

人工智能

Muchen Liu

2024 年 11 月 8 日

目录

1 第一章：人工智能介绍	3
1.1 人工智能的组成部分	4
1.2 人工智能历史	4
1.3 人工智能的应用	7
1.3.1 计算机视觉 (Computer Vision, CV)	7
1.3.2 自然语言处理 (Natural Language Processing, NLP)	7
1.3.3 生成式人工智能 (Generative AI)	7
1.3.4 科学应用的人工智能 (AI for Science)	7
1.4 机器学习	7
1.4.1 背景与定义	7
1.4.2 机器学习与数学	7
1.4.3 机器学习的种类	7
1.4.4 监督学习 (Supervised Learning)	7
1.4.5 无监督学习 (Unsupervised Learning)	7
1.4.6 半监督学习 (Semi-supervised Learning)	7
1.4.7 强化学习 (Reinforcement Learning)	7
1.4.8 机器学习模型的评估	7
2 第二章：Python 介绍	7
2.1 开发环境	8

目录	2
2.1.1 Python 开发环境	8
2.1.2 VS Code	8
2.1.3 Anaconda 和 Jupyter Notebook	8
2.2 Python 语法	8
2.2.1 变量与赋值	8
2.2.2 打印与输入	8
2.2.3 数据类型	8
2.2.4 运算符	8
2.2.5 练习与解答	8
2.3 控制结构	8
2.3.1 条件语句	8
2.3.2 循环	8
2.3.3 for 循环	8
2.3.4 while 循环	8
2.3.5 函数	8
2.3.6 异常处理	8
2.3.7 练习与解答	8
2.4 类	8
2.5 Numpy	8
2.5.1 Numpy 数组	8
2.5.2 Numpy 操作	8
2.5.3 Numpy ndarray	8
2.5.4 Numpy 广播	8
2.6 Matplotlib	8

1 第一章：人工智能介绍

从 2016 年的 AlphaGo(Technologies 2016) 击败李世石开始，到如今以 ChatGPT(OpenAI 2022) 为首的大语言模型持续推进人工智能的飞速发展，这项技术已经脱离了单纯的技术范畴，而是一项具有推动人类社会的、转变社会范式的力量和我们首次达成超人类智慧的可能。

工业革命以来，人们的体力逐步的为工业所取代；而如今，我们站在下一个工业革命的前夕，站在一个全新的、对人们智力的革命。

68 年前，在达特茅斯会议 (Moor 2006) 上，人们对于和人类智慧相似的机器的认识有了最初步的想象；20 年前，人们开始在科幻电影里见到人工智能可能的样子；12 年前，AlexNet 在 ImageNet Challenge(Deng et al. 2009) 中首次达到超越人类水平的识别准确率；9 年前，何凯明副教授团队推出的 ResNet(He et al. 2015) 使得多层神经网络成为可能；8 年前，AlphaGo、AlphaZero 强化学习模型出现，并在围棋等领域超越人类顶尖水平；7 年前，Google Transformer(Vaswani et al. 2017) 模型面世，该模型在随后的几年中一统了人工智能的多个领域；3 年前，AlphaFold(Jumper et al. 2021) 模型在 DeepMind 面世，成功实现了蛋白质结构预测，其作者也获得了今年的诺贝尔化学奖；1 年前，ChatGPT 面世，从旧金山湾区迸发出的人工智能浪潮以前所未有的速度席卷全球，大语言模型随后迅速的带领人们在人工智能最前沿的潮头上高歌猛进，Claude、文心一言、Kimi、Gemini、Grok、runway、Midjourney、Sora、Character.ai 等一众模型以前所未有的速度接踵而至。

随着时间跨度的缩小，人工智能的影响范围从初始的识别手写数字、到围棋游戏、科学预测、成为人类助手、图片视频生成，我们见证着他们的智力水平逐步增加。

同时，社会层面上，人们也时刻关注着这个新科学的发展。人们开始担心：他会取代我的工作吗？我们会不会找不到工作了？我学的东西 AI 是不是都学会了？那我们应该学什么？有什么工作是人工智能无法取代的？诸如此类的担心不但是正常的，更是应当的；当我们站在黎明到来前，我们应当提前进行思考：这份技术带来的到底是什么？对我们的学业、生活、工作、社会到底有什么影响？在这些问题之上，正如 2024 年诺贝尔物理学奖得主、深度学习“教父”Geoffrey Hinton 教授在采访中所述 (Hinton 2024)：“我看

不到一条保证安全的路径。我们正在进入一个充满不确定性的时期。我们要面对的是以前从未处理过的事情。”

但不论如何，为了更好的迎接、改变、创造这个即将来临的时代，我们都应该尽可能的了解人工智能的底层原理，并因此在内心深处生出对未来平和的自信。

在这一章中，我们会深入了解人工智能的子领域和其中的机器学习技术。

1.1 人工智能的组成部分

传统意义上，我们认为人工智能分为弱人工智能和强人工智能。

弱人工智能指的是像工业革命中的机械一样的、辅助人们工作的工具。这也是现阶段所有人工智能模型归属的范畴。弱人工智能的例子有：自动驾驶（辅助人类驾驶的工具）、人脸识别（辅助人类认识人类的工具）、Siri 等智能助手（辅助人类操纵电子设备的工具）等等。

