



知识迁移学习 方法及应用

王雪松 潘 杰 程玉虎 著



科学出版社

知识迁移学习方法及应用

王雪松 潘 杰 程玉虎 著

科 学 出 版 社

北 京

内 容 简 介

知识迁移的目标是使机器具有和人一样的“举一反三”的能力,通过已掌握的知识来完成新的任务。知识迁移是机器突破程式化约束,具有自主学习能力的因素,近年来已成为人工智能的研究热点之一。

从“如何迁移”“迁移什么”与“何时迁移”等角度入手,系统阐述了关联知识的表达形式、迁移的方式与手段以及实现迁移的时机等问题。然而知识迁移方法在提升机器智能性的同时,常常遇到“负迁移”的难题,即迁移后效果反而变差,为此,本书从两方面着手:对迁移任务间相似度的衡量给出一定的标准,筛除相关性差的源任务;避免样本的直接迁移,提取关联任务间的公共特征,即在特征空间实现迁移。为将知识迁移理论推广到实际领域,本书围绕文本分类、人脸识别、色彩调和、目标决策与优化规划等一系列任务展开论述,将知识迁移方法分为分类迁移、决策迁移、色彩迁移与优化迁移等4部分进行分析与研究。此外,各章节内容均涉及相关领域基础知识的介绍,能够为不同层次的读者与研究人员提供入门知识与参考信息。

本书既可作为模式识别、数据挖掘相关专业的研究生辅助教材,也可作为机器学习工程师的参考书目。

图书在版编目(CIP)数据

知识迁移学习方法及应用/王雪松,潘杰,程玉虎著. —北京:科学出版社, 2016.12

ISBN 978-7-03-051395-3

I. ①知… II. ①王… ②潘… ③程… III. ①自学习-研究 IV. ①TP181

中国版本图书馆 CIP 数据核字 (2016) 第 315222 号

责任编辑:惠 雪/责任校对:李 影
责任印制:张 倩/封面设计:许 瑞

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

北京通州皇家印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2016年12月第 一 版 开本:720×1000 1/16

2016年12月第一次印刷 印张:13 1/4

字数:267 000

定价:89.00 元

(如有印装质量问题,我社负责调换)

前言

当今时代是知识爆炸的时代,人们通过不同渠道能够获得各种类型的知识,如何合理利用这些知识实现新的目标和任务,即实现知识迁移,是人类一项基本能力。然而这项能力对于机器而言却并不容易,“如何迁移”“何时迁移”以及“迁移什么”是知识迁移应用于人工智能领域面临的三大难题。为使未来的机器破除程式化的束缚,由单一功能向多功能发展,由单纯执行命令向自主学习进化,由机器思维方式向人类思维方式转变,越来越多的研究人员致力于知识迁移原理的探索与应用,使其成为机器学习乃至人工智能领域的研究热点之一。

知识迁移的目标是从关联任务中提取有效信息来辅助解决新的任务,根本出发点是目标领域的数据稀疏性,即目标训练集缺少有效标签或充足样本,如大量的网页文本信息均是无标签数据。当出现数据稀疏性时,训练模型很容易产生过拟合现象。尽管许多机器学习方法,包括半监督学习、协同训练与主动学习能在一定程度上解决数据稀疏性问题,很多情形下仍需要寻求额外的知识以辅助学习。因而研究知识迁移理论对于填补机器学习的空白与推进人工智能的发展有着至关重要的意义。早在 20 世纪八、九十年代,就有学者开始研究相关问题,称之为自教学习、自适应学习、跨领域学习或领域自适应等。直到 2009 年,香港科技大学教授、IEEE Fellow 杨强在 IEEE 会刊上发表了关于知识迁移的综述后,知识迁移开始有了较为完善的理论框架和标准的概念体系。

