

# 迁移学习简明手册王晋东

---

- 常见度量准则
  - 欧氏距离、马氏距离、闵可夫斯基距离
  - 余弦相似度、互信息、皮尔逊相关系数、Jaccard相关系数
  - KL散度(相对熵)、JS距离
  - 最大均值差异(MMD)
- 迁移学习基本方法
  - 基于样本
  - 基于特征
  - 基于模型(参数)
  - 基于关系
- 数据分布自适应
- 特征选择: SCL(Structural Correspondence Learning)
- 深度迁移学习
  - Fine-tune
    - 网络越深, feature越specific, 迁移效果越差。
    - 基于大型公开数据集训练出的模型通常具有很强的泛化能力, 保留预训练好的网络的前几层(通常学习到的是general feature), 针对自己的任务换上适配的全连接层, 并用自己的小规模数据集进行Fine-tune, 可以取得很棒的效果。
    - 最简单的深度迁移方法: Fine-tune, 学习率要设定得较小。
    - 好处: 节约时间算力, 模型更鲁棒, 泛化能力强。
    - 坏处: 无法很好地处理训练数据与预测数据分布不同的情况。
  - 深度网络自适应
    - 概述
      - 功能: 减少样本数据分布与预测数据分布的距离
      - 核心思考角度:
        - 1、哪些层是可以自适应的(决定网络学习程度)
        - 2、采用何种度量准则(决定网络泛化能力)
      - 设计结构的基本准则: 决定自适应层, 在这些层进行自适应度量, 最后对网络进行Fine-tune
    - 核心方法
      - DDC、DAN、同时迁移领域和任务、深度联合分布自适应、AdaBN
  - 深度对抗网络迁移