概述

教学信息

- 讲授教师：白刚
- 办公地点：计算机东楼564室
- 电子邮箱：baigang@nankai.edu.cn
- 课程信息：http://cc.nankai.edu.cn
  - 授课讲稿：“云班课”平台
  - 课后作业：“云班课”平台
  - 阅读文献：“云班课”平台
- 网络资源
  - 自行搜索中英文《机器学习》课程资源

课程信息

- 教学相长：教与学的“交互与协同”
  - 《礼记•学记》：学然后知不足，教然后知困。知不足，然后能自反也；知困，然后能自强也。故曰教学相长也。
- 学问：学与问的“化学反应”
- 要求：诚信 + 创新 + 实践
- 纪律：
  - 严禁迟到、早退
  - 手机关闭或静音，严禁接打电话
  - 衣冠整洁

网上教学方式

- 授课
  - 视频直播：QQ群直播间
  - 音频PPT：云课堂
  - 签到：云课堂（10:00~10:10)
  - 测验：云课堂
  - 讨论：QQ群 + 微信群
- 课后
  - 作业：云课堂
  - 参考资料：云课堂
  - 答疑：微信群

课程成绩

- 课程成绩：平时作业(60%) + 课程项目(40%)
- 平时作业（个人）
  - 问题描述
  - 基本思路
  - 解题步骤
  - 算法
  - 结果与分析
- 课程项目（小组）
  - 课堂报告
  - 研究报告
  - 项目成果

平时作业内容（推理类）

- 问题描述：10%
  - 使用自己的语言描述所要解决的问题及其难点
- 基本思路：20%
  - 简单叙述解决问题所采用方法的基本原理和特点
- 解题步骤：70%
  - 具体推导内容

平时作业内容（编程类）

- 问题描述：10%
  - 使用自己的语言描述所要解决的问题及其难点
- 基本思路：15%
  - 简单叙述解决问题所采用方法的基本原理和特点
- 解题步骤：35%
  - 处理流程
  - 算法描述
  - 算法实现（Python语言源代码）
- 结果与分析：40%
  - 实验内容与步骤
  - 实验结果
  - 结果分析

平时作业要求

- 独立完成，鼓励讨论，严禁抄袭。
  - 抄袭：把别人的“作品”未加注释的照抄作为自己的成果
  - 凡抄袭成果一律按零分计
- 按时完成且提交，迟交作业按零分计。
- 提交内容：文档（ + 源程序）
  - 文档：简明扼要，重点突出，避免空洞和套话
  - 源程序：必要的注释
- 文档格式：DOCX 或 PDF
- 编程语言：python
- 提交文件名称：学号(作业序号)，如：2120180000(1)
- 提交文件格式：RAR压缩文件格式

课程项目内容

- 课堂报告：20%
  - 总体思路
  - 特征描述
  - 模型选择
  - 实验设计
- 研究报告：60%
  - 问题提出
  - 现状与分析
  - 主要原理
  - 算法描述
  - 结果与分析
- 项目成果：20%
  - 客观评价结果截图

课程项目要求

- 按时完成并提交全部内容
- 提交文档格式：DOCX 或 PDF
- 编程语言：python
- 提交文件名称：学号1+学号2+...+学号n
- 提交文件格式：RAR压缩文件格式
- 加分因素：项目内容公开发表
- 完成形式：2~5人小组（具体人数视最终选课人数确定）

参考书目

- Christopher M. Bishop, ***Pattern Recognition and Machine Learning***, Springer Science+Bussiness Media, LLC, 2006.
- Kevin P. Murphy, ***Machine Learning: A Probabilistic Perspective***, The MIT Press, 2012.
- Ian Goodfellow, Yoshua Bengio, Aaron Courville, ***Deep Learning***, http://www.deeplearningbook.org.
- Tom M. Mitchell, ***Machine Learning***, McGraw-Hill Companies Inc., 1997.