通常将以下两种观点作为典型的知识迁移:①理论上,知识迁移是一种新的学习范式,不同于监督学习、半监督学习或无监督学习;②实践中,知识迁移是一种新的跨领域学习技术,能够很好地解决各类领域差异性问题的,如数据分布、特征与标签空间、源数据噪声以及源领域与目标领域的关联性等。在过去二、三十年中,知识迁移已在包括文本挖掘、语音识别、计算机视觉、评价系统等诸多领域获得成功,并在许多实际场景中取得良好应用。如百科文档到博客文本、搜索图像到网页页面、书籍评价到电影评价等一系列不同领域间知识的提炼、转换与映射,均从属于知识迁移理论。

本书著者长期从事机器学习的研究工作,在国家自然科学基金(项目编号:61273143, 61472424)、教育部“新世纪优秀人才支持计划”、江苏省自然科学基金以及高等学校博士学科点专项科研基金项目资助下,针对不同的问题背景提出了一系列包括特征迁移、模型迁移与样本迁移等在内的知识迁移理论与方法,并将其成功应用于许多复杂的实际问题。

全书围绕文本分类、人脸识别、辅助决策、色彩调和、参数映射与路径寻优等一系列机器学习与计算机视觉问题，分别从分类迁移、决策迁移、色彩迁移与优化迁移等 4 个不同角度阐述知识迁移原理、方法与应用，共 4 部分 11 章。第 1 部分为分类迁移，内容为第 1~5 章，主要是针对人脸识别与领域适应等分类问题的单源迁移与多源集成迁移。包括：基于局部保持投影的特征迁移、加权多源 TrAdaBoost 算法、多源 Tri-Training 知识迁移与基于样本局部判别权重的加权迁移。第 2 部分为决策迁移，内容为第 6~7 章，围绕决策树与强化学习等策略问题展开。包括：基于相似度衡量的决策树迁移与基于极限学习机的多源强化知识迁移。第 3 部分为色彩迁移，内容为第 8~9 章，涉及计算机视觉领域的局部色彩调和问题。包括：基于主动轮廓探索的多源色彩迁移与基于纹理相似性的多源局部色彩迁移。第 4 部分则是针对路径规划问题中的参数迁移与样本优化迁移。

为便于应用本书阐述的算法，在全书最后的附录提供了部分知识迁移算法源程序。

由于时间仓促，且作者水平有限，书中不当之处在所难免，恳切希望得到广大读者的批评和指正。

作 者

2016 年 8 月于中国矿业大学

目 录

前言

第 1 章 知识迁移概述	1
1.1 知识迁移的提出和发展	2
1.1.1 知识迁移的提出	2
1.1.2 知识迁移的发展	3
1.2 知识迁移的概念	4
1.2.1 知识迁移的基本概念	4
1.2.2 知识迁移与机器学习的关系	5
1.3 知识迁移的分类	7
1.3.1 基于实例的迁移	7
1.3.2 基于参数的迁移	8
1.3.3 基于特征的迁移	9
1.3.4 基于相关知识的迁移	9
1.3.5 基于模型的迁移	10
1.4 知识迁移的主要形式与应用	11
1.4.1 分类迁移	11
1.4.2 决策迁移	16
1.4.3 色彩迁移	20
1.4.4 优化迁移	27
1.5 本书主要内容及安排	30
1.6 本章小结	33
参考文献	33
第 2 章 基于局部保持投影的特征迁移	40
2.1 低维特征表示	41
2.2 基于 LPP 的多源特征迁移	43
2.2.1 系统原理图	43
2.2.2 迁移源筛选	46
2.2.3 LPP 特征迁移	47
2.2.4 算法步骤	49
2.3 试验结果与分析	50

