

引出机器学习

自然学习的启示

在自然界中，动物学习的例子为我们提供了深刻的启示。考虑老鼠面对新奇食物的行为：它们会谨慎地先尝试少量，接下来的进食会取决于食物的味道和生理效果。

当食物对动物产生不良影响时，它们会通过学习机制记住这一经验，以获取对食物安全性的专业知识。过去的负面经历会让动物预测未来遇到相同食物时可能带来负面影响。

心理学家斯金纳 (B. F. Skinner) 进行的实验为我们提供了更深的洞察。他观察了饥饿的鸽子在自动装置的影响下表现出的学习行为。每次食物被随机投放，鸽子们会执行特定的动作，这些动作得到强化，使得它们在下次投放时再次执行相同动作。这一连串的事件形成了鸽子的联想，将食物的投放与它们所做的特定动作联系在一起。食物分发器被设定为每隔15秒掉落一粒食物，不管动物当时在做什么。可以看到这便产生了**非关联性强化**。换句话说，不管动物做了什么，**每隔15秒它将得到一份奖励**。

同样的动作。

是什么将导致迷信的学习机制与有用的学习机制区分开来？

导致迷信的学习机制与有用的学习机制有何区别？这个问题对于开发自动学习器至关重要。

有效学习与盲目迷信

1. 有效学习：

有效学习是指通过有目的地获取、理解和应用知识或技能，以提高个体的能力和智力水平。有效学习通常基于科学的方法和教育原则，包括系统性的学习计划、实践和反馈机制。在有效学习中，重要的是理解概念、培养技能，并能够将所学知识应用于实际问题。

2. 迷信：

与此相反，迷信学习可能是缺乏科学基础、依赖不合理方法的学习方式。这可能包括依赖于超自然力量、盲目记忆而非理解、过度依赖运气等。迷信学习方法通常缺乏实质性的支持，并可能导致表面上的学习，但在理解和应用上存在差距。

为什么需要机器学习？ & 机器学习能做什么

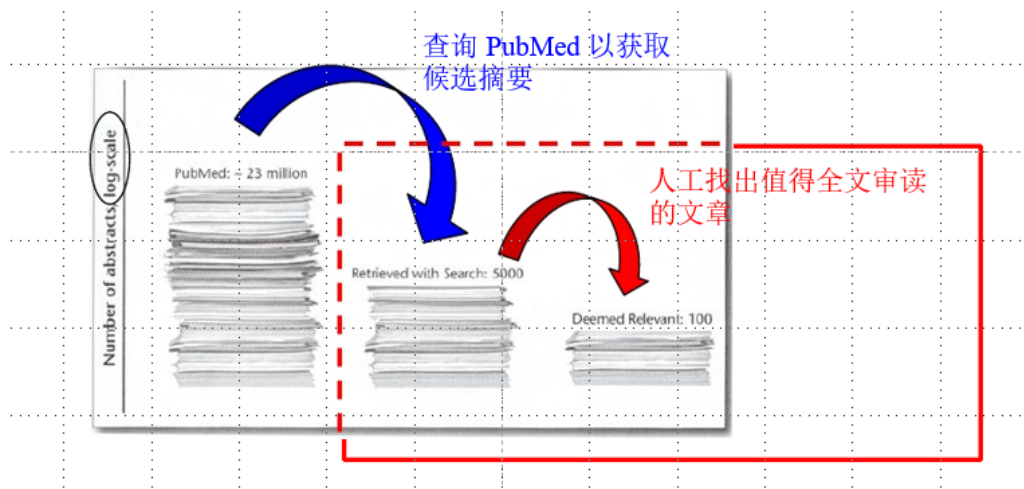
机器学习是从人工智能中产生的一个重要学科分支，是实现智能化的关键；

利用经验来改善系统自身性能

解决复杂的任务 ()