而强人工智能、或通用型人工智能 (Artificial General Intelligence, AGI) 指的是不仅能够显示的使用工具辅助人类，更能自觉的理解工具背后的原理并诞生自己关于怎么使用这些工具的思考 - 也就意味着，拥有和人类同等或超越人类的智力水平。目前学术界和工业界有大约二至三条可能的实现 AGI 的路径：大语言模型（以 OpenAI、Claude 等公司为首）、深度强化学习和泛世界模型（以阿尔伯塔大学 Sutton 教授团队 (Sutton n.d.) 和斯坦福大学李飞飞教授团队 (Li n.d.) 为首）、模型预测控制和世界模型（以纽约大学 Lecun 教授团队 (LeCun n.d.) 为首）。

现阶段人工智能的前沿发展着眼于探索实现 AGI 的道路。

1.2 人工智能历史

人类和人工智能的历史从中世纪就开始了，但我们真正实现这一技术却是在距今不久的时间段内。总的来说，现代人工智能的历史和计算机科学的发展过程是紧密贴合的。

从中世纪的《万物之性》等书，到近代的《弗兰肯斯坦》，我们对于人造人的想象是连续的。而在 20 世纪 40 年代，神经科学证实了大脑是一个

Strong vs Weak AI Comparison Chart	
Weak AI	Strong AI
Weak AI is simply the view that intelligent behavior can be modeled and used by computers to solve complex problems.	Strong AI refers to a hypothetical machine that exhibits human cognitive abilities.
Weak AI refers to systems that are programmed to accomplish a wide range of problems but operate within a predefined range of functions.	Strong AI refers to machines with the mind of their own and which can think and accomplish complex tasks on their own.
Weak AI-powered machines do not have minds of their own.	Strong AI-powered machines can exhibit strong human cognitive abilities.
Alexa and Siri are the best examples of weak AI programs.	Strong AI is a hypothetical concept which does not exist yet in its true form.

图 1: 强人工智能与弱人工智能的区别 (Pangaeax 2022)

由神经元构成的网络之后，一小部分神经科学家、数学家、经济学家、心理学家开始着眼于通过技术手段复刻这个网络。其中，阿兰图灵是最早一批进行哲学性的思考如何去定义智能的人。他提出了图灵测试，即如果一台机器能和人进行无差别的对话，那么我们认为这个机器正在思考。从这个意义上看，ChatGPT 已经比较趋紧这个标准了。

1956 年，在由马文·明斯基和约翰·麦卡锡组织的达特茅斯会议上，参会的科学家首次提出了“人工智能”的理念并将其确立为一门学科，学界广泛认为这是人工智能诞生的时刻。

同年，两位麻省理工的科学家在信息理论特别兴趣小组（Special Interest Group in Information Theory）的一次研讨会上，首次提出逻辑理论家（Logic Theorist），这次会议开始了计算机科学和神经科学新的跨学科范式。

在此后的很长时间，人工智能得到了迅速的发展：说英文、解简单的数学题等等，这也引起了英美学术界和国家机关的重视，多个国家级或大学级别的人工智能实验室成立。

但很快，我们现在熟知的神经网络尽管被首次提出，却没有受到重视，而他的对手符号主义人工智能也没有出现什么令人激动的发展。人类进入了第一个人工智能寒冬（1974-1980）。

这期间，各国各大学对人工智能的资金支持迅速减少。由于人类的算力无法支持任何可能成功的算法真正运转起来，人工智能实验室的资金开

始被削减。在此期间，为数不多的大学还在支持着人工智能领域的研究。

1980 年，专家系统被研发成功-它是一个人工智能程序，但把自己的知识聚焦在某一个具体知识类别上。专家系统带来了广阔的经济效益，各国政府又开始重新支持人工智能的研究。

也是在这个阶段，现在畅行的人工神经网络复苏了。普林斯顿大学教授约翰·霍普菲尔德 (John Hopfield) 证明了一种网络会在学习知识后进行收敛。同一时间杰弗里·辛顿 (Geoffrey Hinton) 和大卫·鲁梅尔哈特 (David Rumelhart) 提出了反向传播算法 - 即教会计算机如何去学习。1990 年，贝尔实验室的杨立昆 (Yann Lecun) 开发出了第一个可以识别手写数字的神经网络。

同时，在这个阶段，另一种形态的人工智能学习模式-强化学习，也被提出了。理查德·萨顿 (Richard Sutton) 和安德鲁·巴托 (Andrew Barto) 首次提出了以人类学习方式训练机器人的思路，也就是说，在表现好的时候收获奖励，在表现不好时收获惩罚。在这一思想下，他们发现 TD (时序差分) 算法，这一算法在 AlphaGo 等现代模型中也被大量应用。