2.3.1	人脸特征迁移	50
2.3.2	不同特征迁移方法效果对比	51
2.3.3	单样本特征迁移精度分析	53
2.4	本章小结	55
	参考文献	55
第 3 章	加权多源 TrAdaBoost 算法	57
3.1	TrAdaBoost 算法简介	57
3.2	加权多源 TrAdaBoost 算法	58
3.2.1	基本符号	59
3.2.2	TrAdaBoost 算法流程	59
3.3	理论分析	61
3.3.1	源领域训练集样本损失值分析	61
3.3.2	目标领域训练集测试误差分析	64
3.3.3	时间复杂度分析	65
3.4	试验及其结果对比	66
3.4.1	双月型数据集试验	66
3.4.2	文本分类试验	70
3.5	本章小结	72
	参考文献	72
第 4 章	多源 Tri-Training 知识迁移	73
4.1	半监督学习简介	73
4.2	多源 Tri-Training 知识迁移	74
4.2.1	Tri-Training 算法	74
4.2.2	多源 Tri-Training 知识迁移	75
4.2.3	算法结构	75
4.2.4	算法流程	75
4.3	试验结果与对比	76
4.4	本章小结	80
	参考文献	81
第 5 章	基于样本分布信息的跨领域知识迁移	82
5.1	领域适应学习	82
5.2	基于样本局部判别权重的加权迁移成分分析	83
5.2.1	问题描述	83
5.2.2	迁移成分分析	84
5.2.3	局部判别权重	84

5.2.4 加权迁移	86
5.3 基于样本分布信息的最大分布加权均值差异	89
5.4 基于 MDWMD 的领域适应学习	91
5.4.1 联合分布调整	91
5.4.2 投影矩阵	92
5.4.3 目标函数	93
5.4.4 算法步骤	93
5.5 试验结果与对比	94
5.5.1 试验数据集	94
5.5.2 试验结果及分析	95
5.6 本章小结	98
参考文献	99
第 6 章 基于相似度衡量的决策树迁移	100
6.1 研究动机	100
6.2 决策迁移数学描述	101
6.2.1 领域与任务	101
6.2.2 决策迁移	102
6.3 多源决策树相似度衡量与误差分析	103
6.4 算法流程与仿真研究	107
6.4.1 算法流程	107
6.4.2 精度与复杂度分析	107
6.4.3 ROC 特性分析	109
6.4.4 文本分类研究	111
6.5 本章小结	112
参考文献	113
第 7 章 基于极限学习机的多源强化知识迁移	114
7.1 强化迁移机制	114
7.2 基于 ELM 的 Q 学习	115
7.3 基于 ELM 的多源迁移 Q 学习	117
7.3.1 系统原理图	117
7.3.2 任务空间与样本空间迁移	119
7.4 算法步骤与仿真分析	121
7.4.1 算法步骤	121
7.4.2 仿真分析	121
7.5 本章小结	126

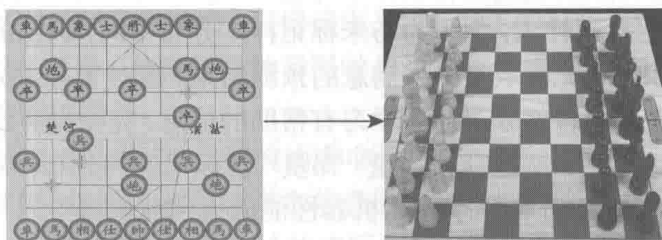
参考文献	127
第 8 章 基于主动轮廓探索的多源色彩迁移	128
8.1 色彩迁移	128
8.2 系统结构	129
8.3 基于主动轮廓探索的多源色彩迁移	131
8.4 试验结果与分析	133
8.4.1 试验设定	133
8.4.2 多源迁移效果影响因素分析	134
8.4.3 与局部迁移算法的对比	137
8.5 本章小结	139
参考文献	139
第 9 章 基于纹理相似性的多源局部色彩迁移算法	141
9.1 局部色彩迁移	141
9.2 多源局部色彩迁移算法	142
9.2.1 爬山分割算法	143
9.2.2 灰度共生矩阵	143
9.2.3 多源局部色彩迁移算法	145
9.2.4 算法流程	146
9.3 试验结果与对比	147
9.3.1 单源局部迁移	147
9.3.2 多源局部迁移	149
9.3.3 灰度图色彩迁移	151
9.4 本章小结	153
参考文献	153
第 10 章 基于图构建的多源参数迁移	155
10.1 参数迁移	155
10.2 ACO 参数邻接迁移矩阵	156
10.2.1 ACO 参数分析	156
10.2.2 邻接迁移矩阵	157
10.3 基于扩展模型图的多源参数迁移	158
10.4 目标任务的参数优化	159
10.5 仿真研究与分析	160
10.5.1 环境设定	160
10.5.2 迁移性能分析	160
10.6 本章小结	164