- 光凭人类就能做到的任务（动物/人类能执行的任务）：文献筛选，语音识别，图像识别，图像处理
 - 例子：'文献筛选'：

在“循证医学” (evidence-based medicine) 中，针对特定的临床问题，先要对相关研究报告进行详尽评估。在一项关于婴儿和儿童残疾的研究中，美国Tufts医学中心筛选了约 33,000 篇摘要尽管Tufts医学中心的专家效率 很高，对每篇摘要只需 30 秒钟，为了降低昂贵的成本，Tufts医学中心引入了机器学习技术。——人类专家只需阅读 **50** 篇摘要，系统的自动筛选精度就达到 **93%**。人类专家阅读 **1,000** 篇摘要，则系统的自动筛选敏感度达到 **95%**(人类专家以前需阅读 **33,000** 篇摘要才能获得此效果)



人类可以做到且做得更好的任务，但面对某些场景下造成的人力资源消耗较大。而机器甚至不属于有机生命体，所以解决繁琐任务的重担就落在它身上了

- 超越人类能力的任务：例如根据前几日的天气预测未来的天气；预测股票趋势（这个目前没人能得到）；将各类数据档案转换成人类可直观理解的数值。

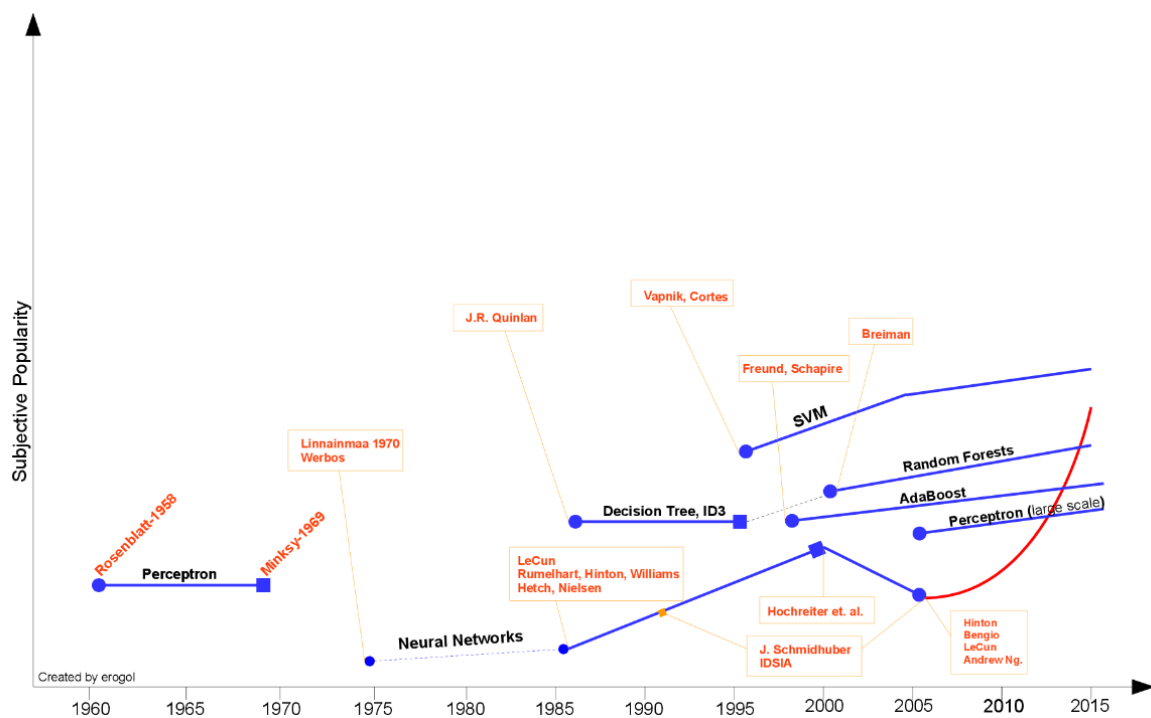
人类无法理解。学习如何在庞大而复杂的数据集中发现有意义的模式

学习检测大型复杂数据集中的有意义模式是一个前景广阔的领域。

在这一领域，学习程序与几乎无限的内存容量和不断提高的计算机处理速度的结合开辟了新天地。

开辟了新天地。

机器学习发展简史



是谁？从哪来？到哪去？

提及机器学习，就必不可少谈到人工智能。人工智能，学习或者智能的任何特性都能够被精确地加以描述，使得机器可以对其进行模拟。

人工智能（Artificial Intelligence, AI）是机器学习的根本框架和更广泛发展的背景。人工智能是一门研究如何使计算机系统具备类似于人类智能的能力的科学和工程领域。这种智能包括理解自然语言、学习、推理、问题解决等一系列复杂的认知能力。

