東莞理工學院
DONGGUAN UNIVERSITY OF TECHNOLOGY

人工智能概论

第二章：人工智能基础

丁烨，计算机科学与技术学院
dingye@dgut.edu.cn

目录

机器学习基础

基本概念

- 机器学习（Machine Learning）
- 人工智能的一个分支
- 实现人工智能的一个途径，即以机器学习为手段解决人工智能中的问题
- 机器学习在近 30 多年已发展为一门多领域交叉学科，涉及概率论、统计学、逼近论、凸分析、计算复杂性理论等多门学科
- 机器学习理论主要是设计和分析一些让计算机可以自动“学习”的算法
- 机器学习算法是一类从数据中自动分析获得规律，并利用规律对未知数据进行预测的算法

机器学习基础

基本概念

- 傍晚小街路面上沁出微雨后的湿润，和熙的细风吹来，抬头看看天边的晚霞，嗯，明天又是一个好天气。走到水果摊旁，挑了个根蒂蜷缩、敲起来声音浊响的青绿西瓜，一边满心期待着皮薄肉厚瓢甜的爽落感，一边愉快地想着，这学期狠下了工夫，基础概念弄得清清楚楚，算法作业也是信手拈来，这门课成绩一定差不了！

机器学习基础

基本概念

- 傍晚小街路面上沁出微雨后的湿润，和熙的细风吹来，抬头看看天边的晚霞，嗯，明天又是一个好天气。走到水果摊旁，挑了个根蒂蜷缩、敲起来声音浊响的青绿西瓜，一边满心期待着皮薄肉厚瓢甜的爽落感，一边愉快地想着，这学期狠下了工夫，基础概念弄得清清楚楚，算法作业也是信手拈来，这门课成绩一定差不了！

机器学习基础

基本概念

- 傍晚小街路面上沁出微雨后的湿润，和熙的细风吹来，抬头看看天边的晚霞，嗯，明天又是一个好天气。走到水果摊旁，挑了个根蒂蜷缩、敲起来声音浊响的青绿西瓜，一边满心期待着皮薄肉厚瓢甜的爽落感，一边愉快地想着，这学期狠下了工夫，基础概念弄得清清楚楚，算法作业也是信手拈来，这门课成绩一定差不了！

机器学习基础

基本概念

- 傍晚小街路面上沁出微雨后的湿润，和熙的细风吹来，抬头看看天边的晚霞，嗯，明天又是一个好天气。走到水果摊旁，挑了个根蒂蜷缩、敲起来声音浊响的青绿西瓜，一边满心期待着皮薄肉厚瓢甜的爽落感，一边愉快地想着，这学期狠下了工夫，基础概念弄得清清楚楚，算法作业也是信手拈来，这门课成绩一定差不了！

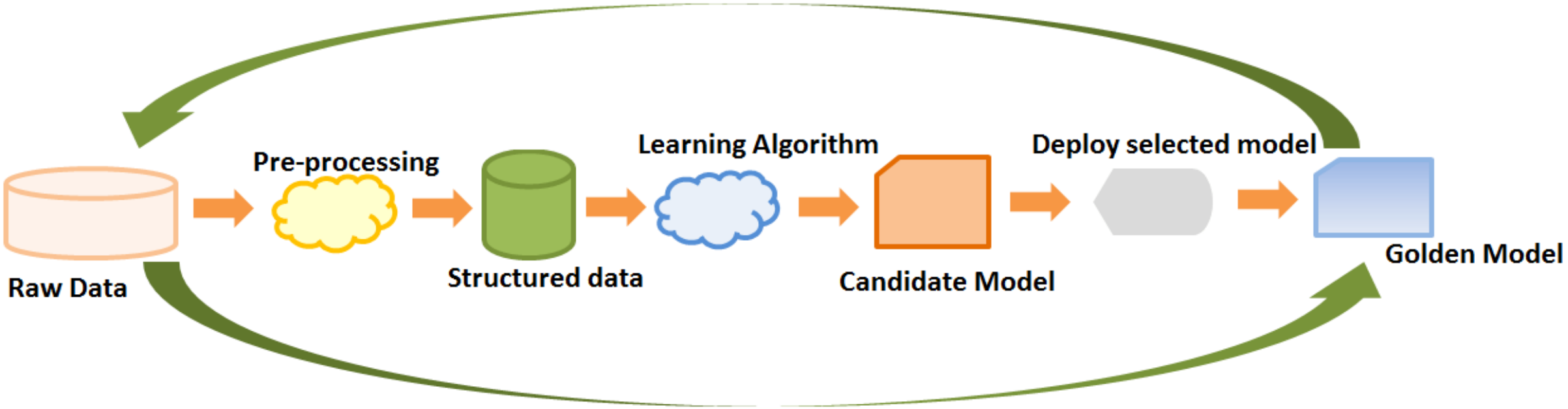
机器学习基础

基本概念

- 我们能做出有效的预判，是因为我们已经积累了许多经验
- 而通过对经验的利用，就能对新情况做出有效的决策
- 机器学习正是这样一门学科
- 它致力于研究如何通过计算的手段，利用经验来改善系统自身的性能

机器学习基础

流程


Raw Data
Pre-processing
Structured data
Learning Algorithm
Candidate Model
Deploy selected model
Golden Model

机器学习基础

流程

- 在计算机系统中，经验通常以「数据（Data）」形式存在
- 因此，机器学习一般会使用「算法（Algorithm）」
- 通过「学习（Learning）」的过程
- 从数据中产生「模型（Model）」
- 学习得到有效的模型之后，在面对新的情况时
- （例如看到一个没剖开的西瓜)
- 模型会给我们提供相应的判断
- （例如好瓜）