参考文献	164
第 11 章 基于样本空间筛选的优化迁移	166
11.1 Ant-Q 优化迁移	166
11.2 多源迁移 Ant-Q 算法	167
11.3 基于先验概率的样本空间筛选	169
11.4 TSP 仿真研究	171
11.4.1 单源迁移 TSP 问题	171
11.4.2 多源迁移 TSP 问题	175
11.5 本章小结	176
参考文献	176
附录 部分知识迁移学习算法 MATLAB 源程序	178
程序 1 第 2 章基于局部保持投影的特征迁移 MATLAB 源程序	178
程序 2 第 6 章基于相似度衡量的决策树迁移部分 MATLAB 源程序	184
程序 3 第 8 章基于主动轮廓探索的色彩迁移部分 MATLAB 源程序	192
索引	199

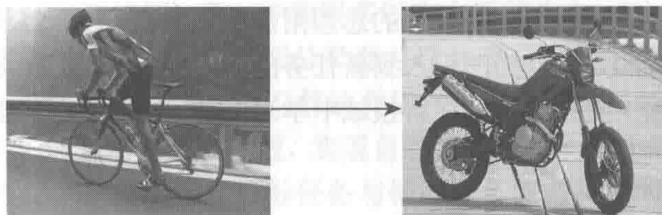
第1章 知识迁移概述

心理学认为,知识迁移是反映人类心理认知过程的一种概念;是人类学习的一种方法;通过一种学习影响另一种学习。古代人类就注意到迁移现象,春秋时期的思想家孔子曾提出“举一反三”的说法,宋代朱熹提出“融会贯通”的思想,这些思想都是说明学习者需汇聚各方面知识,掌握不同门类甚至是不同领域知识的共通性,从而可以将在一个领域内对知识的理解转移至其他领域当中。

在日常生活和学习中,我们会发现,数学学习好的学生物理学习一般也很好;会打羽毛球的学生学习网球会很快;会下象棋的人学习国际象棋会很容易;会骑自行车的人也会很快学会骑摩托车,如图 1-1 所示。因此,我们发现,如果掌握了其中一个领域的知识,在学习另外一个相近领域知识的时候,已掌握的知识会对新知识的学习起到帮助作用,这就是一种知识的迁移。这种知识迁移可以大大增强人类在不同环境中学习新任务的能力。



(a)象棋与国际象棋的迁移



(b)自行车与摩托车的迁移

图 1-1 日常生活中的迁移现象

本章将简要介绍知识迁移的提出和发展、基本概念与分类、主要表现形式及其应用等,最后给出本书的主要结构和章节安排。

1.1 知识迁移的提出和发展

1.1.1 知识迁移的提出

当今世界是信息的世界, 计算机技术发展迅速, 对于获取并存储海量信息已不再是难题, 然而, 如何处理这些获取的海量信息则成为了一个新的挑战。在此基础上, 机器学习的概念与方法被提出并得到了长足的发展。机器学习是模拟人类思维的一种计算机学习方法, 是人工智能的核心, 然而传统机器学习与人类学习相比, 并不具有学习新环境知识的能力, 不能利用已掌握的知识帮助新任务的学习, 而是需要足够的标记样本来训练出可靠的分类模型, 然后将模型用于未标记样本获取其标签值。训练与测试样本必须服从同分布假设, 如果分布不同或训练样本数量不足都难以获得可靠的分类模型, 这极大限制了机器学习的发展及应用。如果让机器拥有人类类似的知识迁移能力, 学习不同领域的知识, 则能够在很大程度上提高机器的学习效率, 丰富学习内容。

