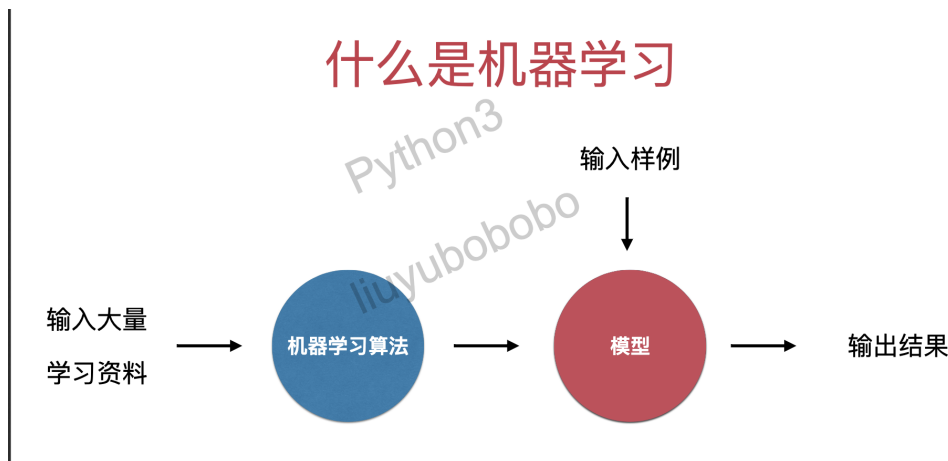


机器学习的基本概念

- 1. 什么是机器学习：让机器去【学习】



- 2. 机器学习的分类

- 方法的分类：

- 监督学习：给机器的训练数据拥有**标准答案**
 - 分类问题：输出是个离散的变量，包括**二分类问题**和**多分类问题**
 - 回归问题：输出是个连续的变量
 - 非监督学习：给机器的训练数据**没有任何标记或答案**
 - 聚类
 - 降维
 - 异常检测
 - 半监督学习：一部分数据有标准答案，另外一部分数据没有
 - 由非监督学习对数据做处理，之后使用监督学习手段做模型的训练和预测
 - 强化学习：根据周围环境的变换情况，采取行动，根据采取行动的结果，学习行动方式

- 其他分类

- 批量学习 vs 在线学习
 - 参数学习 vs 非参数学习

- 3. 机器学习的一些常见名词：

- 特征(feature)：自变量，在数学上是一个矩阵，**用大写X表示**
 - 标记(label)：因变量，监督学习里的标准答案，在数学上是一个列向量，**用小写y表示**
 - 特征向量：每一行数据
 - 特征空间(feature space)：由自变量组成的空间【**分类任务的本质就是在特征空间划分**】
 - 奥卡姆剃刀原则：简单的就是好的

- 没有免费的午餐定理：任意两个算法，他们的期望性能是相同的