第二次人工智能寒冬在 1990-2012 年之间。此时人工智能领域实际上已经做出了很多前所未有的成就：语音识别、自动驾驶等，但是由于经济原因，这些发展没有被广泛称赞。同时，人工智能研究者的信心也在逐年减退，人们开始可以尽量避免使用人工智能这个词，面的显得想无知的梦想家。

1997 年，“深蓝”成功击败国际象棋世界冠军。2005 年，斯坦福大学的机器人在一条陌生的道路上驾驶了 131 英里。

在这个寒冬中，斯坦福大学的李飞飞教授团队建立了首个供人工神经网络训练的大型数据库 ImageNet。

2012 年，人工智能发展的拐点出现。Alex Krizhevsky、Ilya Sutskever 和 Geoffrey Hinton 开发的 AlexNet 深度学习神经网络在 ImageNet 挑战赛中达到了远低于第二名的错误率。从此，深度学习开始发芽。为了解决深度学习神经网络中梯度消失导致的无法学习的问题，我国计算机科学家何凯明副教授团队开发了 ResNet 残差神经网络，再 AlexNet 的基础上，将神经网络的层数增加到了百级。

随着深度学习的发展，各项人工智能技术向我们袭来 (AlphaGo、ChatGPT、AlphaFold, etc)。详见 1.1 的第 3 自然段。

这也就是人工智能的简单历史。

1.3 人工智能的应用

1.3.1 计算机视觉 (Computer Vision, CV)

1.3.2 自然语言处理 (Natural Language Processing, NLP)

1.3.3 生成式人工智能 (Generative AI)

1.3.4 科学应用的人工智能 (AI for Science)

1.4 机器学习

1.4.1 背景与定义

1.4.2 机器学习与数学

1.4.3 机器学习的种类

1.4.4 监督学习 (Supervised Learning)

1.4.5 无监督学习 (Unsupervised Learning)

1.4.6 半监督学习 (Semi-supervised Learning)

1.4.7 强化学习 (Reinforcement Learning)

1.4.8 机器学习模型的评估

2 第二章：Python 介绍

在本章中，我们系统介绍 Python 编程语言，包括开发环境、语法、控制结构及常用库（如 Numpy 和 Matplotlib）。

2.1 开发环境

2.1.1 Python 开发环境

2.1.2 VS Code

2.1.3 Anaconda 和 Jupyter Notebook

2.2 Python 语法

2.2.1 变量与赋值

2.2.2 打印与输入

2.2.3 数据类型

2.2.4 运算符

2.2.5 练习与解答

2.3 控制结构

2.3.1 条件语句

2.3.2 循环

2.3.3 for 循环

2.3.4 while 循环

2.3.5 函数

2.3.6 异常处理

2.3.7 练习与解答

2.4 类

2.5 Numpy

2.5.1 Numpy 数组

2.5.2 Numpy 操作

2.5.3 Numpy ndarray

2.5.4 Numpy 广播

2.6 Matplotlib

References

- Deng, Jia et al. (2009). “ImageNet: A large-scale hierarchical image database”. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. URL: <https://ieeexplore.ieee.org/document/5206848>.
- He, Kaiming et al. (2015). “Deep Residual Learning for Image Recognition”. In: *arXiv preprint arXiv:1512.03385*. URL: <https://arxiv.org/abs/1512.03385>.
- Hinton, Geoffrey (2024). *Geoffrey Hinton on the Dangers of AI*. Interview by CBS News on 60 Minutes. URL: <https://www.cbsnews.com/news/geoffrey-hinton-ai-dangers-60-minutes-transcript/>.
- Jumper, John et al. (2021). “Highly accurate protein structure prediction with AlphaFold”. In: *Nature* 596.7873, pp. 583–589. URL: <https://www.nature.com/articles/s41586-021-03819-2>.
- LeCun, Yann (n.d.). 纽约大学 *Lecun* 教授团队. 纽约大学计算机科学系. URL: <https://cs.nyu.edu/~yann/>.
- Li, Fei-Fei (n.d.). 斯坦福大学李飞飞教授团队. 斯坦福视觉实验室. URL: <https://vision.stanford.edu/>.
- Moor, James (2006). “The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years”. In: *AI Magazine* 27.4, pp. 87–91. URL: <https://aaai.org/ojs/index.php/aimagazine/article/view/1911/1809>.
- OpenAI (2022). *Introducing ChatGPT*. URL: <https://openai.com/index/chatgpt/>.
- Pangaeax (2022). “AI Types and Applications in Business”. In: *Pangaeax Blog*. URL: <https://www.pangaeax.com/2022/05/09/ai-types-and-applications-in-business/>.
- Sutton, Richard S. (n.d.). 阿尔伯塔大学 *Sutton* 教授团队. 阿尔伯塔大学强化学习研究. URL: <https://www.ualberta.ca/ai/research.html>.
- Technologies, DeepMind (2016). “AlphaGo”. In: URL: <https://en.wikipedia.org/wiki/AlphaGo>.

Vaswani, Ashish et al. (2017). “Attention is All You Need”. In: *arXiv preprint arXiv:1706.03762*. URL: <https://arxiv.org/abs/1706.03762>.