

# 迁移学习算法与理论

杜春/李尧

***DBRG***



2019/11/07

# 迁移学习算法与理论



**Associate Professor, Ph.D Supervisor**

School of Software, Tsinghua University

Machine Learning Group

National Engineering Lab for Big Data Software

[longmingsheng@gmail.com](mailto:longmingsheng@gmail.com)

[mingsheng@tsinghua.edu.cn](mailto:mingsheng@tsinghua.edu.cn)

Room 11-413, East Main Building, Tsinghua University, Beijing, China

*Reminder to prospective students: My PhD and Master positions for class 2020 have been filled.*

龙明盛：

2004-2008，清华大学，电机工程与应用电子技术系电气工程及其自动化专业

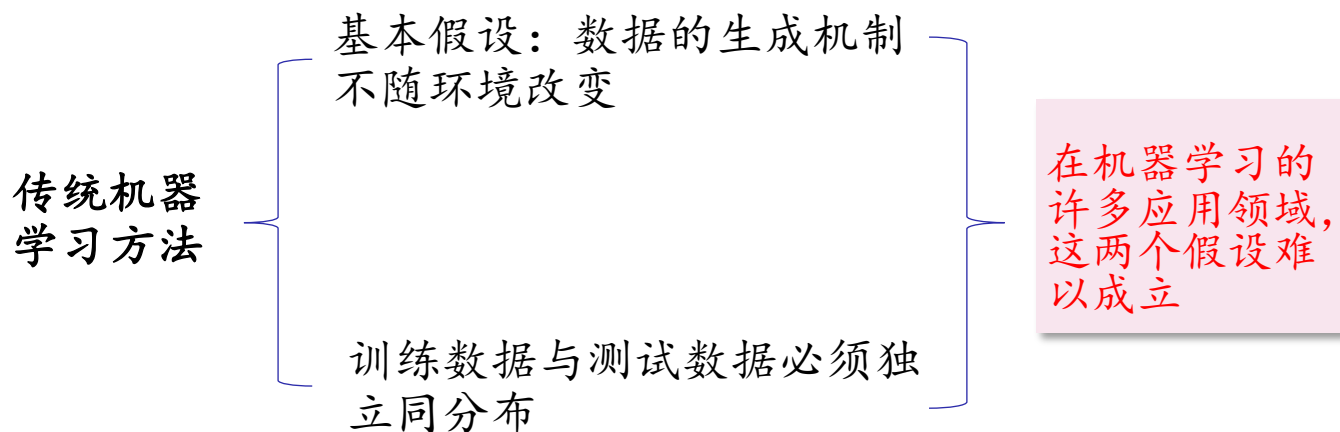
2008-2014，清华大学，计算机科学与技术系博士

2014-2015，UC Berkeley，访问学者

2014-2016，清华大学，软件学院，博士后

研究方向：深度学习，迁移学习，元学习(Meta Learning)，预测学习  
(Predictive Learning)

# 从传统机器学习到迁移学习



如何分析和挖掘非平稳环境中的大规模数据？

## 迁移学习

- 能够在彼此不同但又相互关联的两个领域间挖掘领域不变的本质特征和结构，使得标注数据等有监督信息可以在领域间实现迁移和复用。
- 解决目标任务标注数据稀缺的基本方法
- 存在过拟合、欠拟合、欠适配、负迁移等关键问题与挑战

# 问题定义

## 定义：

给定标注的辅助领域  $\mathcal{D}_s = \{(x_1^{(s)}, y_1^{(s)}), \dots, (x_n^{(s)}, y_n^{(s)})\}$  和学习任务  $\mathcal{T}_s$ ，无标注的目标领域  $\mathcal{D}_t = \{(x_1^{(t)}, \dots, x_{n_t}^{(t)})\}$  学习任务  $\mathcal{T}_t$ ，迁移学习的目标是在  $\mathcal{D}_s \neq \mathcal{D}_t$  或者  $\mathcal{T}_s \neq \mathcal{T}_t$  条件下降低目标领域预测模型  $f_t(x)$  的泛化误差。



辅助领域 = 图书

This is an **excellent** and **broad** survey of the development of civilization with all the punch of high **quality fiction**. A **thrilling** book.

