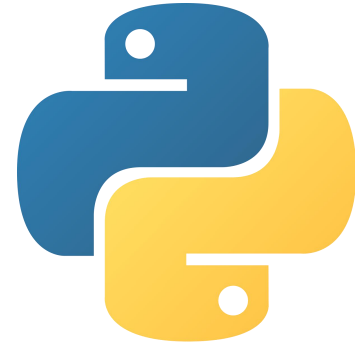


人工智能实践教程

从Python入门到机器学习



第二部分 机器学习 概述

- 所有代码及ppt均可以由以下链接下载
- <https://github.com/shao1chuan/pythonbook>
- <https://gitee.com/shao1chuan/pythonbook>

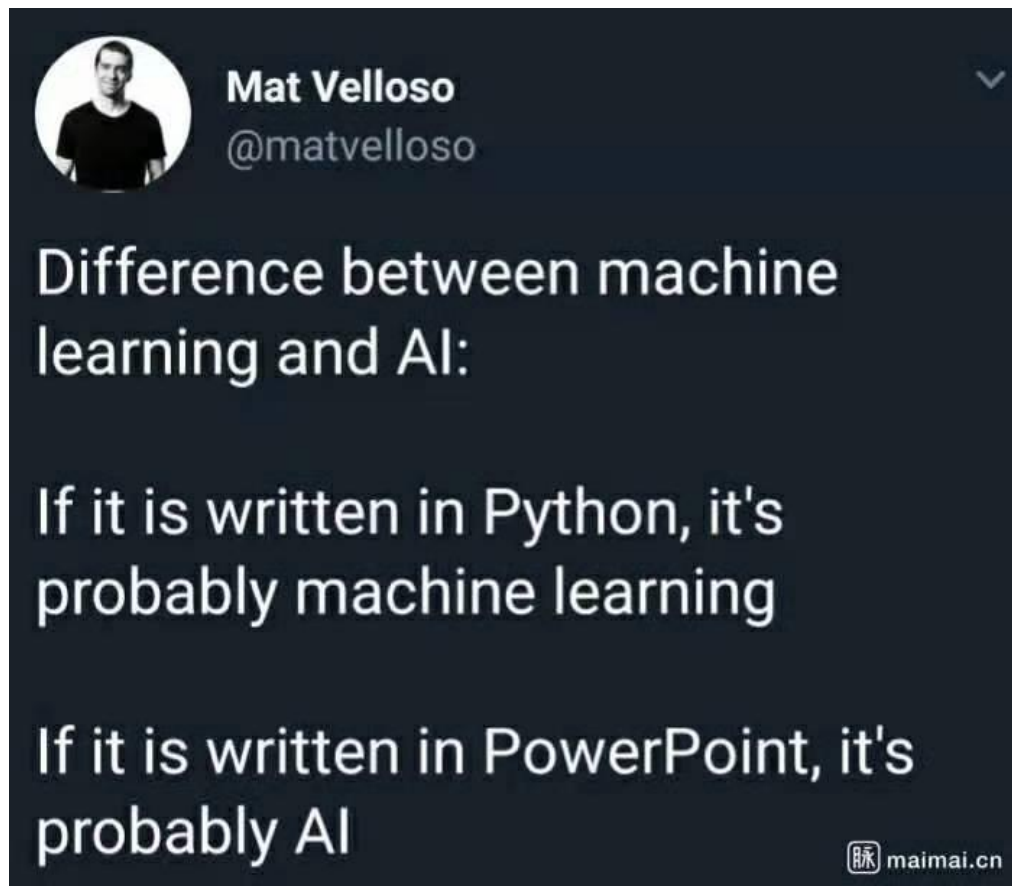


人工智能和机器学习

机器学习和数学

机器学习的应用

机器学习和人工智能的区别



什么是人工智能



Stephen Hawking says A.I. could be 'worst event in the history of our civilization'

- Physicist Stephen Hawking said the emergence of artificial intelligence could be the "worst event in the history of our civilization."
- He urged creators of AI to "employ best practice and effective management."
- Hawking is among a number of voices including Elon Musk who have warned about the dangers of AI.