学术期刊

- Artificial Intelligence
- Computational Intelligence
- IEEE Trans. on Knowledge and Data Engineering
- IEEE Trans. on Pattern Analysis and Machine Intelligence
- IEEE Trans. on Neural Networks
- Journal of Machine Learning Research
- Machine Learning
- 模式识别与人工智能
- 自动化学报

学术会议

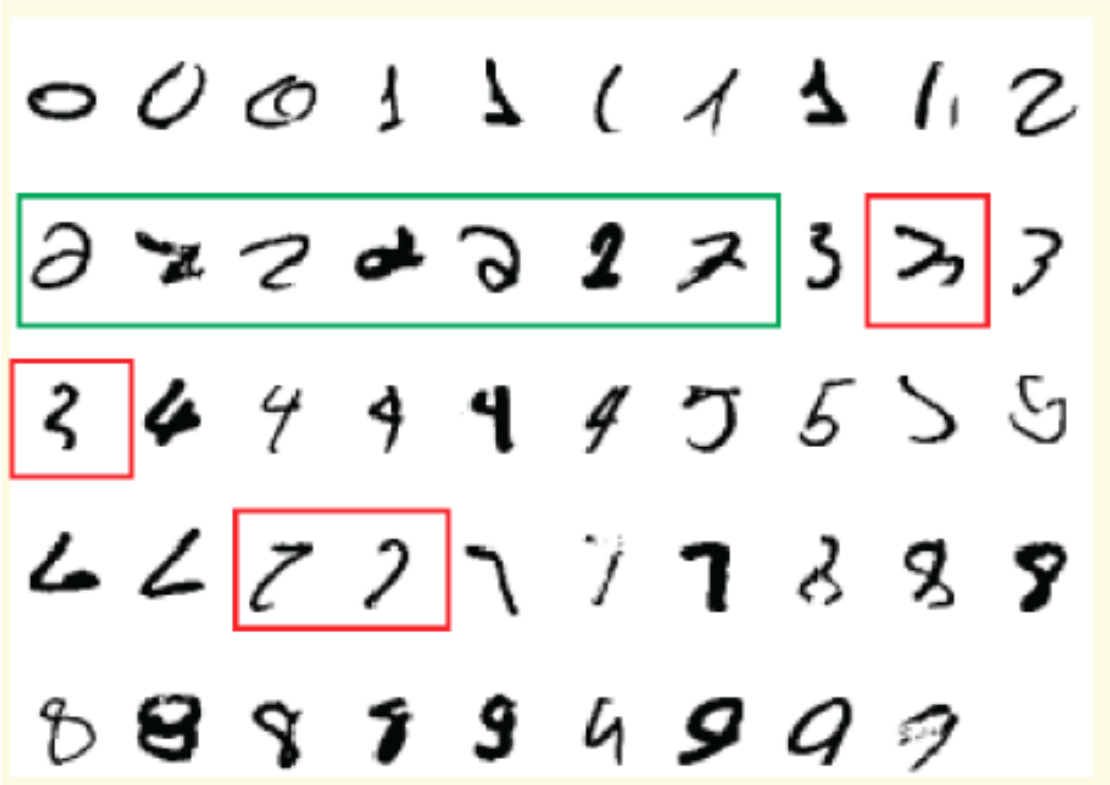
- Annual Conference on Neural Information Processing Systems (NeurIPS)
- International Conference on Machine Learning (ICML)
- The Conference on Uncertainty in Artificial Intelligence (UAI)
- International Conference on Artificial Intelligence and Statistics (AISTATS)

Any question?

