中国科学 : 信息科学
SCIENTIA SINICA Informationis
论文

自主机器学习

朱文武[1], 王鑫[1*], 徐宗本[2]

1. 清华大学计算机科学与技术系, 北京信息科学与技术国家研究中心, 北京 100084
2. 西安交通大学数学与统计学院, 西安 710049
* 通信作者. E-mail: xin_wang@tsinghua.edu.cn

收稿日期: 2025–04–07; 修回日期: 2025–06–21; 接受日期: 2025–07–27; 网络出版日期: 2025–09–10

科技创新 2030 “新一代人工智能” 重大项目 (批准号: 2020AAA0106300)、国家自然科学基金原创探索计划项目 (批准号: 62250008)、国家优秀青年科学基金项目 (批准号: 62222209) 和北京信息科学与技术国家研究中心项目 (批准号: BNR2023RC01003) 资助

摘要 当今机器学习 (machine learning) 能够针对给定任务通过数据训练并构建机器学习模型, 使模型获得在相应任务上的预测和决策能力, 可归纳为任务驱动的机器学习. 现有方法依赖于外界人工引导, 并基于经验对学习过程进行数据与任务指定、模型设定与参数学习, 在真实动态开放环境中, 无法像人类一样自主学习. 本文提出自主机器学习 (autonomous machine learning) 的新概念. 具体而言, 本文将自主机器学习设计为一个由自我驱动、具有动态自主演化的学习过程, 包括自主优化和自主演化. 首先能够主动探索感知环境进行数据自选择、模型自适应、任务自切换, 而不需要人工干预; 同时, 根据环境学习的反馈结果和自我状态进行动态自主演化学习, 进一步提升自我. 此外, 本文还展示了一些自主机器学习的应用案例, 最后讨论了未来的研究方向. 我们期望自主机器学习可以使机器进行像人类一样的自主学习, 并为通用人工智能提供一个新视角.

关键词 自主机器学习, 自我优化, 自主演化

1 引言

机器学习在自然语言处理、计算机视觉、机器人等多个领域已经取得了巨大成功, 成为推动人工智能发展的关键技术. 为了实现通用人工智能 (artificial general intelligence, AGI), 研究者们一直在努力减少学习过程中的人工设计. 无监督学习、半监督学习和自监督学习等方法已经在减少人工标注方面取得了一定进展 [1~5], 而主动学习通过交互式查询用户或其他信息源来标注 “重要” 的新数据点, 从而降低了数据标注的成本 [6~8]. 此外, 自动机器学习 (AutoML) 利用神经架构搜索和超参数优化来减少模型设计与选择中的人工工作量 [9~11], 而元学习则通过元学习器快速调整学习算法, 以适应少量新数据和新任务 [12~14].

尽管现有方法在一定程度上减少了人工干预, 但现行的机器学习范式仍然严重依赖人工设计与人类指导, 尤其是在任务选择、数据标注、模型设定等方面. 举例来说, 在开发救援机器人系统时, 人工专家依然需要设计如何在人工选择的数据集上训练推理模型、如何自动构建知识图谱等. 这种依赖人

引用格式: 朱文武, 王鑫, 徐宗本. 自主机器学习. 中国科学: 信息科学, 2025, doi: 10.1360/SSI-2025-0137
Zhu W W, Wang X, Xu Z B. Autonomous machine learning. Sci Sin Inform, 2025, doi: 10.1360/SSI-2025-0137