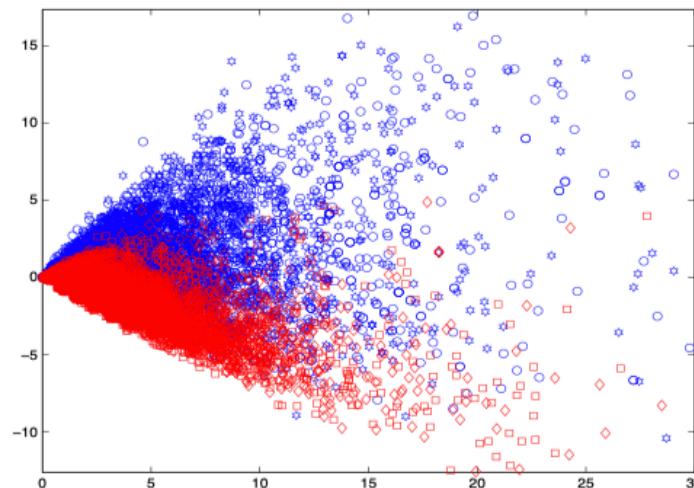
迁移什么？  
如何迁移？

迁移学习

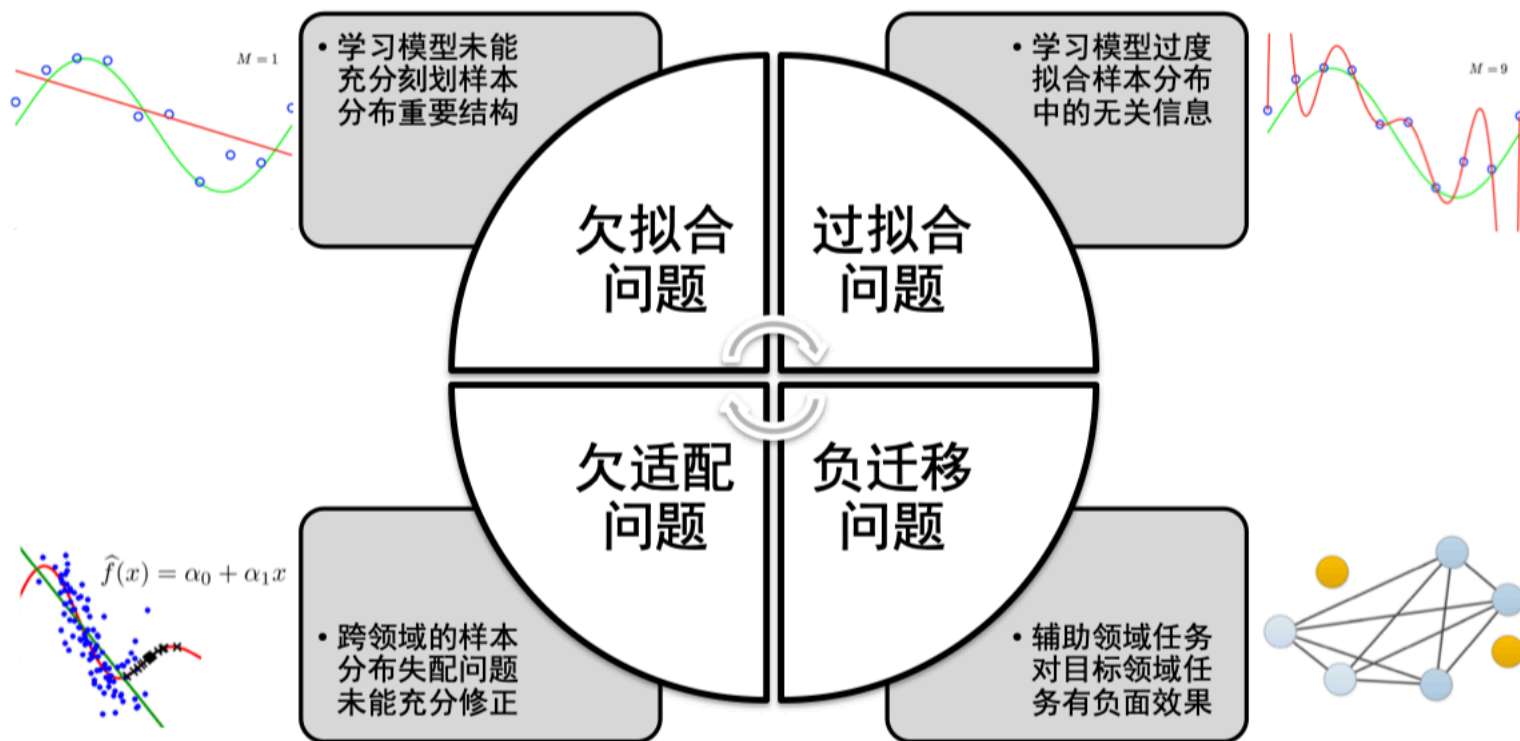


目标领域 = 家具

This is both a **light weight** and **sharp** knife for the kitchen use. An **excellent** bargain considering what I had to pay.



# 迁移学习的主要问题与挑战



迁移学习的主要问题挑战：包括经典机器学习的过拟合、欠拟合问题，以及迁移学习特有的欠适配、负迁移问题。这些问题挑战交错叠加，大大增加了问题解决的难度。

# 深度学习时代的迁移学习

## 背景

继Jason Yosinski在2014年的NIPS上的《How transferable are features in deep neural networks?》探讨了深度神经网络的可迁移性以后，有一大批工作就开始实际地进行深度迁移学习。

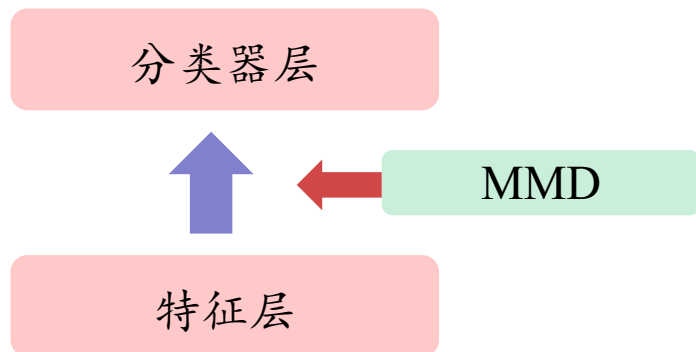
对于一个深度网络，随着网络层数的加深，网络越来越依赖于特定任务；而浅层相对来说只是学习一个大概的特征。不同任务的网络中，浅层的特征基本是通用的。这就启发我们，如果要适配一个网络，重点是要适配高层——那些task-specific的层。