学习问题

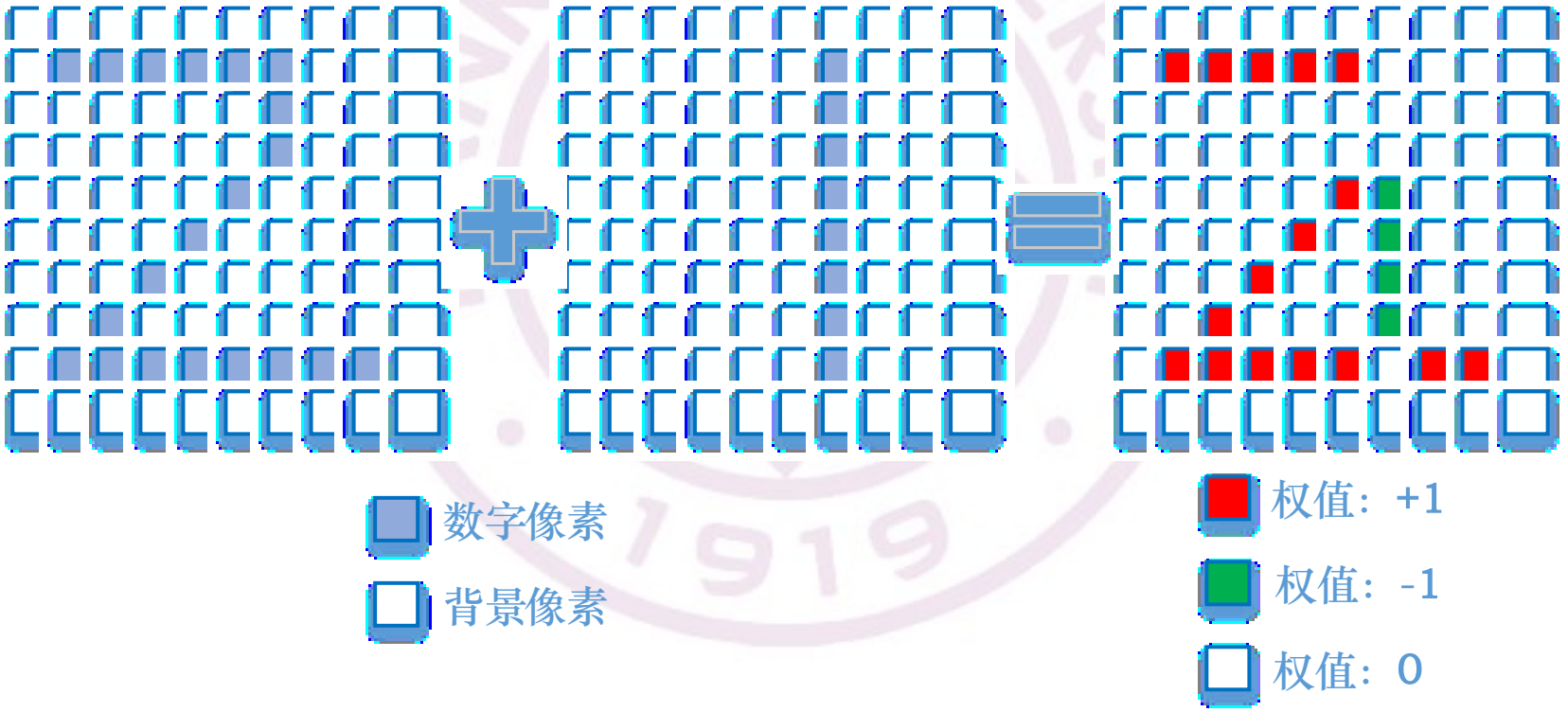
手写数字识别

- 编程实现手写数字识别
  - 确定性编程方法难于完成
  - 利用机器学习算法从训练数据中学习决策函数（建模）



机器学习方法求解

- 训练过程
  - 如果　某个像素属于数字 2
  - 则　权值矩阵对应元素 加 1
  - 否则　权值矩阵对应元素 减 1。



机器学习方法求解

- 测试过程
  - 待识别图像与权值矩阵按元素相乘，然后对所有元素乘积求和；
  - 如果求和为正数，则图像中数字“为 2”，否则图像中数字“不为 2”。
- 方法的缺点：
  - 手写数字在样本图像中的相对位置移动将带来识别错误。
- 有限的训练数据
  - 模型的记忆能力不能够带来好的预测能力
  - 准确预测新数据的关键是模型的泛化(generalization)能力
  - 模型的泛化能力类似于人类的归纳能力

