35–42
- 77 Hara T, Miyao Y, Tsujii J. Evaluating the impact of re-training a lexical disambiguation model on domain adaptation of an HPSG parser. *Trends in Parsing Technology*. Netherlands: Springer, 2010, **43**: 257–275
- 78 Xu R F, Xu J, Wang X L. Instance level transfer learning for cross lingual opinion analysis. In: Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis. Portland, Oregon, USA: ACL, 2011. 182–188
- 79 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, Washington, USA: Omnipress, 2011. 513–520
- 80 Das D, Smith N A. Semi-supervised frame-semantic parsing for unknown predicates. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. Portland, Oregon, USA: ACL, 2011. 1435–1444
- 81 Blitzer J, Dredze M, Pereira F. Biographies, bollywood, boom-boxes and blenders: domain adaptation for sentiment classification. In: Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics. Prague, Czech Republic: ACL, 2007. 440–447
- 82 Bollegala D, Weir D J, Carroll J. Using multiple sources to construct a sentiment sensitive thesaurus for cross-domain sentiment classification. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. Portland, Oregon, USA: ACL, 2011. 132–141
- 83 Pecina P, Toral A, Way A. Towards using web-crawled data for domain adaptation in statistical machine translation. In: Proceedings of the 15th Annual Conference of the European Association for Machine Translation. Leuven, Belgium: Springer, 2011. 63–72
- 84 Axelrod A, He X D, Gao J F. Domain adaptation via pseudo in-domain data selection. In: Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing. Portland, Oregon, USA: ACL, 2011. 355–362
- 85 Ding K, Jin L W. Incremental MQDF learning for writer adaptive handwriting recognition. In: Proceedings of the 2010 International Conference on Frontiers in Handwriting Recognition. Kolkata, India: IEEE, 2010. 559–564
- 86 Zhang X Y, Liu C L. Writer adaptation with style transfer mapping. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(7): 1773–1787



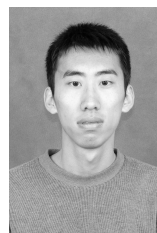
- 87 Faraji-Davar N, Decampos T, Windridge D, Kittler J, Christmas W. Domain adaptation in the context of sport video action recognition. In: Proceedings of the 25th Annual Conference on Neural Information Processing Systems. Sierra Nevada, Spain: Springer, 2011. 61–65
- 88 Duan L X, Tsang I W, Xu D, Chua T S. Domain adaptation from multiple sources via auxiliary classifiers. In: Proceedings of the 26th Annual International Conference on Machine Learning. Montreal, Canada: ACM, 2009. 37–51
- 89 Duh K, Kirchhoff K. Learning to rank with partially-labeled data. In: Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval. Singapore: ACM, 2008. 251–258
- 90 Geng B, Yang L J, Xu C, Hua X S. Ranking model adaptation for domain-specific search. *IEEE Transactions on Knowledge and Data Engineering*, 2012, **24**(4): 745–758
- 91 Tao Jian-Wen, Wang Shi-Tong. Kernel distribution consistency based local domain adaptation learning. *Acta Automatica Sinica*, 2013, **39**(8): 1295–1309  
(陶剑文, 王士同. 核分布一致局部领域适应学习. 自动化学报, 2013, **39**(8): 1295–1309)
- 92 Gao Jun, Huang Li-Li, Sun Chang-Yin. A local weighted mean based domain adaptation learning framework. *Acta Automatica Sinica*, 2013, **39**(7): 1037–1052  
(皋军, 黄丽莉, 孙长银. 一种基于局部加权均值的领域适应学习框架. 自动化学报, 2013, **39**(7): 1037–1052)
- 93 Tao Jian-Wen, Wang Shi-Tong. Kernel support vector machine for domain adaptation. *Acta Automatica Sinica*, 2012, **38**(5): 797–811  
(陶剑文, 王士同. 领域适应核支持向量机. 自动化学报, 2012, **38**(5): 797–811)
- 94 Jenatton R, Audibert J Y, Bach F. Structured variable selection with sparsity-inducing norms. *The Journal of Machine Learning Research*, 2011, **12**: 2777–2824
- 95 Poultney C, Chopra S, Cun Y L. Efficient learning of sparse representations with an energy-based model. In: Proceedings of the 20th Annual Conference on Neural Information Processing Systems. Vancouver, British Columbia, Canada: MIT Press, 2006. 1137–1144
- 96 Christoudias C, Urtasun R, Darrell T. Multi-view learning in the presence of view disagreement. *The Computing Research Repository*, 2012, **12**(6): 1153–1159
- 97 Ganchev K, Graca J, Blitzer J, Taskar B. Multi-view learning over structured and non-identical outputs. *The Computing Research Repository*, 2012, **12**(6): 1206–1214
- 98 Koehn P, Schroeder J. Experiments in domain adaptation for statistical machine translation. In: Proceedings of the 2nd Workshop on Statistical Machine Translation. Prague, Czech Republic: ACL, 2007. 224–227
- 99 He X. Using word dependent transition models in HMM based word alignment for statistical machine translation. In: Proceedings of the 2nd Workshop on Statistical Machine Translation. Prague, Czech Republic: ACL, 2007. 80–87
- 100 Koehn P, Hoang H, Birch A. Moses: open source toolkit for statistical machine translation. In: Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions. Prague, Czech Republic: ACL, 2007. 177–180
- 101 Matsoukas S, Rosti A V I, Zhang B. Discriminative corpus weight estimation for machine translation. In: Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing. Singapore: ACL, 2009. 708–717
- 102 Och F J. Minimum error rate training in statistical machine translation. In: Proceedings of the 41st Annual Meeting on Association for Computational Linguistics. Sapporo Convention Center, Sapporo, Japan: ACL, 2003. 160–167
- 103 Lei D M, Ng A Y, Jordan M I. Latent dirichlet allocation. *The Journal of Machine Learning Research*, 2003, **3**: 993–1022
- 104 Guo H L, Zhu H J, Guo Z L. Domain adaptation with latent semantic association for named entity recognition. In: Proceedings of the 2009 North American Chapter of the Association of Computational Linguistics. Boulder, Colorado, USA: ACL, 2009. 281–289



**刘建伟** 博士, 中国石油大学(北京) 副研究员. 主要研究方向为智能信息处理, 机器学习, 复杂系统分析, 预测与控制, 算法分析与设计. 本文通信作者.

E-mail: liujw@cup.edu.cn

(**LIU Jian-Wei** Ph.D., associate professor in the Department of Automation, College of Geophysics and Information Engineering, China University of Petroleum, Beijing. His research interest covers intelligent information processing, machine learning, analysis, prediction, controlling of complicated nonlinear system, and analysis of the algorithm and the designing. Corresponding author of this paper.)



**孙正康** 中国石油大学(北京) 地球物理与信息工程学院硕士研究生. 主要研究方向为机器学习, 域自适应学习研究.

E-mail: sunzhengkang@126.com

(**SUN Zheng-Kang** Master student in the Department of Automation, College of Geophysics and Information Engineering, China University of Petroleum, Beijing. His research interest covers machine learning and domain adaptation learning.)



**罗雄麟** 博士, 中国石油大学(北京) 教授. 主要研究方向为智能控制和复杂系统分析, 预测与控制.

E-mail: luoxl@cup.edu.cn

(**LUO Xiong-Lin** Ph.D, professor in the Department of Automation, College of Geophysics and Information Engineering, China University of Petroleum, Beijing. His research interest covers intelligent control, and analysis, prediction, controlling of complicated nonlinear system.)