Arjun Kharpal | @ArjunKharpal
Published 2:11 PM ET Mon, 6 Nov 2017 | Updated 5:59 PM ET Mon, 6 Nov 2017

CNBC



Facebook's head of AI wants us to stop using the Terminator to talk about AI

Yann LeCun chats about super-intelligent AI and the future of virtual assistants

By Sarah Frier | @sarahfrier | Oct 26, 2017, 12:36pm EDT

f t share



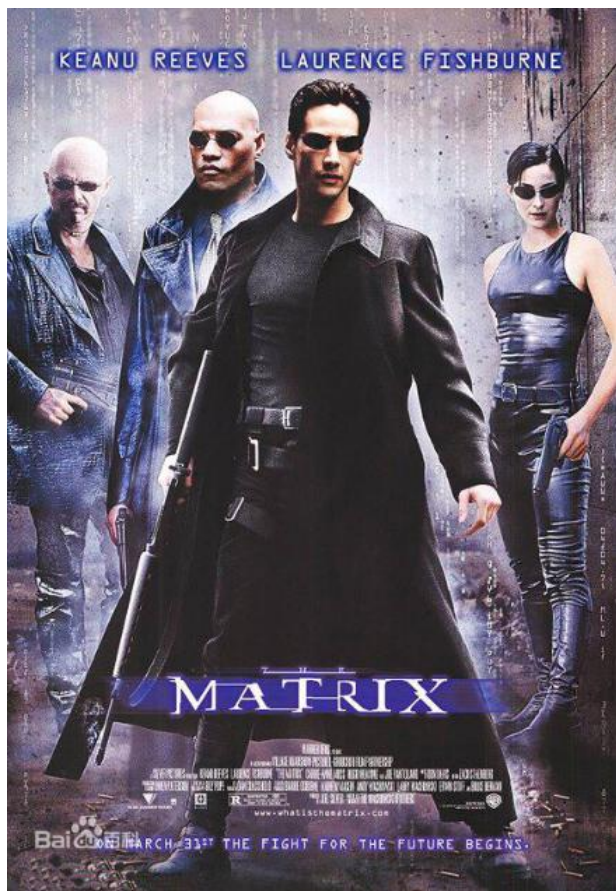
AGI VS AI

Artificial intelligence pioneer says we need to start over



Geoffrey Hinton harbors doubts about AI's current workhorse. (Johnny Quetta / University of Toronto)

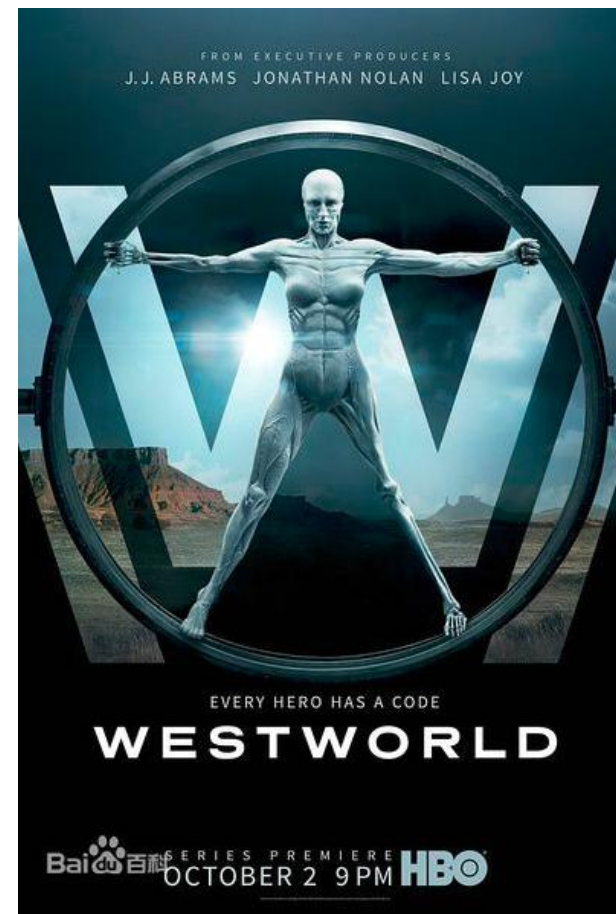
通用人工智能人工智能 (Artificial General Intelligent)