人工智能好像一直作为科幻元素出现在大众视野中。自从AlphaGo战胜李世石之后，人工智能突然间成了坊间谈资，仿佛人类已经造出了超越人类智慧的机器。而人工智能的核心技术机器学习及其子领域深度学习一时间成了人们的掌上明珠。面对这个从天而降的“怪物”，乐观者有之，悲观者亦有之。但追溯历史，我们会发现机器学习的技术爆发有其历史必然性，属于技术发展的必然产物。

- 奠基时期

- Hebb学习理论：假设反射活动的持续性或反复性会导致细胞的持续性变化并增加其稳定性，当一个神经元A能持续或反复激发神经元B时，其中一个或两个神经元的生长或代谢过程都会变化。
- 图灵测试：
- 术语的提出：1952年，Arthur Samuel在IBM研发出了会下跳棋的程序。该程序能够观察位置并学习得出一个模糊的模型，指导其为以后的情况作出更好的移动。Samuel与程序下了很多次棋，并发现程序能够随着时间的增长而下的更好。Samuel借此反驳了公众关于机器执行的命令不能够超出所写代码并如人类一般学习的认知。他创造了“机器学习(machine learning)”这一术语，并定义：

一个通过不明确编程就使计算机拥有某种能力的研究领域。

- 感知机线性分类器&最近邻算法：
- **XOR线性不可分问题**

- 停滞时期

- 问题难以解决：理论研究缓慢，计算机硬件限制
- 资金不足：科研成果难以转化
- 社会批评：大多数哲学问题：‘符号’对机器而言有没有意义？机器到底怎么才算会‘思考’？

- 突破时期

- 提出多层感知机（配合反向传播算法），决策树，玻尔兹曼机

- 成型时期

- 各类优秀模型，算法得以提出

- 爆发时期（深度学习正式起步）

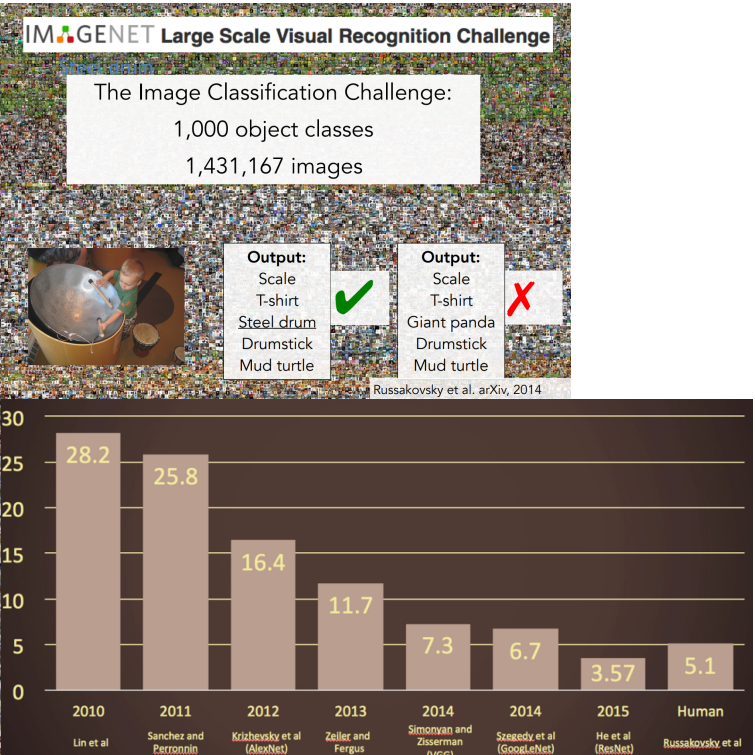
-什么是深度学习？深度学习，深在哪？

深度学习是机器学习的一个分支，其核心特征是使用深层神经网络来建模和解决复杂的任务。深度指的是神经网络中的层数，通常包含多个隐层，使得模型能够学习更加抽象、复杂的特征表示。它学习层次化的特征，逐步从底层的原始特征提取到高层的抽象特征。每一层都通过学习权重和偏差来自动地学习数据的表示，使得模型能够更好地理解和捕捉数据中的模式。这种分层表示的能力使深度学习在处理大规模、高维度数据时表现出色。