机器学习基础

流程

- 在计算机系统中，经验通常以「数据（Data）」形式存在
- 因此，机器学习一般会使用「算法（Algorithm）」
- 通过「学习（Learning）」的过程
- 从数据中产生「模型（Model）」
- 学习得到有效的模型之后，在面对新的情况时
- （例如看到一个没剖开的西瓜）
- 模型会给我们提供相应的判断
- （例如好瓜）

机器学习基础

基本术语

❖ 假定我们收集了一批关于西瓜的数据，例如：

1. 色泽：青绿，根蒂：蜷缩，敲声：清脆
2. 色泽：乌黑，根蒂：蜷缩，敲声：沉闷
3. 色泽：青绿，根蒂：硬挺，敲声：沉闷

❖ 这组记录的集合称为一个“数据集（Data Set）”

机器学习基础

基本术语

- 假定我们收集了一批关于西瓜的数据，例如：

1. 色泽：青绿，根蒂：蜷缩，敲声：清脆
2. 色泽：乌黑，根蒂：蜷缩，敲声：沉闷
3. 色泽：青绿，根蒂：硬挺，敲声：沉闷

- 其中每条记录是关于一个对象（“西瓜”）的描述
- 该记录被称为“实例（Instance）”或“样本（Sample）”

机器学习基础

基本术语

- 假定我们收集了一批关于西瓜的数据，例如：

1. 色泽：青绿，根蒂：蜷缩，敲声：清脆
2. 色泽：乌黑，根蒂：蜷缩，敲声：沉闷
3. 色泽：青绿，根蒂：硬挺，敲声：沉闷

- 每个记录都记载了对象的一些“属性（Attribute）”
- 或“特征（Feature）”
- 例如：色泽、根蒂、敲声

机器学习基础

基本术语

- 假定我们收集了一批关于西瓜的数据，例如：

1. 色泽：青绿，根蒂：蜷缩，敲声：清脆
2. 色泽：乌黑，根蒂：蜷缩，敲声：沉闷
3. 色泽：青绿，根蒂：硬挺，敲声：沉闷

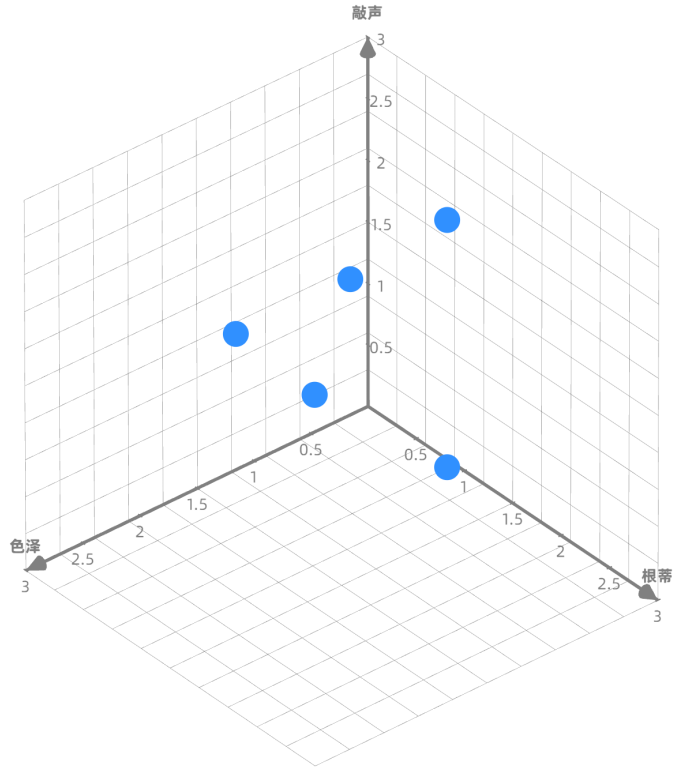
- 特征上的取值，例如“青绿”、“乌黑”等
- 被称为“特征值（Value）”

机器学习基础

基本术语

- 假定我们收集了一批关于西瓜的数据，例如：

1. 色泽：青绿，根蒂：蜷缩，敲声：清脆
2. 色泽：乌黑，根蒂：蜷缩，敲声：沉闷
3. 色泽：青绿，根蒂：硬挺，敲声：沉闷



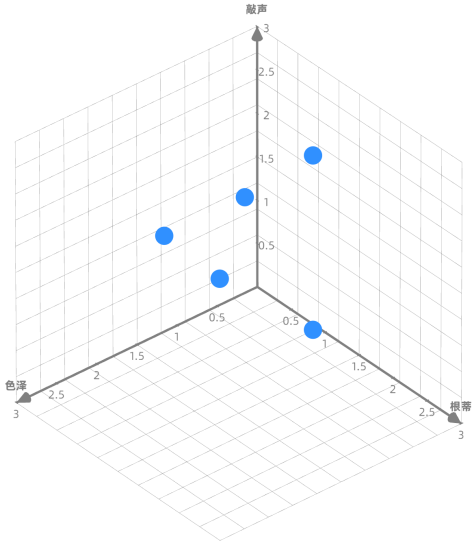
- 特征构建了一个多维空间
- 被称为“特征空间（Feature Space）”
- 每个特征代表着一个“维度（Dimension）”