实际应用中常遇到的问题是, 我们很容易获取大量未标记样本, 然而, 给样本赋予标记的手段往往是手工操作, 这个过程将消耗大量时间与资源。另外, 由于信息的增长和更新, 许多得到的已标记样本与新增加的未标记样本分布存在偏差, 虽然能获得足够的已标记样本, 但因为与未标记样本分布不同, 训练的分类模型一般不具有很好的泛化性, 无法取得令人满意的预测结果。而对于与未标记样本分布不同的相似样本, 常常包含部分对任务学习有帮助的信息, 完全抛弃这部分样本则会造成大量资源浪费。因此, 在获得少量“昂贵”已标记的同分布样本的情况下, 如何利用大量已标记相似分布样本提高机器性能成为当前机器学习与模式识别研究中的热点问题。

知识迁移就是在这样的应用背景下产生的, 目的是打破传统机器学习训练样本与测试样本同分布的假设, 将迁移的思想附加给计算机, 让计算机拥有跨领域学习的能力, 将已有的知识恰当的引入到新任务的学习中, 使相似领域中包含信息的有效重用成为可能, 提高计算机在新领域中学习的能力。知识迁移学习与传统机器学习的关系如图 1-2 所示。

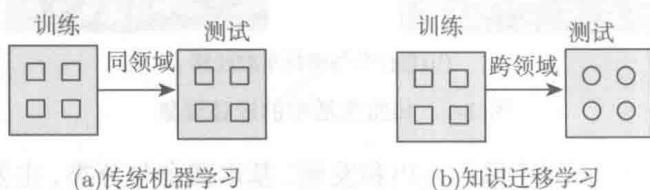


图 1-2 传统机器学习与知识迁移学习的关系

知识迁移研究更关注当训练样本信息不足的情况下,如何通过已有的相似领域知识获得具有良好预测效果以及推广能力的学习模型,这里的相似包括样本分布相似或学习任务相似。知识迁移的理论研究对于深化机器学习理论,扩大其应用领域具有重要现实意义,并且对机器学习存在的许多理论问题,如相似域间信息重用、模型参数设计、知识迁移与传统学习以及半监督学习关系等都具有重要指导意义。

1.1.2 知识迁移的发展

近些年机器学习理论发展迅速,知识迁移在理论和应用中得到快速发展。目前,知识迁移研究成果已广泛应用于网页检索和文本分类、数字图像处理、图像识别、自然语言处理、邮件分类、计算机辅助设计以及无线网络定位等领域中。

实际上,知识迁移最早是用来研究人类心理活动的一种方法。心理学认为,人类学习认知事物并将其“举一反三”的过程就是一种迁移行为。20世纪90年代,知识迁移研究被引入到机器学习理论中,概念一经提出,就在学术界引起广泛关注。最初,关于知识迁移的定义很宽泛,经历了终身学习、迁移学习、归纳迁移、多任务学习、元学习和增量(积累)学习等不同阶段。2005年,美国国防部先进研究项目局给出了一个新的知识迁移定义:知识迁移是系统将在已有环境中认知和学习到的信息应用到新任务的能力。根据该定义,知识迁移旨在将已有源任务中获取的信息重用于目标任务。与多任务学习同时关注复数个学习任务不同,知识迁移更注重目标任务性能的提升。