机器学习的定义

- Arthur Samuel (1959)：一个不用直接编程就能给予计算机学习能力的研究领域。
- Tom Mitchell (1998)：如果计算机程序完成某项任务(**T**)的性能(**P**)能够通过某些经验(**E**)得到改进，则称其可通过经验来学习。
  - 在垃圾邮件检测中，“垃圾邮件检测”为任务 T，“检测的成功率”为性能 P，包含大量正常邮件和垃圾邮件的“邮件样本集”为经验 E。

机器学习类型

机器学习类型

- 主要类型
  - 监督学习(supervised learning)
  - 无监督学习(unsupervised learning)
  - 增强学习(reinforcement learning)
- 其它类型
  - 半监督学习(semi-supervised learning)
  - 元学习(meta learning or learning to learn)

监督学习

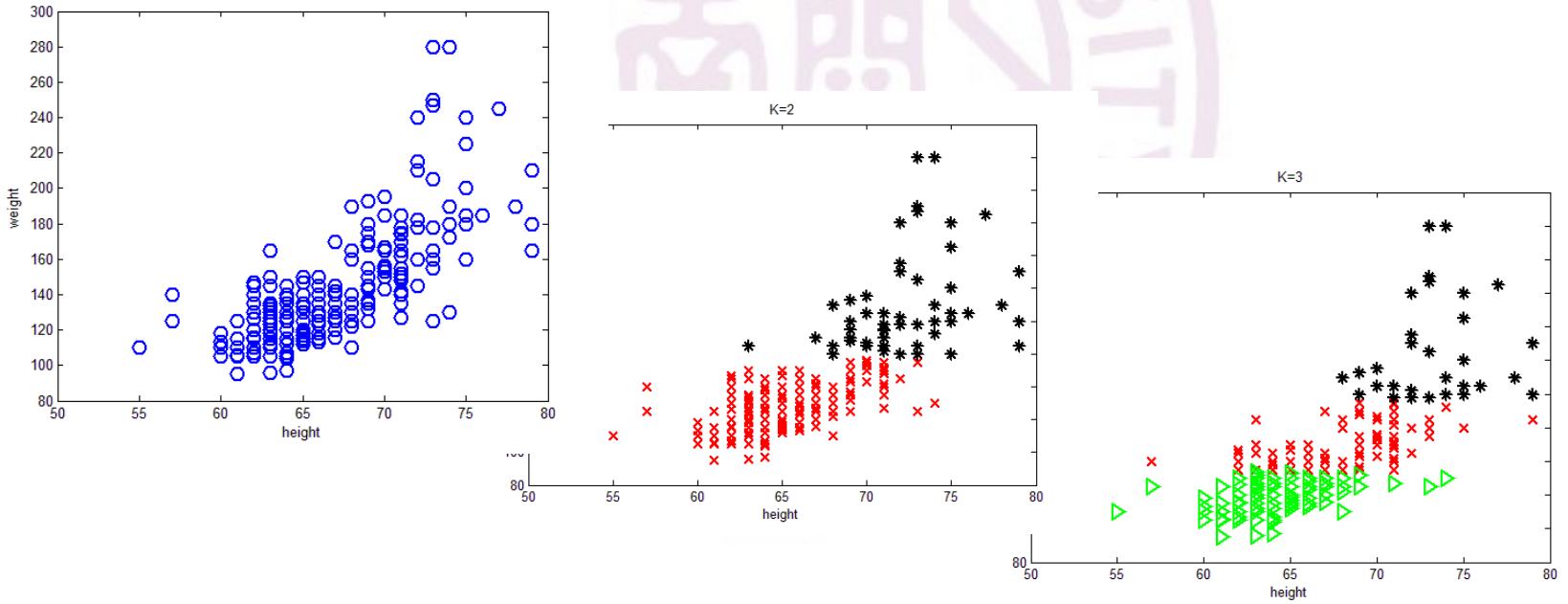
- 已知：训练样本集合 $\mathcal{D}$ $\left.\left.y_i\right)\right\}_{i=1}^{N}$，其中：
  - N：训练样本数目
  - $\mathbf{x}_i$：特征或属性（输入变量）
  - y_i：响应变量（输出变量）
- 求解：从输入 $\mathbf{x}$ 到输出 y 之间的映射函数 $y=\hat{f}(\mathbf{x})$
  - 条件：满足最小化代价函数
- 问题类型
  - 当 y 取离散数值时，一般称为分类(classification)问题；
  - 当 y 取连续数值时，一般称为回归(regression)问题。
- 分类问题
  - 图像分类
  - 语音识别
  - 文本分类

