

前导知识

自从 1946 年第一台真正意义上的计算机发明以来，人们一直致力于建造更加智能的计算机。随着算力的提高和数据的增长，人工智能（Artificial Intelligence, AI）获得了空前的发展，在一些任务上的表现甚至已经超越人类，比如围棋、象棋，以及一些疾病诊断和电子游戏等。人工智能技术还能被广泛用于其他应用中，比如药物发现、天气预测、材料设计、推荐系统、机器感知与控制、自动驾驶、人脸识别、语音识别和对话系统。

近十年来，很多国家，比如中国、英国、美国、日本、德国，对人工智能进行了大量的投入。与此同时，还有很多科技巨头，比如 Google、Facebook、Microsoft、Apple、百度、华为、腾讯、字节跳动和阿里巴巴等，也都积极地参与其中。人工智能在我们的日常生活中正变得无处不在，如自动驾驶汽车、人脸 ID 和聊天机器人。毫无疑问，人工智能对人类社会的发展至关重要。

在我们深入阅读本书之前，第一步应该先了解人工智能领域不同的子领域，如机器学习（Machine Learning, ML）、深度学习（Deep Learning, DL）、强化学习（Reinforcement Learning, RL），以及本书的主题——深度强化学习（Deep Reinforcement Learning, DRL）。图 1 用韦恩图（Venn Diagram）展示了它们之间的关系，下面将会逐一介绍它们。

人工智能

虽然科学家一直以来都在努力让计算机变得越来越智能，但是“智能”的定义直到今天依然是非常模糊的。在这个问题上，Alan Turing 最早在他 1950 年曼城大学时的文章 *Computing Machinery and Intelligence* 中介绍了图灵测试（Turing Test）。图灵测试可以用来衡量机器模拟人类行为的能力大小。具体来说，它描述了一个“imitation game”，一个质问者向一个人和一台计算机提出一系列问题，用以判断哪个是人，哪个是机器。当且仅当质问者不能分辨出人和机器时，图灵测试就通过了。

人工智能的概念最早是由 John McCarthy 在 1956 年夏天的达特茅斯（Dartmouth）会议上提出的。这次会议被认为是人工智能正式进入计算机科学领域的开端。最早期的人工智能算法主要用于解决可以被数学符号和逻辑规则公式化的问题。

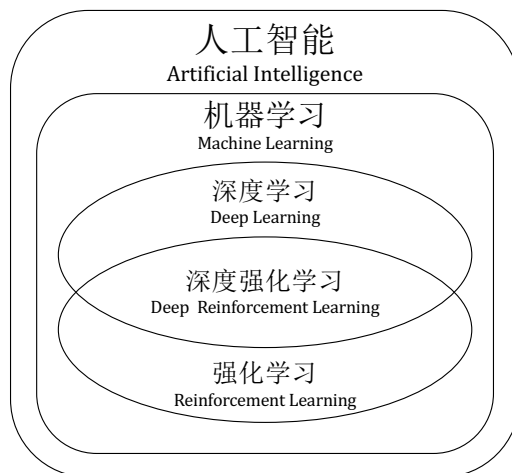


图 1 人工智能、机器学习、深度学习、强化学习及深度强化学习之间的关系

机器学习

机器学习 (Machine Learning, ML) 的概念和名字是由 Arthur Samuel (Bell Labs, IBM, Stanford) 在 1959 年首次提出来的。一个人工智能系统需要具备从原始数据中学习知识的能力, 这个能力就称为机器学习。很多人工智能问题可以被这样解决: 通过设计有针对性的模式识别算法来从原始数据中提取有效特征, 然后用机器学习算法使用这些特征。

比如, 在早期的人脸识别算法中, 我们需要特殊的人脸特征提取算法。最简单的方法就是使用主成分分析 (Principal Component Analysis, PCA) 降低数据的维度, 然后把低维度特征输入一个分类器获得结果。长期以来, 人脸识别需要纯手工设计的特征工程算法。针对不同问题设计特征提取算法的过程非常耗时, 而且在很多任务中设计有针对性的特征提取算法的难度非常大。比如, 语言翻译的特征提取需要语法的知识, 这需要很多语言学专家帮助。然而, 一个通用的算法应该具备从对不同任务自行学习出特征提取算法, 以大大降低算法开发过程中所需的人力的先验知识。

学术界有很多研究, 使得机器学习能自动学习数据的表征。表征学习的智能化不仅可以提升性能, 还能降低解决人工智能问题的成本。

深度学习

深度学习是机器学习中的一个子领域, 与其他算法不同, 它主要基于人工神经网络 (Artificial Neural Network, ANN) (Goodfellow et al., 2016) 来实现。我们之所以称它为神经网络, 是因为它是由生物神经网络启发设计的。Warren Sturgis McCulloch 和 Walter Pitts 在 1943 年共同发表的 *A Logical Calculus of the Ideas Immanent in Nervous Activity* (McCulloch et al., 1943) 被视为人工神经

网络的开端。至此，人工神经网络作为一种全自动特征学习器，使得我们不需要对不同数据开发特定的特征提取算法，从而大大提高了开发算法的效率。

深度神经网络（Deep Neural Network, DNN）是人工神经网络的“深度”版本，有很多的网络层，深层的网络相比浅层的网络具有更强的数据表达能力。图 2 展示了深度学习方法与非深度学习方法的的主要区别。深度学习方法让开发者不再需要针对特定数据来设计纯手工的特征提取算法。我们因此也称这些学习算法为端到端（End-to-end）方法。但值得注意的是，很多人质疑，深度学习方法是一个黑盒子（Black-box），我们并不知道它是如何学到数据特征表达的，往往缺乏透明性和可解析性。