机器学习基础

基本术语

- 假定我们收集了一批关于西瓜的数据，例如：

1. 色泽：青绿，根蒂：蜷缩，敲声：清脆
2. 色泽：乌黑，根蒂：蜷缩，敲声：沉闷
3. 色泽：青绿，根蒂：硬挺，敲声：沉闷



- 每个西瓜都可在这个特征空间中找到自己的坐标位置
- 由于空间中的每个点对应一个坐标向量
- 特征值构成了一个“特征向量（Feature Vector）”

机器学习基础

流程

- 在计算机系统中，经验通常以「数据（Data）」形式存在
- 因此，机器学习一般会使用「算法（Algorithm）」
- 通过「学习（Learning）」的过程
- 从数据中产生「模型（Model）」
- 学习得到有效的模型之后，在面对新的情况时
- （例如看到一个没剖开的西瓜）
- 模型会给我们提供相应的判断
- （例如好瓜）

机器学习基础

基本术语

- 使用“算法（Algorithm）”从数据中学得模型的过程被称为：
- “学习（Learning）”或“训练（Training）”
- 训练过程中使用的数据被称为“训练集（Training Set）”
- 学得的“模型（Model）”归纳了关于数据的某种潜在规律
- 这种潜在规律自身，被称为“真相（Ground Truth）”
- 训练的目的就是为了找出或逼近真相

机器学习基础

基本术语

- 假定我们收集了一批关于西瓜的数据，例如：

1. 色泽：青绿，根蒂：蜷缩，敲声：清脆
2. 色泽：乌黑，根蒂：蜷缩，敲声：沉闷
3. 色泽：青绿，根蒂：硬挺，敲声：沉闷

- 如果希望学得一个能帮助我们判断没剖开的瓜是不是“好瓜”的模型
- 仅有前面的示例数据显然是不够的

机器学习基础

基本术语

- 我们需要获得训练样本的结果信息，例如：
- 色泽：青绿，根蒂：蜷缩，敲声：清脆；好瓜
- 这里关于示例结果的信息，称为“标签（Label）”

- 不同于特征数据可以通过观察获得
- 标记数据通常需要“领域专家（Domain Expert）”的协助才能获取

机器学习基础

基本术语

- 若欲预测的标签是离散值，例如“好瓜”、“坏瓜”
- 此类学习任务称为“分类（Classification）”

- 若欲预测的标签是连续值，例如：西瓜成熟度 0.95 、0.37 等
- 此类学习任务称为“回归（Regression）”

机器学习基础

基本术语

- 除此之外，我们还可以对西瓜做“聚类（Clustering）”
- 将训练集中的西瓜分成若干“簇（Cluster）”
- 这些自动形成的簇可能对应一些潜在的概念划分
- 例如“浅色瓜”、“深色瓜”等
- 在聚类任务中，我们并不事先知道聚类标签是什么
- 聚类有助于我们了解数据内在的规律，为更深入地分析数据建立基础

机器学习基础

基本术语

- 根据训练数据是否拥有标签信息，学习任务可大致划分为两大类
- “监督学习（Supervised Learning）”
- “无监督学习（Unsupervised Learning）”

- 分类和回归是监督学习的代表
- 聚类则是无监督学习的代表

机器学习基础

流程

- 在计算机系统中，经验通常以「数据（Data）」形式存在
- 因此，机器学习一般会使用「算法（Algorithm）」
- 通过「学习（Learning）」的过程
- 从数据中产生「模型（Model）」
- 学习得到有效的模型之后，在面对新的情况时
- （例如看到一个没剖开的西瓜)
- 模型会给我们提供相应的判断
- （例如好瓜）

机器学习基础

基本术语

- 模型训练后，可以对“测试集（Testing Set）”的数据进行分析
- 分析过程被称为“预测（Prediction）”

机器学习基础

基本术语

- 机器学习的目标是使学得的模型能很好地适用于“新样本”
- 而不是仅仅在训练样本上工作得很好
- 学得模型适用于新样本的能力，称为“泛化（Generalization）”能力
- 具有强泛化能力的模型能很好地适用于整个样本空间
- 一般而言，训练样本越多，我们得到的信息也越多
- 这样就越有可能通过学习获得具有强泛化能力的模型

目录

模型评估与选择

归纳偏好

- 假定我们收集了一批关于西瓜的数据，例如：

1. 色泽：青绿，根蒂：蜷缩，敲声：清脆；好瓜
2. 色泽：乌黑，根蒂：蜷缩，敲声：清脆；好瓜
3. 色泽：青绿，根蒂：硬挺，敲声：沉闷；坏瓜
4. 色泽：乌黑，根蒂：硬挺，敲声：沉闷；坏瓜

- 你可以很容易的通过“相关性分析”得到一个模型

模型评估与选择

归纳偏好

❖ 假定我们收集了一批关于西瓜的数据，例如：

1. 色泽：青绿，根蒂：蜷缩，敲声：清脆；好瓜
2. 色泽：乌黑，根蒂：蜷缩，敲声：清脆；好瓜
3. 色泽：青绿，根蒂：硬挺，敲声：沉闷；坏瓜
4. 色泽：乌黑，根蒂：硬挺，敲声：沉闷；坏瓜

❖ 你可以很容易的通过“相关性分析”得到一个模型

模型评估与选择

归纳偏好

❖ 假定我们收集了一批关于西瓜的数据，例如：

1. 色泽：青绿，根蒂：蜷缩，敲声：清脆；好瓜
2. 色泽：乌黑，根蒂：蜷缩，敲声：清脆；好瓜
3. 色泽：青绿，根蒂：硬挺，敲声：沉闷；坏瓜
4. 色泽：乌黑，根蒂：硬挺，敲声：沉闷；坏瓜