《骇客帝国》



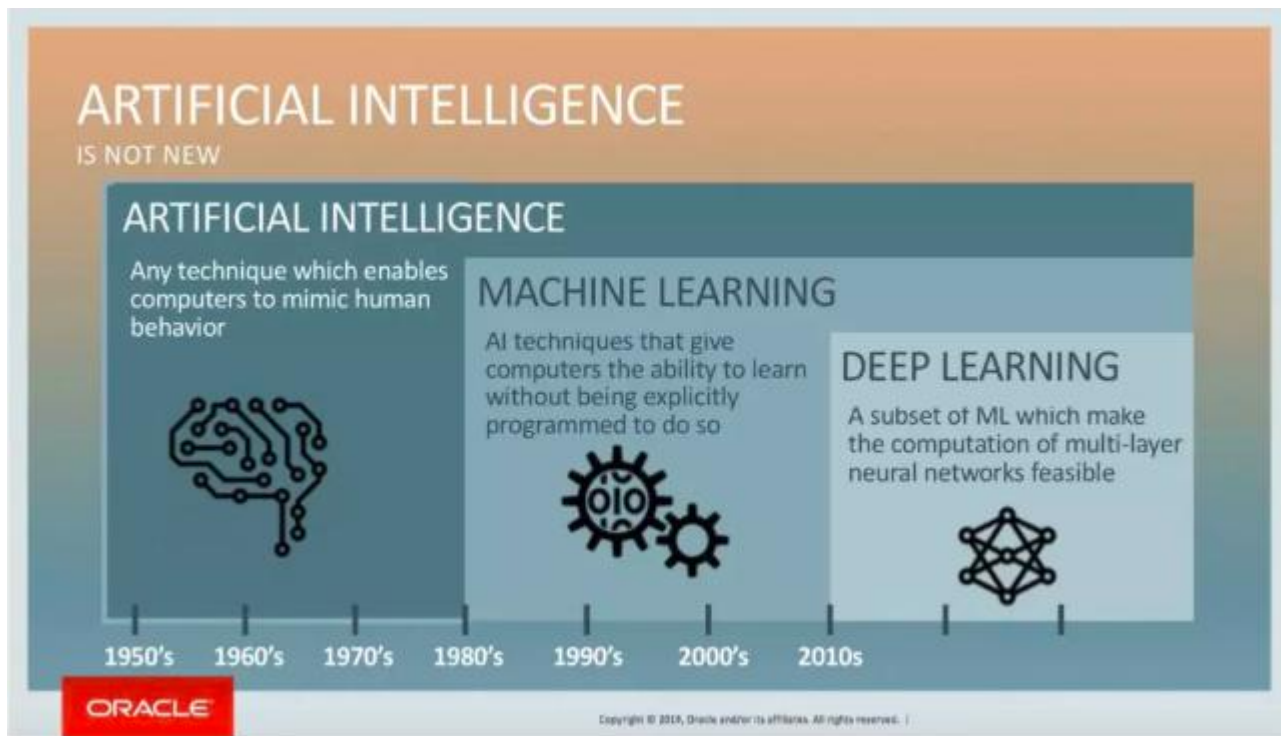
《超严骇客》



《西部世界》

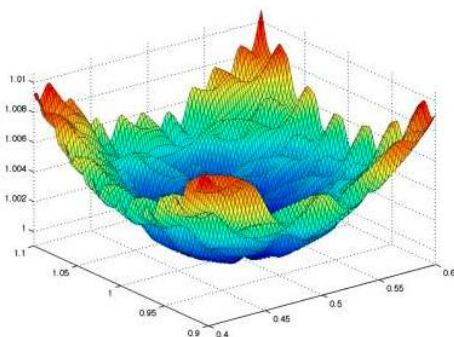
《黑镜》

人工智能技术

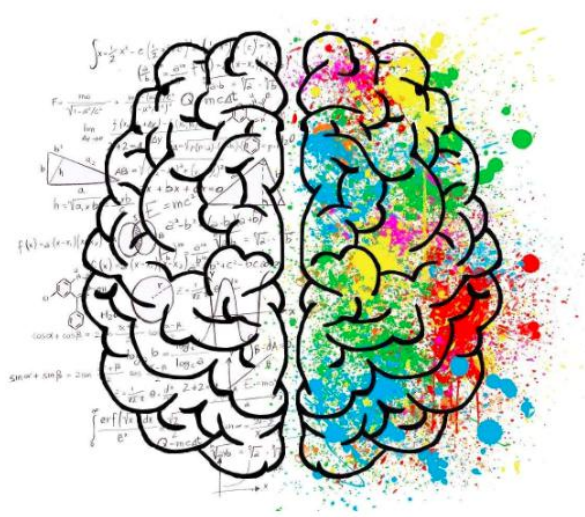


只要能**让计算机模拟人类行为**的技术，都叫**人工智能**

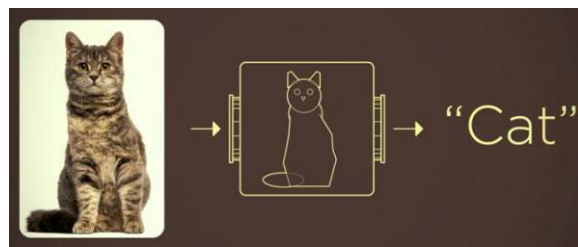
人工智能的分类领域



模式识别



语音识别



图像识别



自然语言处理

人工智能的分类领域

图像识别应用领域

自动驾驶 -> 深度神经网络/图像识别/模式识别 -> 特斯拉,
百度
陕西汽车控股集团有限公司 (本土)

图像识别/图像处理 -> 深度神经网络/LSTM -> 学科网
华为, Oppo (本土)
大疆 (本土)

工业自动化 -> 机器学习/图像识别 -> 三一重工, 长安大学 (本土)

光学影像 -> 深度学习/模式识别 -> 西工大博士 李学龙教授 (本土)

等等....

人工智能的分类领域

NLP应用领域

(大) 数据分析 -> 大数据 (Hadoop/Spark/MR, 深度学习/LSTM, NLP) -> 字节跳动 (互联网),
汇丰 (本土)
华为

文本分析 -> 模式识别/NLP/LSTM -> 爱彼迎, 大众点评
SAP (本土)

等等....

人工智能的分类领域

语音识别应用领域

个人助手 -> 深度学习 -> 苹果 (Siri) , 微软 (Cortana)