知识迁移自提出以来,涉及研究内容非常广泛。例如,对各种传统机器学习算法进行修改和扩展,以融入源领域样本信息的研究,或者利用源领域中获取的知识在目标领域重构学习器等。知识迁移的实际应用内容也非常丰富,概括来说有:基于语义的文本迁移和情感分类、基于颜色空间转换的色彩迁移以及利用概率统计理论的关联任务样本迁移等。第一种应用多以文本迁移为主,Yu等考虑语句中的多元关键字,将语义分割,构建不同分类器并迁移至新语句中实现文本数据挖掘功能^[1];Blitzer等改进结构对应学习算法并提出一种确定域与域之间的相关性的度量方法,解决了情感语义标注问题,实现自动情感分类^[2];刘成等构造衡量任务相似性的相关性矩阵,提出通过目标任务与辅助任务联合决策的方式提高目标任务的分类精度,显著改善图像分类的准确率^[3]。第二种应用则是将色彩分布作为一种可利用的知识进行迁移,代表是Reinhard算法,其在图像颜色空间转换的基础上,对空间像素值进行线性变换,成功将颜色图像的颜色基调转移至形状图像,实现图像之间颜色的迁移^[4]。第三种应用则是基于模型与样本权重的样本迁移,Yang等考虑用户动态和环境因素,利用隐马尔可夫模型捕捉到的时间和拓扑结构问题域,根据当前接收的RSS信号预测用户位置来解决用户无线定位问题^[5];王雪松

等通过贝叶斯理论分析源任务与目标任务的相似性,按照迁移价值对源任务排序,筛选有用样本,指导 Agent 在路径规划问题中快速合理做出决策^[6]; Dai 等利用 Boosting 机制,对目标领域样本和源领域样本分别赋予不同的权重,改变不同样本在训练中的重要性,将源领域样本添加至训练集共同训练,很好地实现了 Web 网页的自动分类^[7];杜俊卫使用聚类的方法选择与目标领域相似的源领域文本,并将这些文本应用到目标领域的学习中,增加训练样本的个数,大大提高了垃圾邮件过滤技术的性能^[8]。

如今,知识迁移的研究内容已频繁出现在一些国际顶级期刊、学术会议中,如 IJCAI^[9]、AAAI^[10]、ICML^[11]、ICAPS^[12]、CVPR^[13]、NIPS^[14]、ICCV^[15]、KDD^[16]、Machine Learning^[17]与 Journal of Machine Learning Research^[18]等。由于知识迁移可以降低人工标注代价并在提高学习性能方面具有极大优势,因此越来越受到国内外学者青睐。知识迁移在理论研究方面与实际应用中均获得巨大成果,而且在机器学习中的研究地位逐渐升高。

1.2 知识迁移的概念

本节介绍知识迁移的一些基本概念以及与传统机器学习之间的关系。

1.2.1 知识迁移的基本概念

1. 领域

领域 D 通常包含两部分:特征空间 χ 和边缘概率分布 $P(X)$,即 $D = \{\chi, P(X)\}$,其中 $X = \{x_1, x_2, \dots, x_n\} \subset \chi$ 。例如,在 Web 网页分类问题中, χ 即为网页特征空间, X 为网页集合, x_n 表示第 n 个网页, $P(X)$ 则是网页的类属。一般来说,如果两个领域特征空间 χ 不同或者边缘概率分布 $P(X)$ 不同,则认为它们是不同的。在知识迁移中,领域通常分为源领域与目标领域。

2. 任务

对于给定的一个领域 $D = \{\chi, P(X)\}$,对应的任务 T 包含两部分:标记空间 Y 和预测函数 $f(\cdot)$,即 $T = \{Y, f(\cdot)\}$,其中 $f(\cdot)$ 是通过训练样本 $\{x_i, y_i\}$ 学习得到的,可以对新的样本 x 预测其标签 $f(x)$, $x_i \in X$, $y_i \in Y$ 。从概率的角度讲, $f(x) = P(y|x)$ 。例如,在网页分类中, Y 可以看作所有网页的标签集合, y_i 为其中一类类属, $f(x)$ 即是判断样本 x 属于哪一类。与领域类似,如果两个任务不同,则它们可能是标记空间 Y 不同,或者是预测函数 $f(\cdot)$ 不同。