❖ 你可以很容易的通过“相关性分析”得到一个模型

模型评估与选择

归纳偏好

❖ 假定我们收集了一批关于西瓜的数据，例如：

1. 色泽：青绿，根蒂：蜷缩，敲声：清脆；好瓜
2. 色泽：乌黑，根蒂：蜷缩，敲声：清脆；好瓜
3. 色泽：青绿，根蒂：硬挺，敲声：[illegible]；坏瓜
4. 色泽：乌黑，根蒂：硬挺，敲声：[illegible]；坏瓜

❖ 你可以很容易的通过“相关性分析”得到一个模型

模型评估与选择

归纳偏好

- 归纳偏好（Inductive Bias）
- 机器学习算法在学习过程中对某种类型假设的偏好
- 归纳偏好可看作学习算法自身在一个可能很庞大的假设空间中对假设进行选择的启发式“价值观”

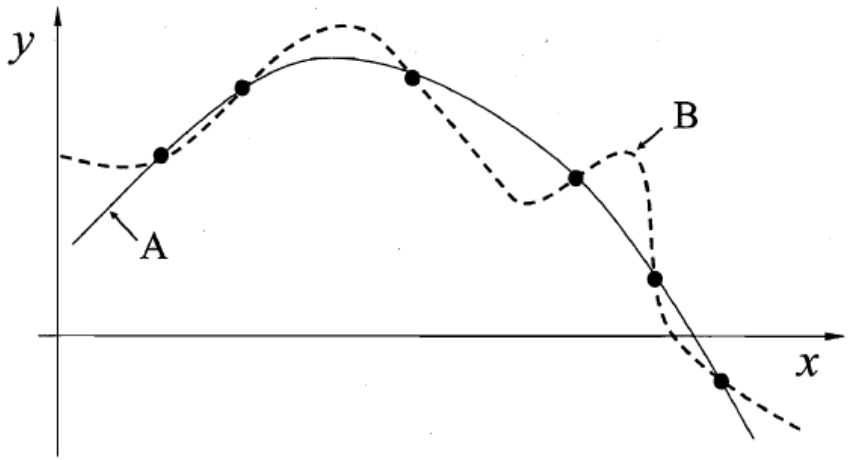


图 1.3　存在多条曲线与有限样本训练集一致

模型评估与选择

归纳偏好

- 奥卡姆剃刀（Occam’s Razor）原则
- 14 世纪修士奥卡姆的威廉（William of Occam）提出的逻辑学法则
- 如果关于同一个问题有许多种理论，每一种都能作出同样准确的预言，那么应该挑选其中使用假定最少的
- 尽管越复杂的方法通常能做出越好的预言，但是在不考虑预言能力（即结果大致相同）的情况下，假设越少越好
- 在所有符合实验数据的模型中，简单的模型优于复杂模型