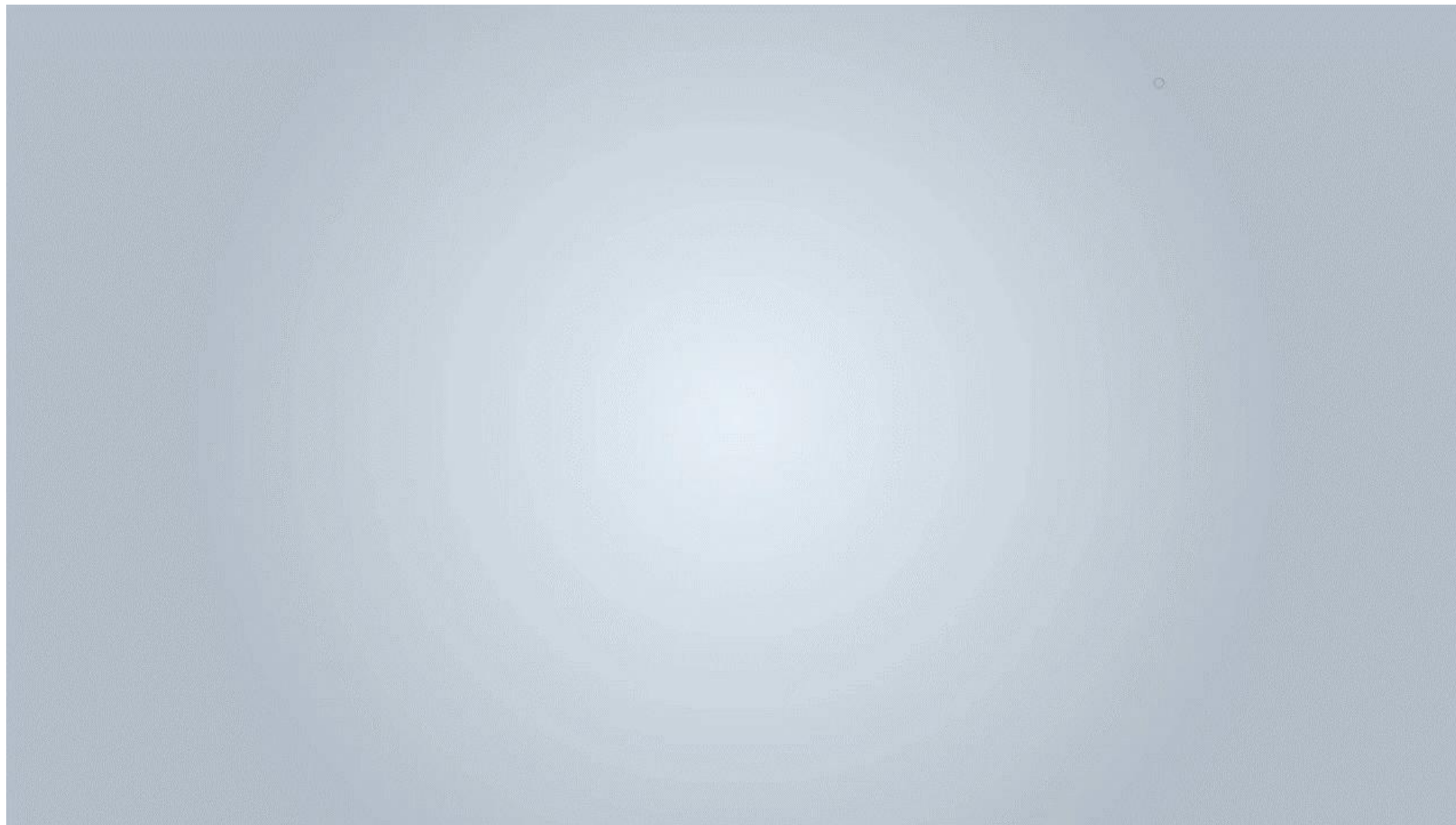
自动话务员 -> 深度学习/模式识别/LSTM -> 科大讯飞

声纹识别 -> 深度学习/LSTM -> SpeakIn

人工智能的分类领域



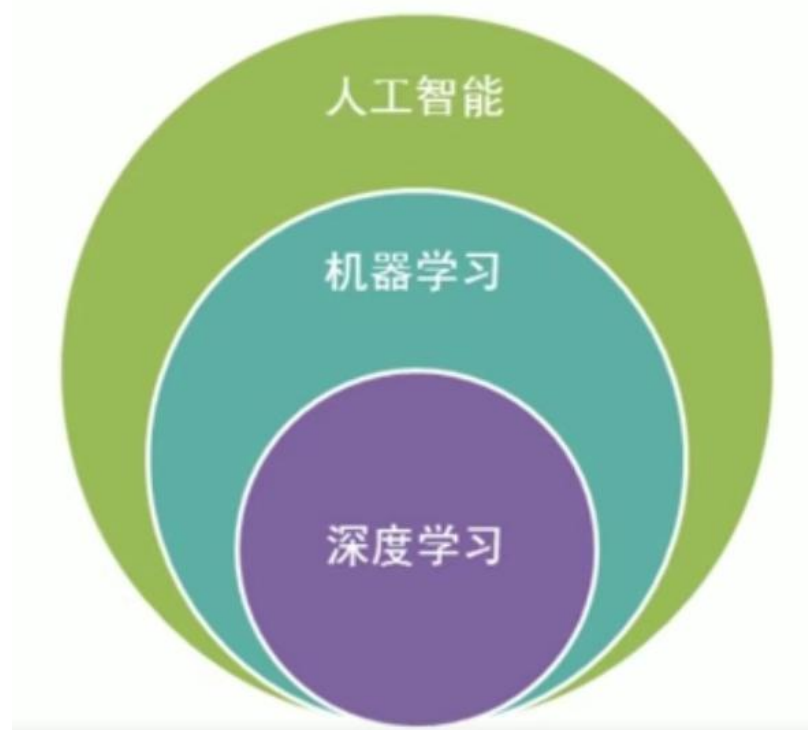
人工智能的分类领域



机器学习与人工智能

机器学习是现阶段人工智能的核心**技术**，是通过统计学原理，来对数据进行分析与建模

- 模式识别=机器学习
- 数据挖掘=机器学习+数据库
- 统计学习近似等于机器学习
- 计算机视觉=图像处理+机器学习
- 语音识别=语音处理+机器学习
- 自然语言处理=文本处理+机器学习