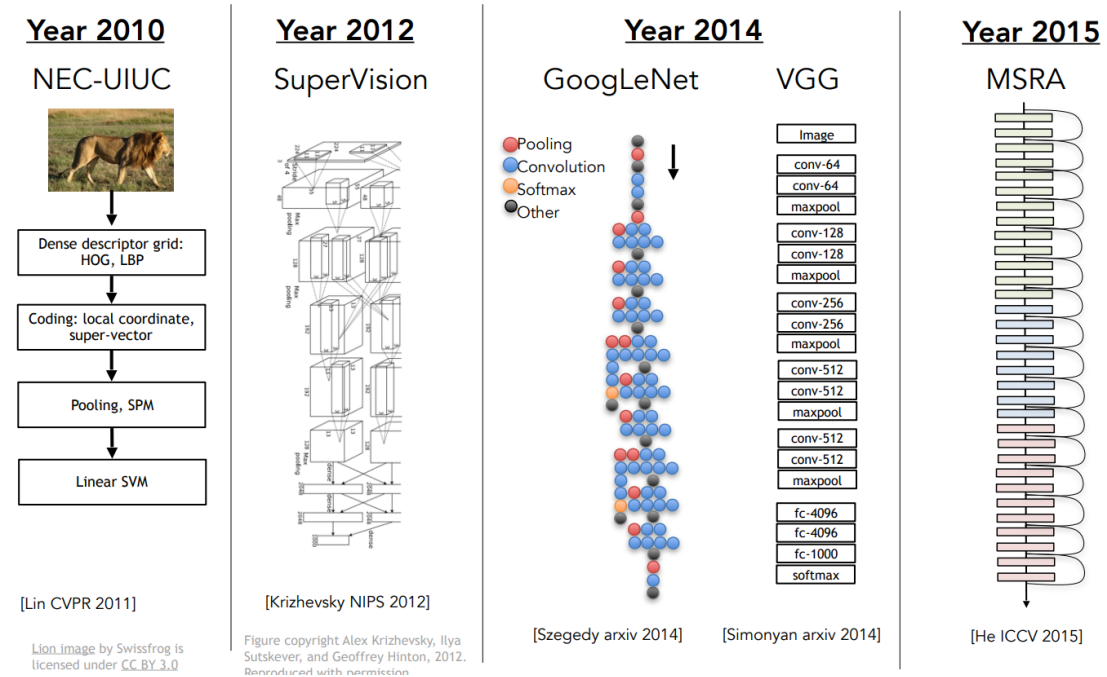
- 我想让机器自动帮我做事
- 计算机硬件发展
- 神经网络（MLP->ANN->SNN）不断被研究
- 反向传播立大功。各类优秀理论的提出
- 优质数据集不断产生
- 为深度学习发展贡献了巨大推动力的是当时的一个大奖赛：大规模图像分类挑战赛 ImageNet（现以停办）：

想象一下，数以百万计的图像被投射到科技的舞台上，挑战着计算机系统对这些图像进行高效分类的能力。ImageNet挑战赛为深度学习提供了一个庞大而真实的测试场景，让算法在海量图像的洪流中自我锤炼。

在这个竞技场上，深度学习模型开始逐渐展露头角，超越了传统方法。深度卷积神经网络（CNN）等先进架构的崭露头角，而这些模型在ImageNet的战场上取得了一次次的胜利。图像分类的准确性突飞猛进，让深度学习在计算机视觉领域崭露头角，也为在其他领域的应用奠定了基础。



IMAGENET大奖赛介绍 & 往年优胜错误率



优秀网络的提出

现在

Generative Pre-trained Transformer (GPT) 给我们带来了无限的可能

Diffusion (扩散模型) ~~ai~~绘画

NeRF (二维图像构建三维场景) :重建城市级别的逼真场景，重建精细的三维模型

SRGAN: 超分辨率

医学图像分割，医学数据集分类

等等

推荐了解

[人工智能发展史](#)

[深度学习发展](#)

[很炫酷的NeRF](#)

Title : 机器学习/深度学习概述

AU : Hiragi

Time : 2023/12/4