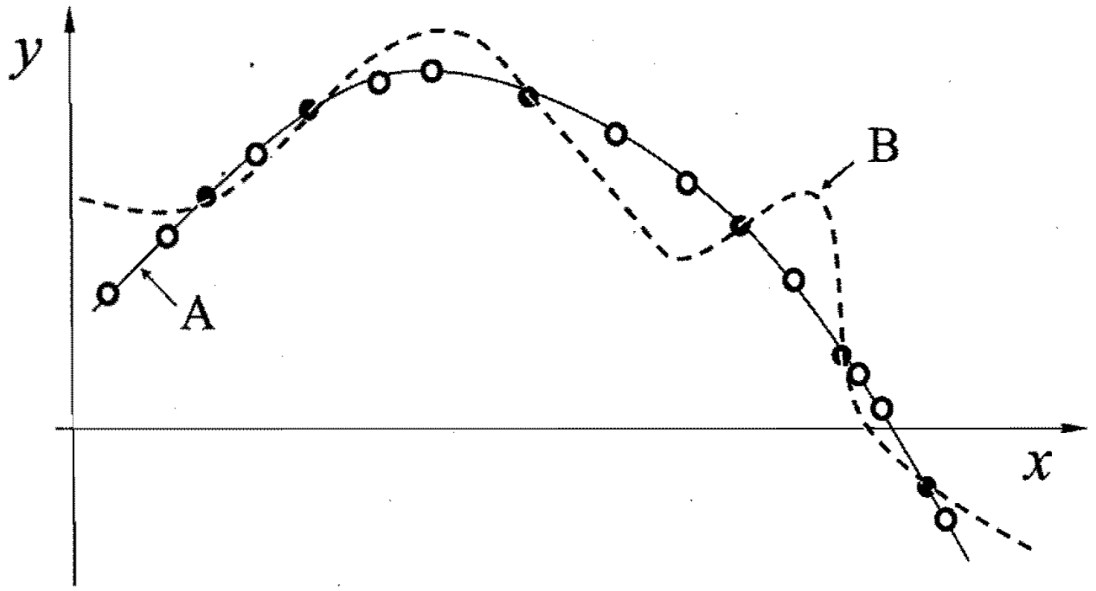
模型评估与选择

归纳偏好

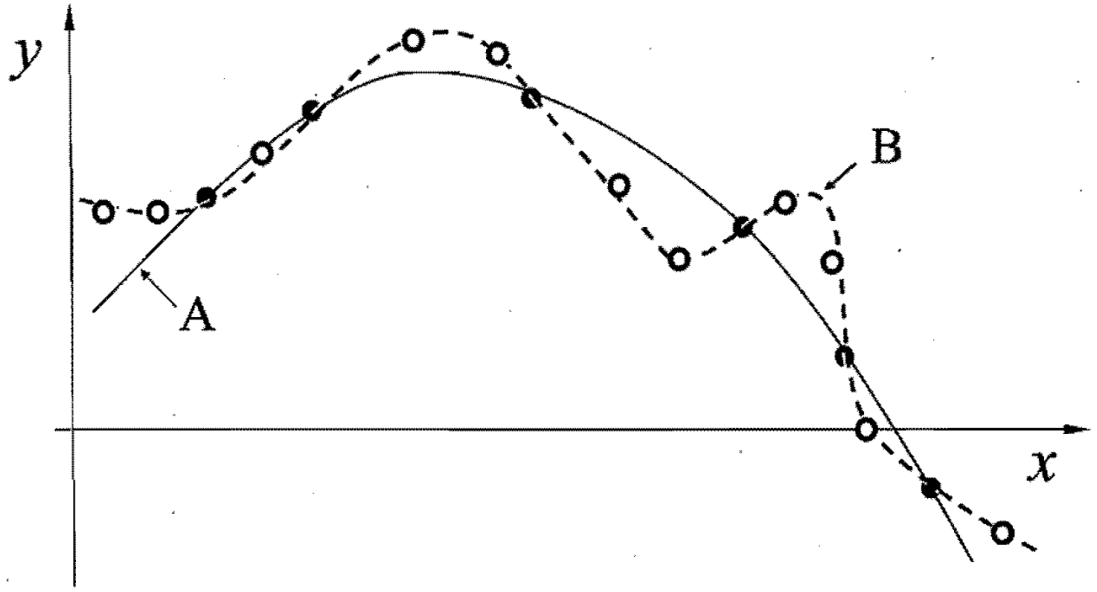
- 奥卡姆剃刀并非唯一可行的原则
- 结果“大致”相同是很难量化的
- 归纳偏好对应了算法本身所做出的关于“什么样的模型更好”的假设
- 这个假设是否成立，即：算法的归纳偏好是否与问题本身匹配
- 大多数时候直接决定了算法能否取得好的性能

模型评估与选择

归纳偏好



(a) A 优于 B



(b) B 优于 A

图 1.4 没有免费的午餐. (黑点: 训练样本; 白点: 测试样本)

模型评估与选择

经验误差与过拟合

- 因此，我们需要一个合理的评估“什么样的模型更好”的方法
- 对于分类问题来说，通常我们把分类错误的样本数占样本总数的比例称为“错误率（Error Rate）”
- 即如果在 m 个样本中有 a 个样本分类错误，则错误率 $E = \frac{a}{m}$
- 相应的，$1 - a/m$ 被称为“精度（Accuracy）”

模型评估与选择

经验误差与过拟合

- 那么，在机器学习的过程中，错误率是不是越小越好呢？
- 对应奥卡姆剃刀原则：结果大致相同的情况下，模型太复杂了
- 如果模型将训练样本学得太好，很可能会导致泛化能力下降
- 这种现象在机器学习中称为“过拟合（Overfit）”

模型评估与选择

经验误差与过拟合

模型评估与选择

经验误差与过拟合

- 模型选择的几个关键问题：
- 如何获得测试结果？评估方法
- 如何评估性能优劣？性能度量
- 如何判断实质差别？比较检验

模型评估与选择

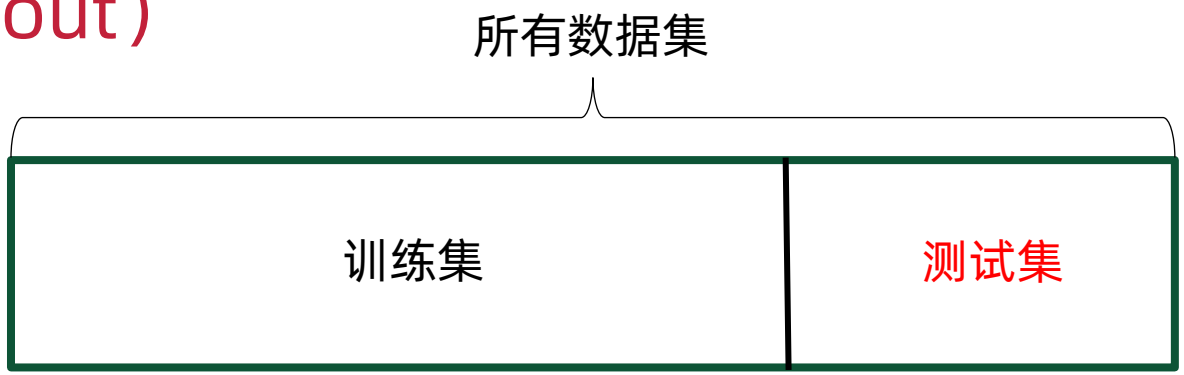
评估方法

- 使用“测试集（Testing Set）”来测试模型对新样本的判别能力
- 以测试集上的“测试误差（Testing Error）”作为泛化误差的近似值
- 在对模型的泛化误差进行评估时，测试集应该尽可能与训练集互斥
- 常见方法：
- 留出法（Hold-out）
- 交叉验证法（Cross Validation）
- 自助法（Bootstrap）

模型评估与选择

评估方法

- 留出法（Hold-out）

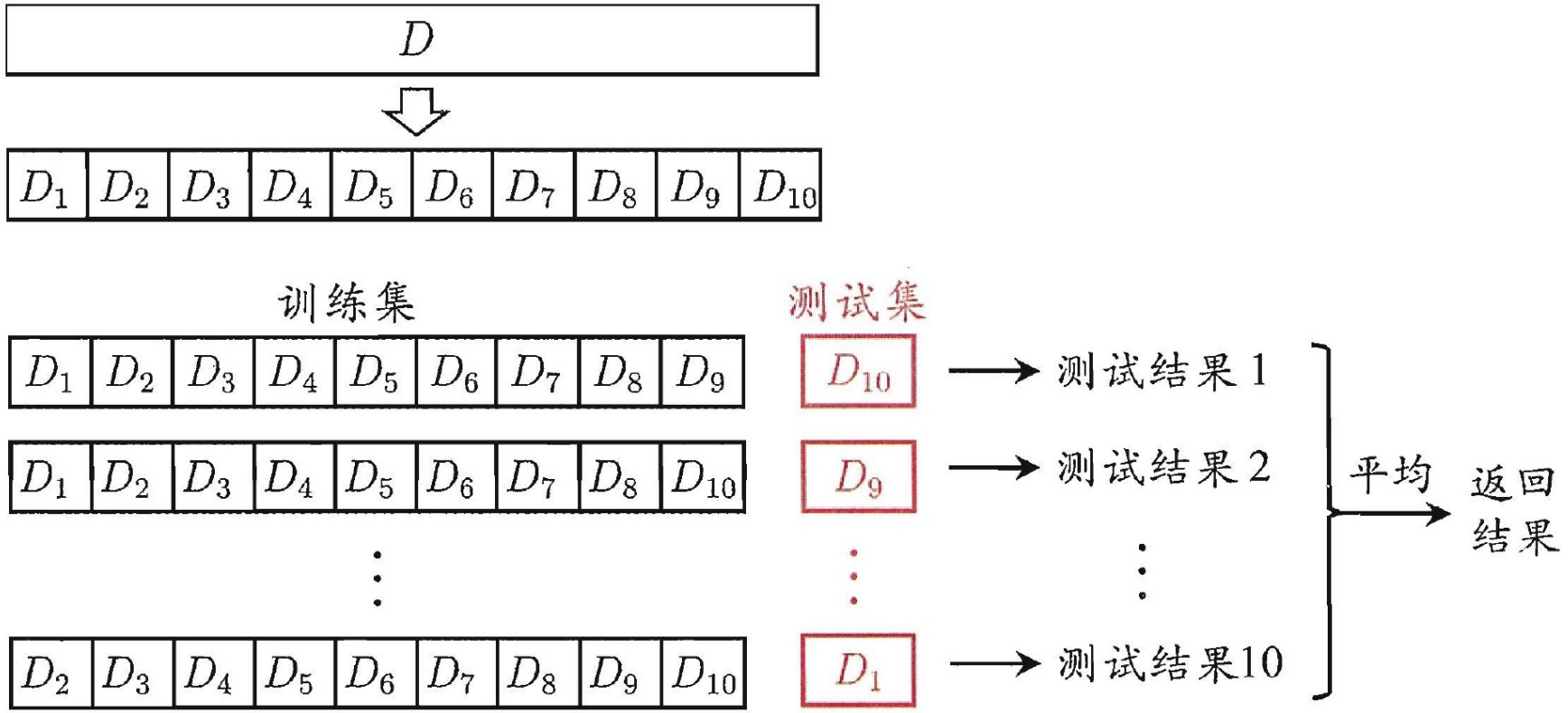


- 保持数据分布一致性 （例如: 分层采样)
- 单次使用留出法得到的估计结果不够稳定可靠，所以一般采取多次重复划分去平均值作为留出法的评估结果（例如：100 次随机划分）
- 测试集不能太大、不能太小（常见：1/5 至 1/3）

模型评估与选择

评估方法

- 交叉验证法（Cross Validation）



模型评估与选择

评估方法

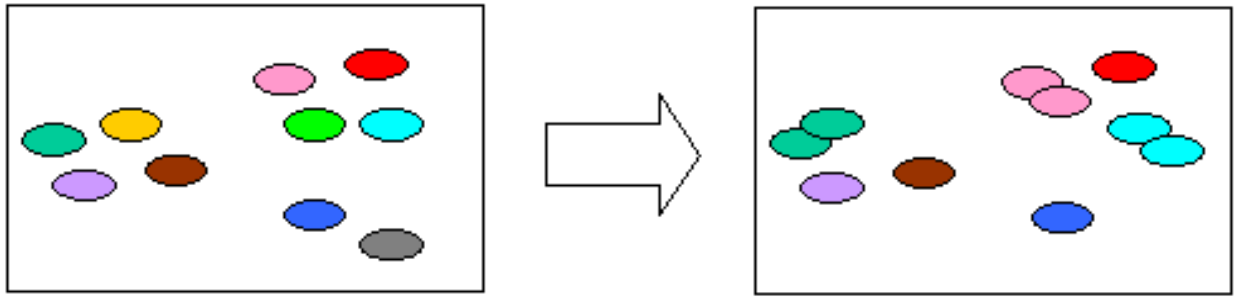
- p 次 k 折交叉验证：
- 数据集划分为 k 个子集，使用不同的划分重复 p 次
- 最终的评估结果为这 p 次交叉验证结果的均值

