

- 1. 初识机器学习
- 2. 什么是机器学习
 - 2.1 监督学习
 - 2.2 无监督学习

1. 初识机器学习

- **Artificial Intelligence**: Human Intelligence Exhibited by Machines.
 - 人工智能 (AI)：机器展示的人类智能。
- **Machine Learning**: An Approach to Achieve Artificial Intelligence.
 - 机器学习 (ML)：一种实现人工智能的方法。
- **Deep Learning**: A Technique for Implementing Machine Learning.
 - 深度学习 (DL)：一种实现机器学习的技术。

2. 什么是机器学习

什么是机器学习，这个问题目前还没有明确、统一的回答。本课程中列举了两种定义：

- 在没有明确设置的情况下，使计算机具有学习能力的研究领域。—— **Arthur Samuel**

Machine Learning: The field of study that gives computers the ability to learn without being explicitly programmed.

- 计算机程序从经验 **E** 中学习，解决某一任务 **T**，进行某一性能度量 **P**，通过 **P** 测定在任务 **T** 上的表现因经验 **E** 而提高。—— **Tom Mitchell**

Well-posed Learning Problem: A computer program is said to learn from experience **E** with respect to some class of tasks **T** and performance measure **P**, if its performance at tasks in **T**, as measured by **P**, improves with experience **E**.

目前，机器学习主要的两类方式：

- **Supervised learning**: 监督学习
- **Unsupervised learning**: 无监督学习
- **Others**:
 - **Semi-supervised learning**: 半监督学习
 - **Reinforcement learning**: 强化学习

2.1 监督学习

监督学习：我们已经得到了一个数据集、并且已经知道正确的输出是什么样子（数据**含有标签**），而且预设输入和输出之间一定存在某种联系。

- **regression problem**: 回归问题，我们尝试预测出**连续的**输出，类似房价问题。
- **classification problem**: 分类问题，我们尝试预测出**离散的**输出，类似肿瘤问题。

2.2 无监督学习

无监督学习: 使用的数据集**没有任何标签**，在我们不知道输入 / 输出是什么的前提下，从数据中提取出结构 (structure) 。

- **clustering problem**: 聚类问题，自动将数据聚类。