机器学习的基本方法 - 数据训练

逻辑推理和数据训练

机器学习的基本思想是通过大量数据的学习，找到一定的规律（这个规律就叫做模型），当有新数据出现时，把新数据带入模型

优点：

- (1) 速度快
- (2) 高度抽象化，简单易用

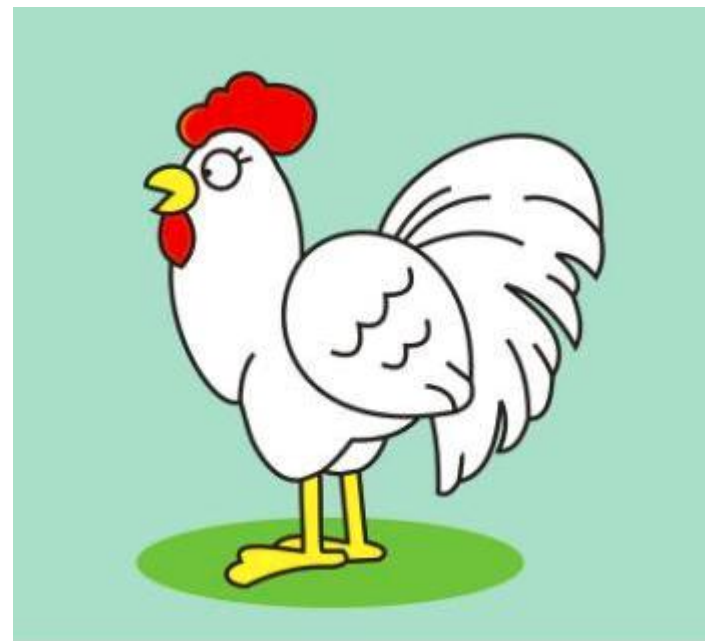
缺点：

- (1) 极度依赖数据
- (2) 模型的好坏至关重要

机器学习的第一个难点：不要日本菜



三体里的“射手”和“农场主”（罗素鸡）



现阶段的主流基于统计的人工智能算法，是通过大量的案例，靠对数据的特征进行归类，来实现识别效果。这个做法，就是罗素鸡。

如何克服这个问题？

答案：深度学习和强化学习！

深度学习和普通机器学习的区别：通过更高的复杂度，加强了让机器自我归纳总结的能力，并且通过规则的设计，让算法的扩展性更大

普通机器学习算法：一般就一个公式，一层推导，比较直接

深度学习算法：多个公式，多层推导，层层递进

强化学习和普通机器学习的区别：强化学习模拟人的认知过程，加入评价体系

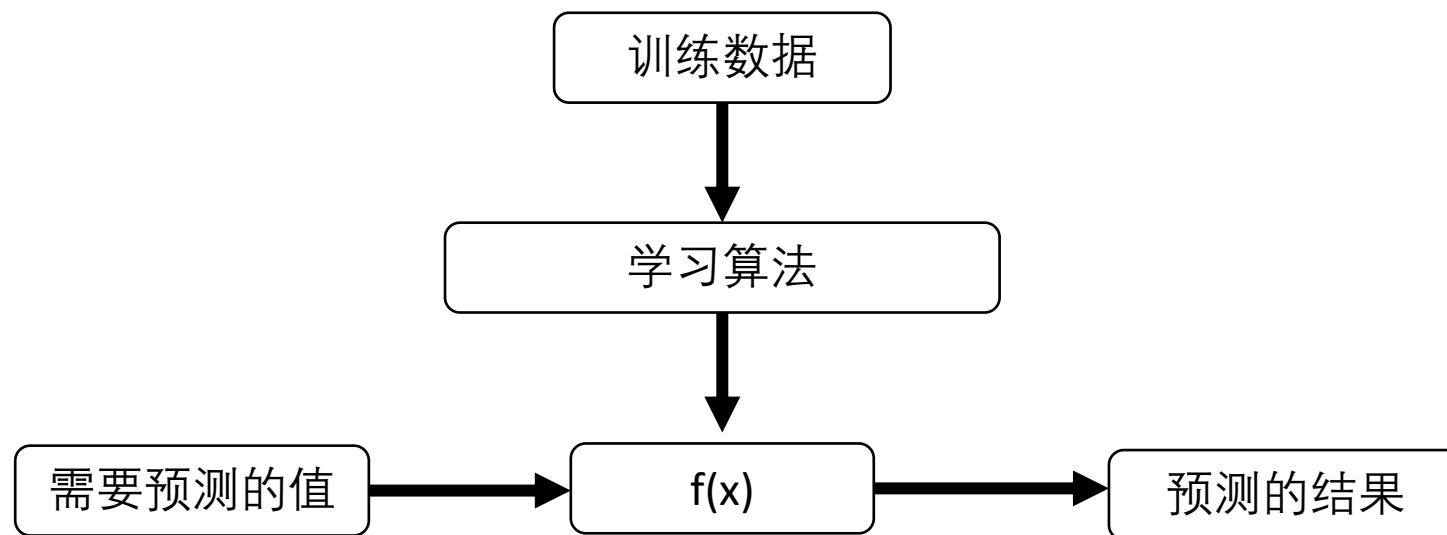
机器学习的第二个难点：四川火锅和日料

- A. “四川火锅比日料更好，因为它很辣”
- B. “四川火锅比日料更好，因为它不辣”

Q: “它”指的是谁？



机器学习的基本步骤



一个语音识别项目的流程：“填槽” (Slot Filling)



机器学习需要数学吗？

数学不仅是计算工具，更重要的是，他是一种语言

高斯分布函数与导数



机器学习主要涉及的数学知识

高等数学

线性代数

概率论与数理统计

不要怕，不是整本书，只是部分章节，而且是最基础的章节

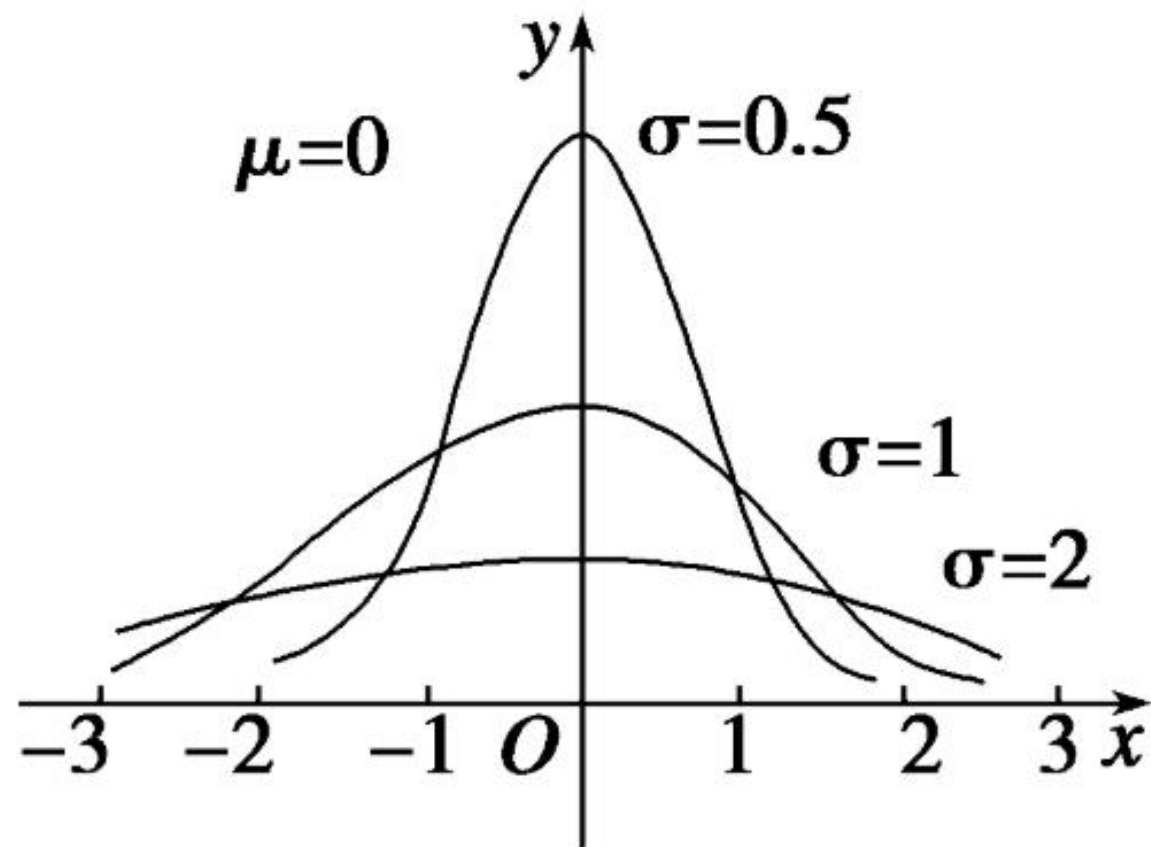
概率与高斯分布函数

概率：上帝的骰子

中心极限定理与高斯分布函数

中心极限定理：

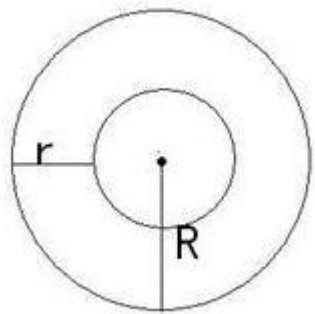
在自然界与生产中，一些现象受到许多相互独立的随机因素的影响，如果每个因素所产生的影响都很微小时，总的影响可以看作是服从正态分布的



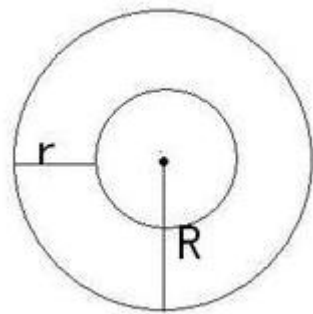
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

导数

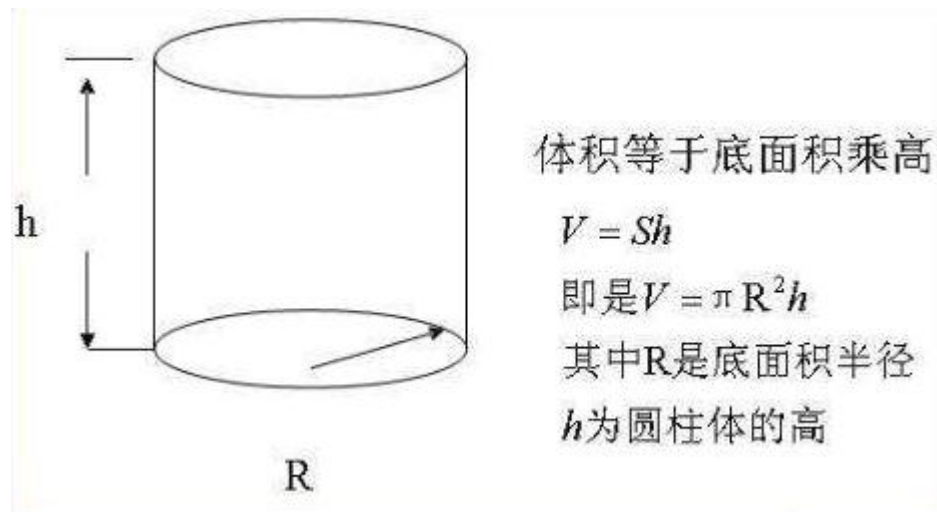
在维度间描述趋势的语法



周长公式: $2\pi r$



面积公式: πR^2



体积公式: $\pi R^2 h$