---

Yosinski, Jason, et al. "How transferable are features in deep neural networks?." *Advances in neural information processing systems*. 2014.

# DaNN(Domain Adaptive Neural Network)

MMD(Max mean discrepancies), 用于计算源域与目标域的特征距离, 并将该参数引入损失函数。



整个网络的优化目标:

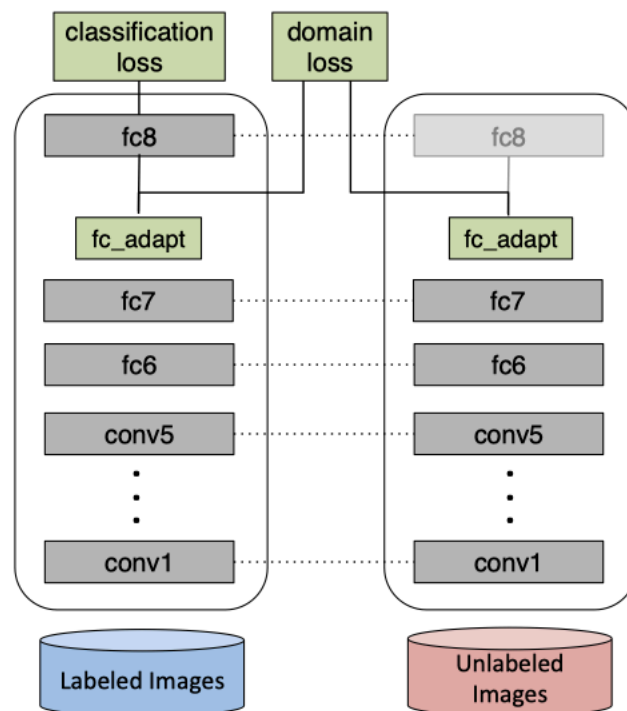
$$\ell = \ell_c + \lambda \ell_D$$

$\lambda$ 是权衡网络适配权重的一个参数, 可以人工指定。

---

Ghifary, Muhammad, W. Bastiaan Kleijn, and Mengjie Zhang. "Domain adaptive neural networks for object recognition." *Pacific Rim international conference on artificial intelligence*. Springer, Cham, 2014.