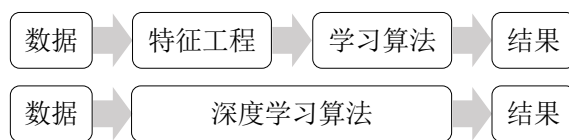


图 2 深度学习方法与非深度学习方法的差别

虽然现在看来，深度学习非常流行，但是在人工神经网络早期发展阶段，受制于当时计算机算力和黑盒子问题，实际应用很少，并未受到学术界的广泛关注。

这种情况直到 2012 年才得到了改变，当年一个叫 Alexnet (Krizhevsky et al., 2012) 的模型在 ImageNet 图像分类竞赛 (Russakovsky et al., 2015) 中取得了超过其他方法 10% 以上的性能。从此，深度学习开始受到越来越多的关注，深度学习方法开始在很多不同领域超越非深度学习方法，比如大家熟悉的计算机视觉 (Girshick, 2015; Johnson et al., 2016; Ledig et al., 2017; Pathak et al., 2016; Vinyals et al., 2016) 和自然语言处理 (Bahdanau et al., 2015)。

强化学习

深度学习虽然具有了很强大的数据表达能力，但不足以建立一个智能的人工智能系统。这是因为人工智能系统不仅需要从给定的数据中学习，而且还要像人类那样学习与真实世界交互。强化学习作为机器学习的一个分支，即可让计算机与环境进行交互学习。

简单来说，强化学习把世界分为两个部分：环境（Environment）与智能体（Agent）。智能体通过执行动作（Action）来与环境交互，并获得环境的反馈。在强化学习中，环境的反馈是以奖励（Reward）形式体现的。智能体学习如何“更好”地与环境交互，以尽可能获得更大的奖励。这个学习过程建立了环境与智能体间的环路，通过强化学习算法来提升智能体的能力。

深度强化学习

深度强化学习结合了深度学习和强化学习各自的优点来建立人工智能系统，主要在强化学习中使用深度神经网络的强大数据表达能力，例如价值函数（Value Function）可以用神经网络来近似，以实现端到端的优化学习。

DeepMind 是一家成立于伦敦、以科研为主导的人工智能技术公司，在深度强化学习历史上具有非常重要的地位。2013 年，仅在 AlexNet 提出一年以后，他们就发表了论文 *Playing Atari with Deep Reinforcement Learning*，该文基于电子游戏的原始画面作为输入，学习了 7 种游戏。DeepMind 的方法不需要手工设计特征提取算法，在 6 个游戏中优于之前的方法，甚至在 1 个游戏中赢了人类。

2017 年，DeepMind 的 AlphaGO 围棋算法在中国打败了世界第一围棋大师——柯洁。该事件标志着人工智能具备比人类更好表现的潜力。深度强化学习是机器学习的一个子领域，具有实现通用人工智能（Artificial General Intelligence, AGI）的潜力。但是还有很多挑战需要我们解决，才能真正地实现这个理想的目标。

TensorLayer

强化学习的算法很多，而且从学习算法到实现算法有一定的距离。因此，本书中很多章节有实现教学，我们会展示一些算法中的关键部分是如何实现的。自从深度学习变得流行以来，出现了很多开源的框架，比如 TensorFlow、Chainer、Theano 和 PyTorch 等，以支持神经网络的自动优化。在本书中，我们选择 TensorLayer，一个为科研人员和专业工程师设计的深度学习与强化学习库。该库获得了 ACM Multimedia 2017 年度最佳开源软件奖。在本书定稿时，TensorLayer 2.0 支持 TensorFlow 2.0 作为后端计算引擎，而在下一版本中，TensorLayer 将会支持更多的其他计算引擎，如华为 MindSpore，以更好地支持国内外的 AI 训练芯片。更多关于 TensorLayer 的最新信息，请访问 GitHub 页面¹。

参考文献

BAHDANAU D, CHO K, BENGIO Y, 2015. Neural machine translation by jointly learning to align and translate[C]//Proceedings of the International Conference on Learning Representations (ICLR).

GIRSHICK R, 2015. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision (ICCV). 1440-1448.

GOODFELLOW I, BENGIO Y, COURVILLE A, 2016. Deep learning[M]. MIT Press.

¹链接见读者服务

- JOHNSON J, ALAHI A, FEI-FEI L, 2016. Perceptual Losses for Real-Time Style Transfer and Super-Resolution[C]//Proceedings of the European Conference on Computer Vision (ECCV).
- KRIZHEVSKY A, SUTSKEVER I, HINTON G E, 2012. Imagenet classification with deep convolutional neural networks[C]//Proceedings of the Neural Information Processing Systems (Advances in Neural Information Processing Systems). 1097-1105.
- LEDIG C, THEIS L, HUSZAR F, et al., 2017. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- MCCULLOCH W S, PITTS W, 1943. A logical calculus of the ideas immanent in nervous activity[J]. The bulletin of mathematical biophysics, 5(4): 115-133.
- PATHAK D, KRAHENBUHL P, DONAHUE J, et al., 2016. Context encoders: Feature learning by inpainting[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2536-2544.
- RUSSAKOVSKY O, DENG J, SU H, et al., 2015. Imagenet Large Scale Visual Recognition Challenge[J]. International Journal of Computer Vision (IJCV), 115(3): 211-252.
- VINYALS O, TOSHEV A, BENGIO S, et al., 2016. Show and tell: Lessons learned from the 2015 mscoco image captioning challenge[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI).