

经验/性能指标提升

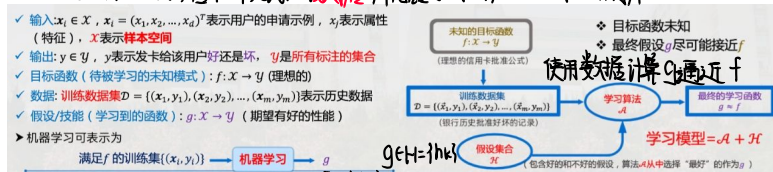
机器学习从大量数据中获得知识或技能的过程

② 不通过显著式编程来赋予计算机学习能力的领域
人为给出一些机械步骤/判断依据

③ 计算机被称为可学习, 指它能针对某个任务 T 和某个性能指标 P , 从经验 E 中学习, 并且它在 T 上被 P 所衡量的性能指标, 会随经验 E 增加而提高

使用机器学习必要条件:

- ① 有可被学习的潜在模式, 性能 P 能被提升
- ② 不容易被显著式编程
- ③ 存在关于潜在模式的数据, 机器学习有经验 E 可供输入



机器学习分类

- 监督学习: 分类, 离散标注; 回归, 连续标注
- 无监督学习: 常见的“聚类”问题
- 半监督学习: 部分数据有标签
- 强化学习

三盘棋: 1952年 Arthur Samuel 西洋跳棋; 1997年 国际象棋 AI 深蓝; 2016年 围棋 Alpha Go

- > 机器学习 (ML): 使用数据来计算假设 g 使它逼近目标函数 f
 - > 数据挖掘 (DM): 使用大量数据来寻找有趣的性质
 - > 模式识别 (PR): 从数据中自动识别出模式或规律
 - > 人工智能 (AI): 机器展现出来的智能
- 具体说明:
- 若有有趣的性质与逼近目标函数的假设相同, 则 ML=DM
 - 若有有趣的性质与逼近目标函数的假设相关, 则 ML 与 DM 可互补
 - ML 是从数据中自下而上归纳出假设 g
 - PR 通常从数据中分析出一个模型, 再用这个模型自上而下去演绎搜索相关模式
 - 逼近目标函数的假设显示出某种智能的行为, 则 ML 可帮助 AI 的实现

> 实际中难以区分 ML 与 DM

> 模式识别是问题, 机器学习是方法

> 机器学习是帮助实现人工智能的一条可能的路径