- 留一法（Leave One Out）：
- 若数据集包含 m 个样本，当 k = m 时，即为留一法
- 留一法的评估往往被认为是比较准确的
- 但在数据量大时开销往往很大

模型评估与选择

评估方法

- 自助法（Bootstrap）



- 每次进行有放回的采样，使得数据集 A 中部分样本会在数据集 B 中多次出现，有些样本则不会出现
- 自助法在数据集较小、难以有效划分训练集和测试集时很有用
- 但产生的数据集改变了初始数据集的分布，会引入估计偏差

模型评估与选择

评估方法

- 大多数学习算法都有些参数需要设定，参数配置不同，学得模型的性能往往有显著差异
- 算法的参数：一般由人工设定，亦称“超参数”
- 模型的参数：一般由学习确定
- 先产生若干模型，然后基于某种评估方法进行参数选择
- 参数调得好不好可能对最终性能有关键影响

模型评估与选择

性能度量

- 性能度量（Performance Measure）
- 衡量模型泛化能力的评价标准
- 反映了任务需求
- 使用不同的性能度量往往会导致不同的评判结果
- 什么样的模型是“好”的，不仅取决于算法和数据，还取决于任务需求
- 回归任务常用：均方误差（Mean Squared Error）
- 分类任务常用：错误率及精度

模型评估与选择

性能度量

- 查准率（Precision）：$\mathrm{P}=\frac{TP}{TP+FP}$
- 查全率（Recall）：$\mathrm{P}=\frac{TP}{TP+FN}$

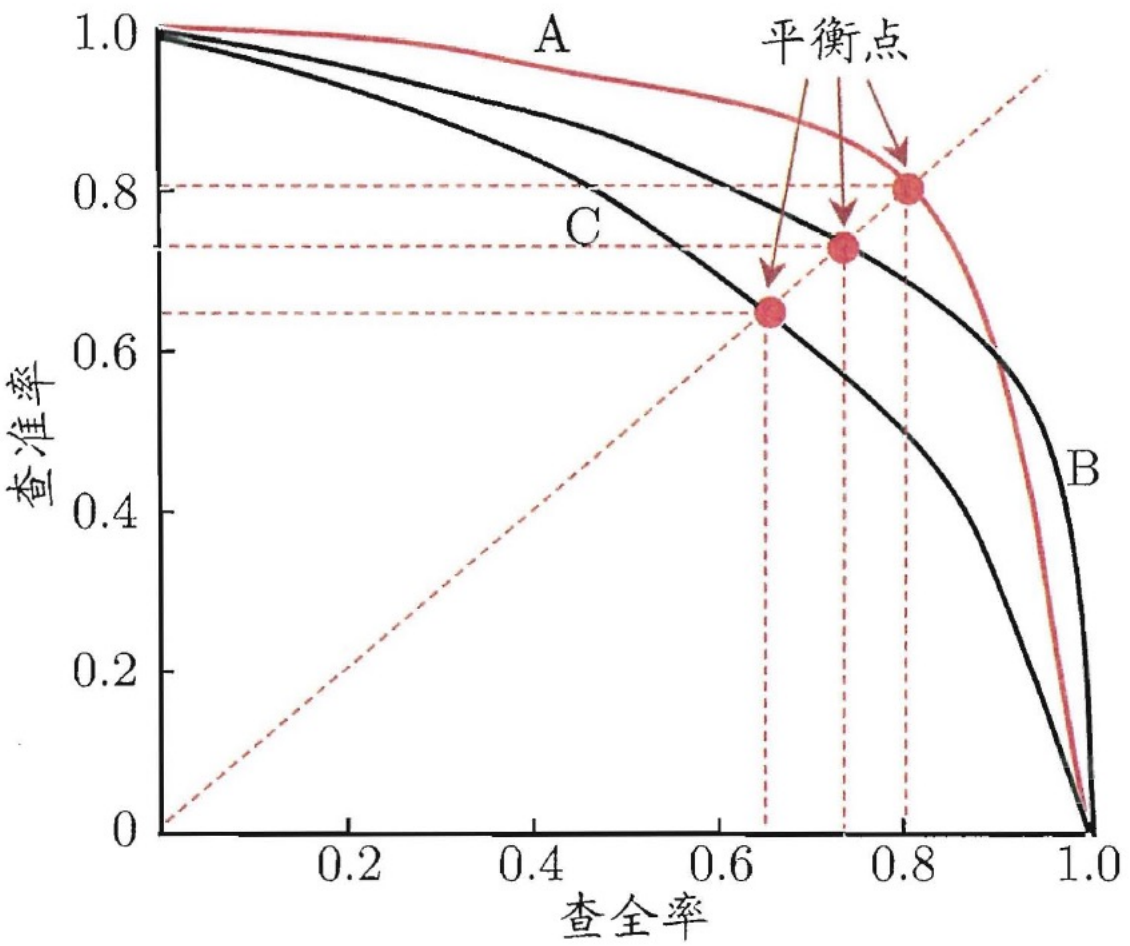
真实情况	预测结果	
	正例	反例
正例	TP（真正例）	FN（假反例）
反例	FP（假正例）	TN（真反例）

- 查准率和查全率为一对矛盾量
- 一般来说：
- 查准率高时，查全率往往偏低
- 查全率高时，查准率往往偏低

模型评估与选择

性能度量

- P-R 图直观地显示出模型在样本总体上的查全率、查准率
- 在进行比较时，若一个模型的 P-R 曲线被另一个模型的 P-R 曲线完全“包住”
- 则可断言后者的性能优于前者