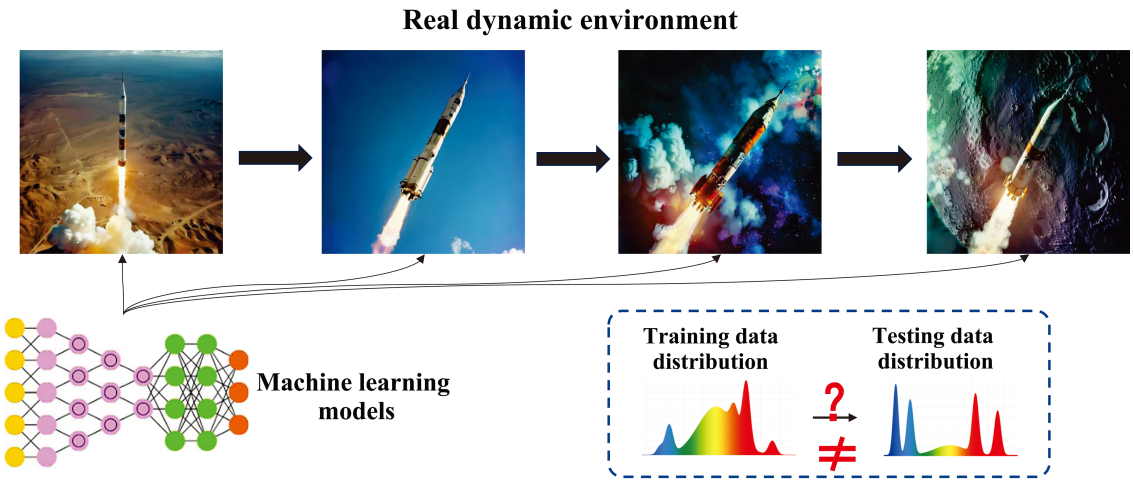

图 1　(网络版彩图) 传统机器学习模型面临动态环境下数据分布不一致性的挑战.

Figure 1　(Color online) Traditional machine learning models face the challenge of inconsistent data distributions in dynamic environment.

工设计的方式不仅耗费大量的人力资源与时间成本, 而且缺乏足够的自主性. 为了突破这一瓶颈, 探索是否可以让机器自主选择学习任务、数据和模型, 并自主地控制学习过程, 成为了学术界日益关注的研究方向.

自 20 世纪 70 年代以来, 人类自主学习 (self-directed learning) 一直是教育科学领域的重要课题 [15,16], 特别是在成人教育 [17,18] 和在线教育 [19,20] 等需要高度学习主动性的场景中. 研究表明, 主动进行自主学习的个体比被动等待指导的学习者能更好地实现学习目标 [21]. 这一点为我们提供了启示: 如果机器能够像人类一样进行自主学习, 并主动选择学习任务、数据、模型与策略, 是否能够为机器学习带来更多的自主性与灵活性?

基于这一思考, 本文提出了自主机器学习 (autonomous machine learning) 的新概念. 不同于传统机器学习依赖人工设计的方式, 自主机器学习通过自我驱动来实现高度自主的学习过程, 其核心是动态的自我优化和自主演化. 在这个过程中, 机器能够主动探索并感知环境, 自主选择数据、适应模型、切换任务, 无需人工干预. 同时, 机器根据环境反馈和自我状态, 通过动态自主演化进一步提升其学习能力. 这种自我驱动的学习过程不仅使机器在没有人类干预的情况下进行学习, 而且在每次学习后, 能够基于反馈进行调整和优化, 从而不断演化以实现自我提升. 因此, 本文提出的自主机器学习不仅是传统机器学习方法的拓展, 更是实现机器自主学习和自我优化的全新尝试, 为通用人工智能的实现提供了一种新的思路.

本文余下内容结构如下. 第 2 节介绍了自主机器学习的基本概念和框架. 第 3 节通过讨论自主机器学习的几个应用, 如自主救援机器人、无人驾驶和智能无人机, 展示了自主机器学习相对于传统机器学习的优势. 第 4 节讨论了未来的方向, 包括使自主机器学习具备鲁棒性、可解释性、推理能力和意识驱动能力.

2　自主机器学习理论

2.1　真实动态环境

在真实世界中, 我们所遇到的环境总是不断动态变化的, 如图 1 所示, 在一个火箭从地球发射到月球的过程中, 会经历地表、大气层、太空、最后到达月球表面等不同的阶段, 这也就需要其能够具备适应不同环境的能力, 能够在动态变化的环境中作出正确的反应. 类似地, 我们也希望机器学习具备

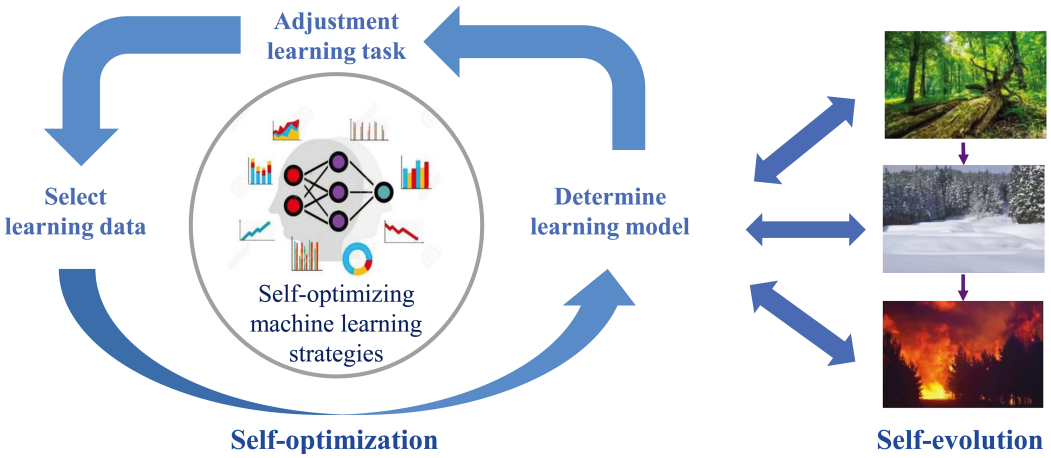

图 2　(网络版彩图) 自主机器学习框架.

Figure 2　(Color online) The framework of autonomous machine learning.

适应动态变化环境的能力. 然而, 传统的机器学习普遍都是学习预先固定的数据分布, 在动态环境中往往面临着训练与测试分布不一致的挑战, 难以适应变化之后的数据新分布.

2.2　自主机器学习

受教育学中人类自主学习概念的启发, 我们提出了一个新的机器学习理论框架, 称为自主机器学习. 传统的机器学习大多是人工驱动的, 只能处理定义明确的特定任务, 但无法自动适应不断变化的环境. 与之相反, 自主机器学习在学习过程中采取主动, 并追求终身自我完善. 具体而言, 自主机器学习通过自我驱动来进行具有高度自主性的自主学习过程, 其学习结果反过来又可以进一步提升自我. 这一过程的核心在于自我优化和自主演化, 自主机器学习主动探索感知环境, 对数据选择、模型适配、任务切换三大核心过程进行动态自我优化, 同时根据环境反馈和自我状态, 通过动态自主演化进一步提升学习能力, 在自我驱动指导下以一种更为自主的方式来进行学习. 自主机器学习框架如图 2 所示.

2.3　自我优化

在与外部动态环境进行交互时, 自主机器学习可以通过自我优化学习策略, 如图 3 所示, 自主选择学习任务、数据、模型, 得到的结果又会反过来为提升自我提供反馈, 从而有别于传统的机器学习范式. 我们将详细介绍自我优化学习策略的每个组成部分, 即数据自选择、模型自适应以及任务自切换.

数据自选择. 人类有找到与给定任务最相关的材料并进行学习的能力, 这样我们就能够以快速且高效的方式提高我们在给定任务上的能力. 人类通常为给定的问题选择合适的材料进行学习, 自主机器学习也应该具备这样的能力, 能够在意识到最终目标、选定的任务和自身当前状态的情况下选择合适的数据进行学习.

形式上, 给定任务序列 $\mathbb{T}$、机器状态 $\mathbb{S}$、机器能力 $\mathbb{C}$、评估结果 O 以及最终目标 T^*, 自主机器学习会为每个任务选择最合适的数据集, 即 $\mathbb{D}=\{D_i\}_{i=0}^{N_T}=f_D(T^*,T_i,\mathbb{C},\mathbb{S},O)$, 其中 $T_i\in\mathbb{T}$. 为了获得更高的效率, 数据选择的过程不仅依赖于所选任务, 还依赖于最终目标, 因为它的目的是选择与最终目标最相关的数据, 同时舍弃不相关的、有噪声甚至有害的数据. 因此, 自主机器学习的设计目标是一种高效且快速的方式改进自己. 数据选择的结果也将影响模型的设计和优化策略的选择.

模型自适应. 当给定不同的学习任务时, 人类能够找到潜在的可选解决方案. 类似地, 在调整了学习任务之后, 自主机器学习需要为每个任务选择学习模型. 形式上, 自主机器学习根据任务序列 $\mathbb{T}$、机器能力 $\mathbb{C}$、机器状态 $\mathbb{S}$ 和评估结果 O 来设计模型 $\mathbb{M}=\{M\}_{i=0}^{N_T}=f_M(T_i,D_i,\mathbb{C},\mathbb{S},O)$. 学习模型的设计同样会影响优化策略和评估指标的选择, 而所选模型以及相应的模型性能将会为任务选择组件提供

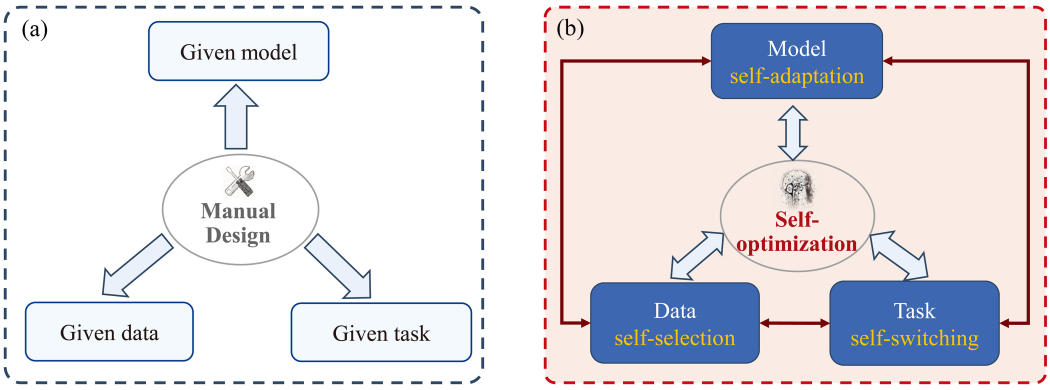

图 3　(网络版彩图) 传统机器学习 (a) 与自我优化学习策略 (b) 的对比.

Figure 3　(Color online) Comparison of (a) conventional machine learning and (b) self-optimization learning strategies.

反馈.

由于人们的能力和可用的学习时间不同, 每个人都有自己的学习速度和风格. 因此, 人们倾向于选择最适合自己的学习策略. 受此启发, 自主机器学习算法在机器自我驱动的指导下选择学习策略. 到目前为止, 研究人员已经提出了许多种优化策略 [22~25], 以减少在数据、监督、损失和优化上的人为干预. 优化策略集在自我驱动中被编码和更新, 它会影响学习的速度、程度和成本等. 优化策略在自主机器学习中是最为灵活的, 可以根据任务和模型来进行设置. 优化策略集也能够成为选择任务和设计模型的关键因素. 形式上, 所选的优化策略是通过考虑任务、模型以及机器的自我状态来进行建模的: $\mathbb{P} = F_P(\mathbb{T}, \mathbb{M}, \mathbb{S})$. 最终, 所选优化策略、结果与评估将为进一步更新自主学习过程中的所有组件提供反馈.

设置适当的评估指标在机器学习中至关重要. 传统的机器学习通过人工设计来设定一种或多种评估指标. 然而, 目前存在着各种各样的评估指标, 甚至连研究人员也不清楚在不同的情况下应该使用哪种指标. 此外, 在模型训练的过程中, 不同的评估指标往往会导向不同的优化方向. 我们希望自主机器学习能够以自主的方式来进行评估, 例如, 从大量的可选指标中自适应地选择评估指标, 甚至生成新的评估指标. 当实现了没有明确定义的复杂目标时, 机器将借助成功的自我评估于世界范围内实现终身的自我完善.

任务自切换. 人类通常可以将一个复杂的最终目标分解为多个细化的任务, 这些任务可能是实现最终目标的先决条件或对其有帮助. 例如, 如果我们计划做一顿丰盛的饭菜, 我们将会将这个最终目标分解为以下几个任务: 购买食材、准备食材、烹饪和调味. 其中的每个任务又可以被分解成更小的任务, 并且我们将会将注意力放在我们完成得最差的任务上, 从而能够在很大程度上提升我们完成最终目标的能力.

受此启发, 自主机器学习的设计目标是将一个复杂的目标自动分解为一系列细化的子任务. 自主机器学习可以选择学习任务的一个子集, 并在自我驱动的指导下决定它们的执行顺序. 形式上, 我们将最终目标定义为 $T^* = \{T_C^*, T_O^*, T_E^*, T_S^*\}$. 注意, 机器学习任务图 G_T 被包含在机器自我状态中. 因此, 给定 T^*、G_T、机器能力 $\mathbb{C}$、机器状态 $\mathbb{S}$ 和评估结果 O, 自主机器学习能够生成一个任务序列 $\mathbb{T} = \{T_i\}_{i=0}^{N_T} = f_T(T^*, G_T, \mathbb{C}, \mathbb{S}, O)$ 的函数 f_T, 其中有 N_T 个可以重复的任务从 G_T 中被选择出来. 这个函数可以通过不同任务的输入、输出、评估指标和实验结果之间的相似性匹配来建模. 所选任务将会影响对数据、模型和优化策略的选择.

自我优化的数学表述. 我们将通过整合在前文中介绍过的元素和过程, 提出一种用于自我优化的数学表述. 具体而言, 我们提出了一个基于多级优化的框架来优化自主机器学习. 在这个框架中, 存在着多个联合优化问题, 每个问题都对应于 2.2 节中描述的一个过程. 这些过程被组织为一个有向图

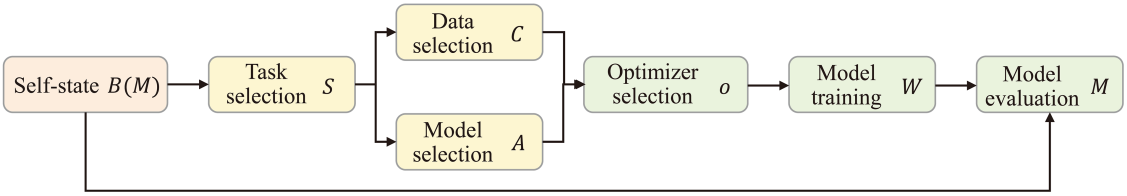

图 4　(网络版彩图) 自我优化过程的有向图表示.

Figure 4　(Color online) Directed acyclic graph representation of the self-optimization process.

(directed acyclic graph, DAG), 如图 4 所示. 如果从过程 A 到过程 B 有边, 则 B 依赖于 A, 即 A 优化问题的最优解被用作 B 优化问题中的一个变量. 这些优化问题一共被分为 6 个层次.

在第 1 层中, 我们通过解决如下优化问题来构建自我状态:

$$B^*(M) = \text{argmin}_B\ L_{\text{sac}}(B, M), \tag{1}$$

其中, B 是自我状态, L_{sac} 是自我状态构建的损失函数, M 是元参数. 该优化在 B 上进行. M 在这个阶段被暂时固定, 将会在之后的阶段中被更新. 注意, 最优解 B^* 是关于 M 的一个函数, 因为 B^* 是关于损失函数的函数, 而损失函数是关于 M 的函数.

在第 2 层中, 我们进行任务选择. 给定已构建的自我状态 $B^*(M)$、可选任务集 $\mathcal{T} = \{t_n\}_{n=1}^N$ 以及目标应用的文本描述 E, 任务选择的目标是选择任务的子集 $\mathcal{S} \subseteq \mathcal{T}$, 并将它们组织为一个有向图. 这个任务有向图体现了任务之间的依赖关系和任务的执行顺序. 在这个阶段, 我们解决如下的优化问题:

$$\mathcal{S}^*(B^*(M)) = \text{argmin}_{\mathcal{S}\subseteq\mathcal{T}}\ L_{ts}(\mathcal{T}, \mathcal{S}, E, B^*(M)), \tag{2}$$

其中, L_{ts} 是任务选择的损失函数, 规定了如何选择最优任务子集并将其组织为有向图的标准. 它被定义在全体任务集 $\mathcal{T}$、可选任务子集 $\mathcal{S}$、应用描述 E 和自我状态 $B^*(M)$ 上. 该优化是在 S 上进行.

在第 3 层中, 存在着两个过程: 一个是训练数据选择, 另一个是模型选择. 数据选择被定义如下: 对于所选任务子集 $\mathcal{S}^*(B^*(M))$ 中的每个任务 $s(M)$, 我们从任务 $s(M)$ 的训练数据 $D_{s(M)}$ 中选择训练样本子集 $C_{s(M)} \subseteq D_{s(M)}$. 任务 $s(M)$ 的数据选择相当于解决如下优化问题:

$$C^*_{s(M)} = \text{argmin}_{C_{s(M)}\subseteq D_{s(M)}}\ L_{ds}(D_{s(M)}, C_{s(M)}), \tag{3}$$

其中, L_{ds} 是数据选择的损失函数, 规定了如何选择最优训练数据子集的标准. 它被定义在任务 $s(M)$ 的全体训练数据集 $D_{s(M)}$ 和可选数据样本 $C_{s(M)}$ 上, 该优化在 $C_{s(M)}$ 上进行.

模型选择被定义如下: 对于所选任务子集 $\mathcal{S}^*(B^*(M))$ 中的每个任务 $s(M)$, 给定模型用来解决任务 $s(M)$ 所用架构和超参数的搜索空间, 我们从中选择最优的架构和超参数. 相应的优化问题如下:

$$A^*_{s(M)} = \text{argmin}_{A_{s(M)}}\ L_{ms}(A_{s(M)}, s(M)), \tag{4}$$

其中, L_{ms} 是模型选择的损失函数, 规定了如何设置最优的架构和超参数的标准. $A_{s(M)}$ 是搜索空间中的可选架构和超参数, 该优化是在 $A_{s(M)}$ 上进行的.

在第 4 层中, 仅有一个过程, 即优化器选择. 对于每个所选任务 $s(M)$, 给定可选优化器集 $\mathcal{O} = \{o_n\}_{n=1}^P$, 我们从中选择最优的 $o^*_{s(M)} \in \mathcal{O}$ 用于在所选数据上训练所选模型. 相应的优化问题如下:

$$o^*_{s(M)} = \text{argmin}_{o_{s(M)}\in\mathcal{O}}\ L_{os}(o_{s(M)}, A^*_{s(M)}, C^*_{s(M)}), \tag{5}$$

其中, L_{os} 是规定了如何选择最佳优化器的标准的损失函数, 优化器的选择依赖于所选模型 $A^*_{s(M)}$ 和所选数据 $C^*_{s(M)}$.

在第 5 层中, 仅有一个过程, 即使用所选优化器来训练所选模型在所选数据上的权重参数, 等同于解决如下优化问题:

$$W^*_{s(M)} = \text{argmin}_W \ L_{wt}(W_{s(M)}, A^*_{s(M)}, C^*_{s(M)}, o^*_{s(M)}). \tag{6}$$

在第 6 层中, 仅有一个过程, 即在验证集 F 上评估在第四层中所训练的模型. 元参数 M 将通过最小化验证损失被更新, 等同于解决如下优化问题:

$$\text{min}_M \ L_{val}(\{W^*_{s(M)} | s(M) \in S^*(M)\}, F). \tag{7}$$

其中, L_{val} 是验证集上的损失函数.