# DDC(Deep Domain Confusion)

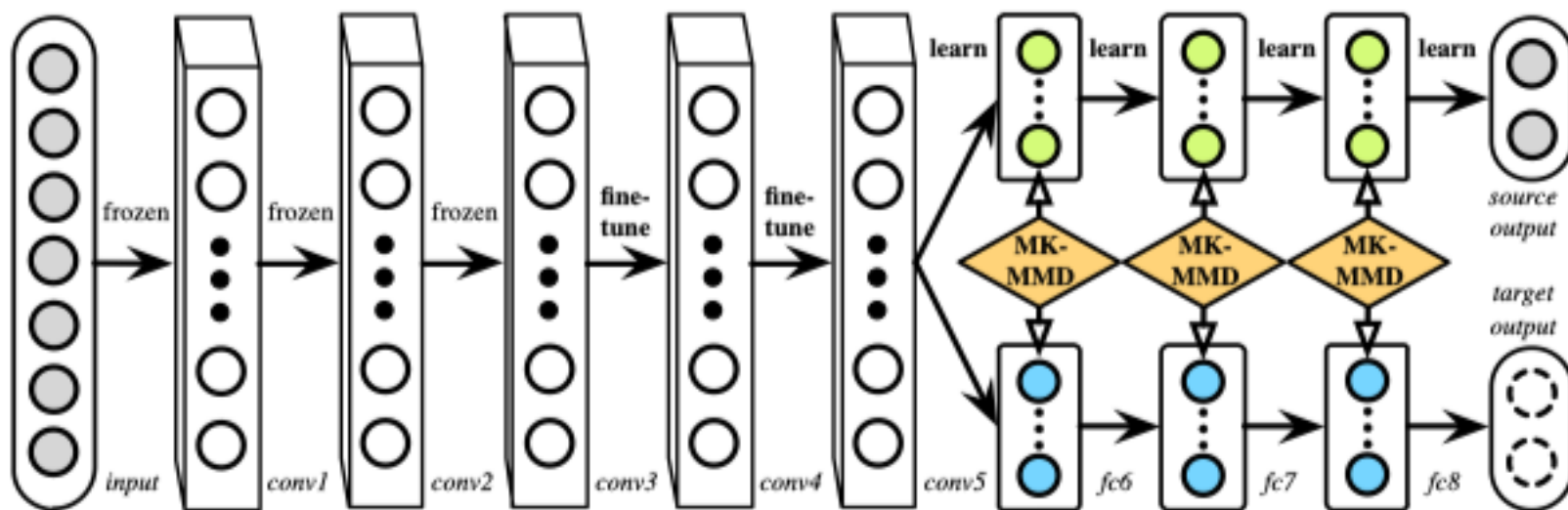


在原有的AlexNet网络的基础上，对网络的fc7层（分类器前一层）后加一层适配层（adaptation layer）。适配层的作用是，单独考察网络对源域和目标域的判别能力。

Tzeng, Eric, et al. "Deep domain confusion: Maximizing for domain invariance." *arXiv preprint arXiv:1412.3474* (2014).



# DAN(Deep Adaption Network)



原来的DDC方法只适配了一层，如果后面多层都task-specific怎么办？

MMD计算需要将source和target用一个相同的映射映射在一个再生**希尔伯特核**空间(RKHS)，关于核的选取，高斯核或者线性核？

Long M, Cao Y, Wang J, et al. Learning transferable features with deep adaptation networks[C]International Conference on Machine Learning. 2015: 97-105.

# 展望与思考

- 迁移学习是否一定有用?
  - 相关性的定义
  - AlphaGo zero vs AlphaGo Lee

欢迎批评指正！