3. 知识迁移

知识迁移就是利用从源领域 D_S 和源任务 T_S 中获取的相关知识提高目标任务

T_T 在目标领域 D_T 上的预测准确率, 其中 $D_S \neq D_T$ 或 $T_S \neq T_T$ 。

从知识迁移定义可以看出, 其解决的是不同领域之间学习的问题, 当源领域 D_S 与目标领域 D_T 相同, 且源任务 T_S 与目标任务 T_T 相同时, 学习问题就变成了传统机器学习。

4. 相关性

知识迁移的研究对象是分布不同但具有相关性的领域, 迁移的有效性在很大程度上依赖于领域间的相似程度。领域间的相关性越大, 领域分布越一致, 迁移效果就越明显。反之, 领域分布差异越大, 相关性越弱, 迁移效果就越差。因此衡量领域间的相关性尤为重要。

5. 负迁移

通过相关性的知识, 我们了解到源领域与目标领域间的关联程度直接影响迁移的效果。当领域间相关性很小, 存在较大分布差异时, 源领域中的知识可能没有提高目标任务学习的准确率, 反而降低了学习器的性能, 此时就产生了负迁移现象。

例如, 我们学会骑自行车, 再学习骑三轮车, 会感觉车身不受控制, 掌握不了平衡。因为在骑自行车转方向时身体需要向转弯方向倾斜, 而骑三轮车时不需要身体倾斜来维持平衡, 相反身体会由于离心作用偏向外侧。如果仍然依靠自行车技术学习三轮车, 则很难把握车身平衡, 如图 1-3 所示。

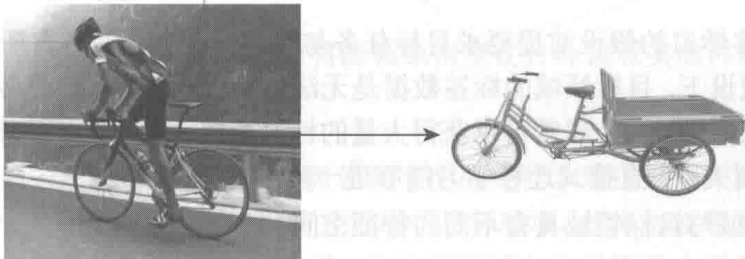


图 1-3 日常生活中的负迁移现象

当前大部分知识迁移工作都围绕在如何计算源领域与目标领域的相关性并寻找对目标任务有用的信息, 以及如何将这部分知识迁移至目标领域等方面, 而针对避免负迁移现象这个关键问题却研究较少。

1.2.2 知识迁移与机器学习的关系

知识迁移与传统机器学习之间的关系总结如表 1-1 所示, 针对源领域与目标领域间的不同关系, 将知识迁移分为三大类: 归纳迁移、无监督迁移与直推迁移^[19]。

表 1-1 传统机器学习与迁移学习之间的关系

学习设定		训练样本	源与目标域	源与目标任务
传统机器学习		大量	相同	相同
知识迁移	归纳迁移	少量	相同 or 相关	不同&相关
	无监督迁移	无	相同 or 相关	不同&相关
	直推迁移	少量	不同&相关	相同

1. 归纳迁移

对于归纳迁移而言，通常认为目标任务与源任务并不相同，而源领域与目标领域的限制条件则较为宽松。一般来说，目标领域的标签数据应对目标域预测函数 $f_T(\cdot)$ 的归纳总结具有指导意义。根据标签数据能否在源领域中获得，归纳迁移能够进一步分为以下两种情形：

(1) 源领域具有能大量访问的标签样本。归纳迁移在该种情形下与多任务学习的设定相类似，两种学习算法的区别在于前者仅对从源任务中高效迁移知识以解决目标任务感兴趣，而后者通常会以平等的地位学习源任务与目标任务。

(2) 源领域无法获得标签数据。这种情形归纳迁移学习的设定类似于自教学习。自教学习是由 Raina 等^[20]首先提出的，其要求源领域与目标领域具有不同的标签空间，这意味着目标领域无法直接从源领域一侧获取可利用的信息，正好与归纳迁移中源领域无法获得标签的设定类似。

2. 直推迁移

直推迁移学习的假设前提要求目标任务与源任务相同，而各自所处的领域不同。在这种假设下，目标领域的标签数据是无法访问的，这使目标任务的学习变得困难，但是在源领域中，通常可以获得大量的标签数据。同样地，根据源领域与目标领域的不同关系，直推式迁移学习能够进一步分为以下情形：

- (1) 源领域与目标领域具有不同的特征空间，即 $X_S \neq X_T$ ；
- (2) 源领域与目标领域所处的特征空间相同，即 $X_S = X_T$ ，但其相应的边界概率分布不同，有 $P(X_S) \neq P(X_T)$ 。

情形 (2) 下的直推式迁移学习本质上属于一种具有迁移知识能力的领域自适应问题，其在样本偏差选择，文本分类与协方差转移等方面均有相关应用。

3. 无监督迁移

与归纳迁移学习的设定相类似，目标任务不同于源任务但具有相关性。其区别在于无监督迁移学习关注于解决目标领域无监督的学习任务，如聚类、降维、与密度估计。这种情形下，源领域与目标领域均无法获得标签数据。

基于如表 1-1 设定下的各类迁移学习，其关联设定以及源域、目标域标签的可见性与解决的任务类型，如表 1-2 所示。

表 1-2 迁移学习不同设定

学习设定	关联设定	源域标签	目标域标签	任务
归纳迁移学习	多任务学习	可获得	可获得	回归，分类
	自教学习	无法获得	可获得	回归，分类
直推迁移学习	领域适应/样本偏差选择/ 协方差转移	可获得	无法获得	回归，分类
无监督迁移学习	—	无法获得	无法获得	聚类，降维

1.3 知识迁移的分类

本节内容总结了知识迁移方面的主要工作与进展，并将知识迁移方法划分为 5 类。

- (1) 基于实例的迁移。研究如何利用辅助数据样本。
- (2) 基于参数的迁移。研究如何寻找目标领域模型与源领域模型共享的一些有益于迁移的参数。
- (3) 基于特征的迁移。研究如何通过特征转换或特征学习联系源领域与目标领域。
- (4) 基于相关知识的迁移。研究如何建立目标领域与源领域相关知识的映射关系。
- (5) 基于模型的迁移。研究如何将源领域模型在目标领域实现再利用。

1.3.1 基于实例的迁移

基于实例的迁移学习的特点是源领域中的部分样本通过赋予不同的权重显示其与目标领域之间不同的相关性，将之重用于目标领域，样本加权和重要性抽样是基于实例的迁移学习的主要研究内容。虽然源领域中的训练样本不能直接参与目标领域的训练，但是有部分样本仍然与目标样本分布相似，可用通过加权重用于目标领域，因此，实例迁移学习最直观也最容易理解。基于实例的迁移学习可以用在归纳迁移学习和直推迁移学习中。

对于样本迁移方法，文献 [7] 提出一种 TransferBoost 算法，其采用一种新的基于集合的 Boosting 形式重新利用源数据样本学习目标任务。TransferBoost 算法基于迁移能力分别提升各源任务性能，提高具有正迁移能力的任务权值并降低负迁移的任务权值。同时，TransferBoost 也能够实现对每个源任务有规律的提升，从而调节个体样本的权值。基于这些权值，TransferBoost 能通过源数据增强目标训