监督学习

- 回归问题
  - 股票股价预测
  - 机械手位置预测
  - 建筑物内温度预测

无监督学习

- 已知：训练样本集合 $\mathcal{D}$
- 求解：概率密度估计 $p(\mathbf{x};\theta)$
  - 监督学习为 $p(y|\mathbf{x};\theta)$
- 无监督学习的目标是发现数据中的“有趣结构”，是人类学习的一般形式。



无监督学习

- 应用问题
  - Autoclass系统根据聚类天体物理量发现一种新类型的恒星；
  - 在电子商务领域，根据购物行为将用户分组，然后给不同组用户发送有针对性的定制广告；
  - 在生物学领域中，将流式细胞数据分组，以发现细胞的不同子群。

增强学习

- 已知：训练样本集合 $\mathcal{D}$ $\left.\left.z_i(y)\right)\right\}_{i=1}^{N}$
  - N：训练样本数目
  - $\mathbf{x}_i$：特征或属性（输入变量）
  - z_i：奖励或惩罚（二值变量），满足

$$z_i(y)=\begin{cases}1 & \text{if } \hat{y}=y_i\\ 0 & \text{otherwise}\end{cases}$$

- 求解：从输入 $\mathbf{x}$ 到输出 y 的映射函数 $y=\hat{f}(\mathbf{x})$
- 应用问题
  - 棋类对弈学习
- 关键问题
  - 信用分配(credit assignment)

元学习

- Meta learning or learning to learn (from Wikipedia.org)
  - It is a subfield of Machine Learning where automatic learning algorithm are applied on meta-data about machine learning experiments. Although different researchers hold different views as to what the term exactly means, the main goal is to use such meta-data to understand how automatic learning can become flexible in solving different kinds of learning problems, hence to improve the performance of existing learning algorithms.

机器学习范式

- 有监督学习(Supervised Learning)

  Given $= \{\mathbf{X}_i, \mathbf{Y}_i\}$, learn $f(\cdot): \mathbf{Y}_i = f(\mathbf{X}_i)$, s.t. $^{\text{new}} = \{\mathbf{X}_j\} \Rightarrow \{\mathbf{Y}_j\}$

- 无监督学习(Unsupervised Learning)

  Given $= \{\mathbf{X}_i\}$, learn $f(\cdot): \mathbf{Y}_i = f(\mathbf{X}_i)$, s.t. $^{\text{new}} = \{\mathbf{X}_j\} \Rightarrow \{\mathbf{Y}_j\}$

- 增强学习(Reinforcement Learning)

  Given $\mathcal{D}$ actions, rewards, simulator/trace/real game$\}$

  learn $\begin{matrix}\text{policy: } e, r \to a \\ \text{utility: } a, e \to r\end{matrix}$, s.t. $\{\text{env, new real name}\} \Rightarrow a_1, a_2, \ldots$

- 主动学习(Active Learning)

  Given $\mathcal{D} \sim$ $\sim$

  s.t. $\mathcal{D}$ $(\cdot)$, policy, $\{\mathbf{Y}_j\}$

几个问题

解决机器学习问题的关键步骤

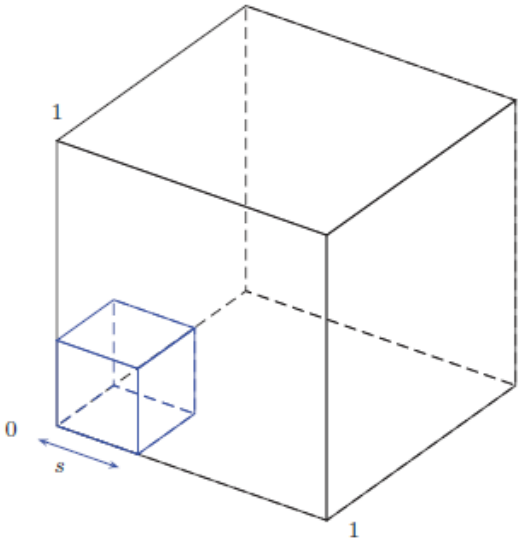
- 数据：方法所依赖的“过去经验”是什么
- 前提：需要对问题做出的前提假设是什么
- 表示：如何表示要解决的问题
- 计算：如何从数据中获取知识
- 评估：获取知识的效果如何
- 模型选择：问题是否可以解决得更好

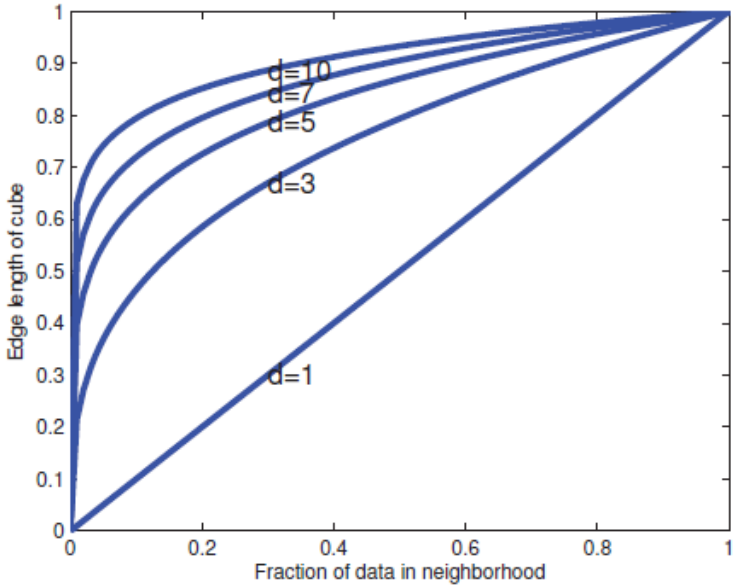
维数灾难

- 问题
  - 为了准确地估计样本的概率密度，一般通过增大估计点邻域尺度来使得样本点数目达到期望的比例 f；
  - 假设邻域为 D 维空间超立方体，则超立方体的边长期望为 $e_D(f) = f^{1/D}$，如：

$$e_{10}(0.1) = 0.1^{1/10} = 0.79, e_{10}(0.01) = 0.63$$

  - 但，边长期望的增加破坏了密度估计的“近邻”性。



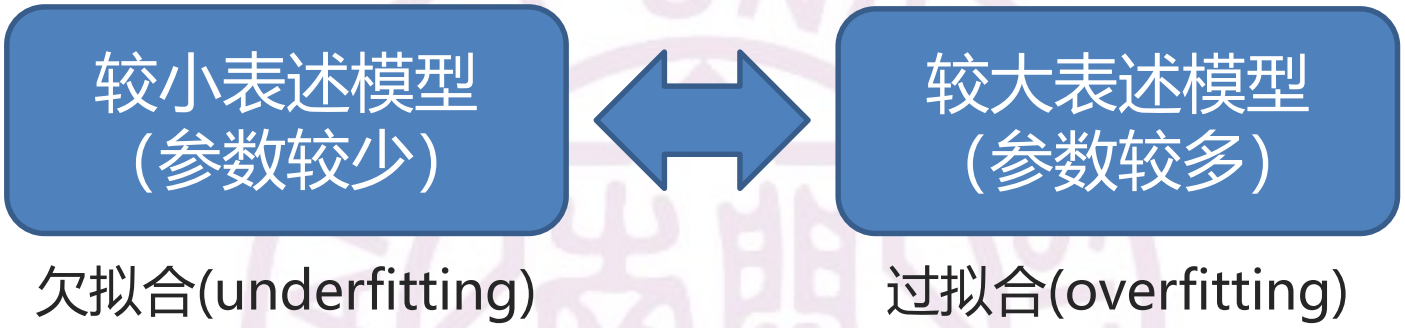


维数灾难

- 应对办法：对数据分布的性质做出某些假设
  - 这些假设称为归纳偏置(inductive bias)
  - 这些假设通常被嵌入到参数模型中

过度拟合

- 模型包含的参数越多就能够越好地拟合训练数据
  - 有限的训练数据使得模型过分地关注数据的独特性，而忽视共性。



模型选择

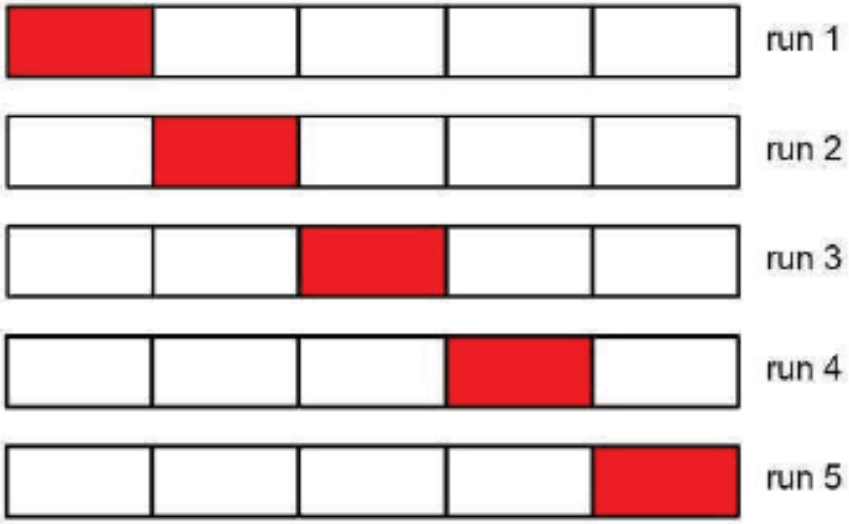
- 问题：从多个模型中选择合适的模型(right one)
- 方法：使用训练样本集合计算每个模型的错误率

$$\text{error}\left(f,\mathcal{D}\right)=\frac{1}{N}\sum_{i=1}^{N}\mathbb{I}\left(f\left(x_i\right)\neq y_i\right)$$

- 泛化误差(generalization error)：对未知数据的期望错误率
  - 通过大规模测试样本集合的性能评估来近似模型的泛化误差
  - 一般基于验证数据集来实现
- 随机拆分训练数据集：将训练样本集合分成两个部分(8:2)
  - 用于模型训练的部分称为训练集合(train set)
  - 用于模型评估的部分称为验证集合(validation set)

模型选择

- 交叉验证(cross validation, CV)



  - 极端情形是余一交叉验证(leave-one out cross validation)

没有免费午餐定理

- 没有免费午餐定理：许多机器学习研究关注设计不同的模型和不同的学习算法来拟合它们所要处理的数据，但是不存在通用的最佳模型。
- 原因：适合某个领域的假设集合未必适合其它领域的问题
- 结果：需要开发不同的模型和学习算法，以适应现实世界中种类繁多的数据。
- 对于每个模型也存在着许多不同的模型训练算法，即学习算法。这些学习算法往往需要在速度-准确性-复杂性之间做出权衡。

深度网络的瓶颈(Prof. Alan Yuille)

- 三大瓶颈：源自“组合爆炸”造成巨大信息量的概念
  - 需要大量标注数据——造成研究倾向数据资源丰富的领域，而不是重要的领域。而且，迁移学习、少样本学习、无监督学习和弱监督学习在性能上还无法与监督学习相比。
  - 过度拟合基准数据——造成泛化能力较差，在实际应用中出现偏差将带来非常严重的后果。如：训练自动驾驶系统的数据集中从来没有坐在路中间的婴儿。
  - 对图像变化过度敏感——模型敏感于对抗性攻击，可能会改变深度网络对一个物体的认知。
- 两条应对之道
  - 利用组合模型培养泛化能力（应用生成模型和因果模型）
  - 使用组合数据测试潜在故障（关注最坏情形而不是平均情形）
- 组合性(Compositionality)是指，一个复杂的表达，它的意义可以通过各个组成部分的意义来决定。

诚信 思路 实践