模型评估与选择

性能度量

- F1 是比 P-R 更加常用的度量方式：

$$F1 = \frac{2 \times P \times R}{P + R} = \frac{2 \times TP}{\text{样例总数} + TP - TN}$$

- F_β 可以表示出对查重率或查准率的不同偏好：

$$F_\beta = \frac{(1 + \beta^2) \times P \times R}{(\beta^2 \times P) + R}$$

模型评估与选择

比较检验

- 在某种度量下取得评估结果后，是否可以直接比较以评判优劣？
- No！因为：
- 测试性能不等于泛化性能
- 测试性能随着测试集的变化而变化
- 很多机器学习算法本身有一定的随机性

模型评估与选择

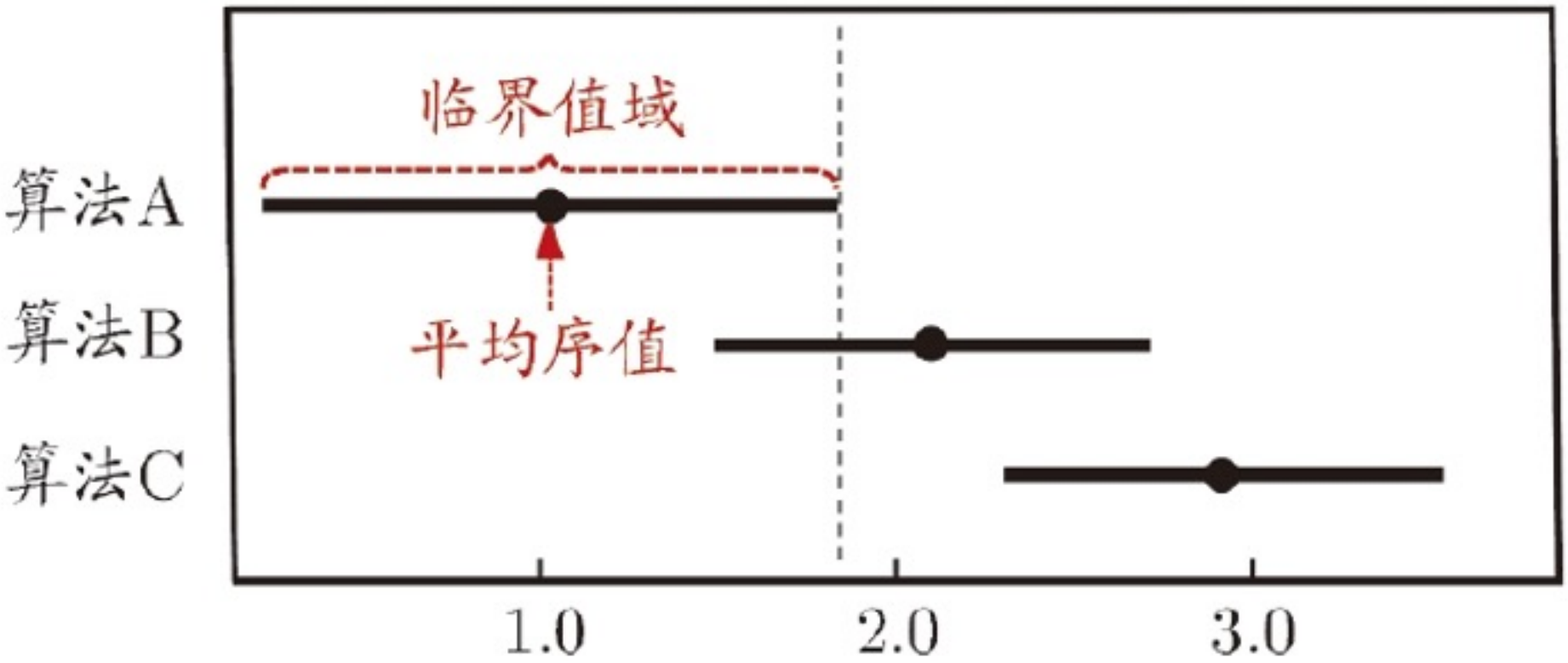
比较检验

- 统计假设检验（Hypothesis Test）
- 比较两个模型：
- 交叉验证 t 检验、McNemar 检验
- 比较多个模型：
- Friedman 检验
- Nemenyi 后续检验

模型评估与选择

比较检验

- Friedman 检验



模型评估与选择

总结

- 模型选择的几个关键问题：
- 如何获得测试结果？评估方法
- 如何评估性能优劣？性能度量
- 如何判断实质差别？比较检验

- 明白如何对模型进行评估和选择之后
- 我们才能有针对性地学习各种模型

总结

- 线性模型
- 决策树
- 神经网络
- 支持向量机
- 贝叶斯分类器
- 聚类与无监督学习
- 降维与度量学习
- 概率图模型
- 深度学习

总结

- 「西瓜书」周志华《机器学习》，清华大学出版社
- https://item.jd.com/12762673.html

- 「南瓜书」谢文睿、秦州《机器学习公式详解》，人民邮电出版社
- https://github.com/datawhalechina/pumpkin-book/

总结

- Scikit-Learn
- https://scikit-learn.org

- TensorFlow
- https://www.tensorflow.org

TensorFlow

東莞理工學院
DONGGUAN UNIVERSITY OF TECHNOLOGY

Thank You!

丁烨，计算机科学与技术学院
dingye@dgut.edu.cn