将上述环节整合在一起, 得到了如下的多级优化问题:

$$\begin{aligned}
&\text{min}_M \ L_{val}(\{W^*_{s(M)} | s(M) \in S^*(M)\}, F) \\
&\text{s.t. } W^*_{s(M)} = \text{argmin}_W \ L_{wt}(W_{s(M)}, A^*_{s(M)}, C^*_{s(M)}, o^*_{s(M)}), \\
&\quad o^*_{s(M)} = \text{argmin}_{o_{s(M)} \in \mathcal{O}} \ L_{os}(o_{s(M)}, A^*_{s(M)}, C^*_{s(M)}), \\
&\quad A^*_{s(M)} = \text{argmin}_{A_{s(M)}} \ L_{ms}(A_{s(M)}, s(M)), \\
&\quad C^*_{s(M)} = \text{argmin}_{C_{s(M)} \subseteq D_{s(M)}} \ L_{ds}(D_{s(M)}, C_{s(M)}), \\
&\quad \mathcal{S}^*(B^*(M)) = \text{argmin}_{\mathcal{S} \subseteq \mathcal{T}} \ L_{ts}(\mathcal{T}, \mathcal{S}, E, B^*(M)), \\
&\quad B^*(M) = \text{argmin}_B \ L_{\text{sac}}(B, M).
\end{aligned} \tag{8}$$

2.4　自主演化

在当前动态环境中, 机器系统面临着持续更新自我认知以适应环境变化的挑战. 这一挑战的核心在于如何使机器能够自适应地感知环境的持续变化, 并根据这些变化自适应更新其对世界的理解以及自身的行为模式. 具体而言, 这涉及机器如何通过调整训练数据、任务目标及模型参数来实现对环境变化的响应. 当前面临的主要难题之一是, 缺乏一套完整的机器自主演化学习理论方法, 这种方法需要能够在没有人为干预的情况下, 让机器自主地识别环境变化, 并据此优化自身的性能.

为了应对上述挑战, 机器必须具备几个关键能力. 首先, 它需要拥有强大的环境感知能力, 能够实时且准确地捕捉周围环境的变化. 其次, 基于所感知的信息, 机器应能自我评估并动态更新其内部的认知模型, 确保其理解和反应方式始终符合最新的环境条件. 此外, 机器还需具备一种机制, 可以根据环境变化自动调整其使用的训练数据集、任务定义以及模型架构, 从而实现更加灵活和有效的操作. 这种自主联合动态演化的方法将使得机器不仅能在静态环境下高效工作, 也能够在面对未知或动态变化时展现出渐进式的智能提升.

为了解决这些问题, 我们提出了自主演化学习框架, 通过环境学习来充分建模和感知动态环境, 并通过数据 – 任务 – 模型的动态演化来记住旧知识和适应新环境, 框架图如图 5 所示. 形式上, 自主演化根据机器自我状态 B 与环境状态 $\mathcal{E}$, 通过机器与环境持续交互, 进行任务选择 $\mathcal{S}$、记忆选择 $\mathcal{C}$、模型选择 A, 通过在验证集 F 上评估, 实现元参数的持续更新:

$$\text{min}_M \ L_{val}(B, \mathcal{E}, \mathcal{S}, \mathcal{C}, A | F), \tag{9}$$

其中 L_{val} 是验证集上的损失函数. 我们希望未来的探索将集中在开发支持自主演化学习的新理论和技术上. 这包括跨学科的合作研究, 结合人工智能、生物学、心理学等领域的知识, 共同探讨生物体如何在自然环境中实现动态自我认知的更新, 从中获取灵感用于指导机器设计. 同时, 算法层面的创新也是必不可少的. 目的是赋予机器无需明确指导即可自主决定何时及如何更新其认知模型的能力. 此外,

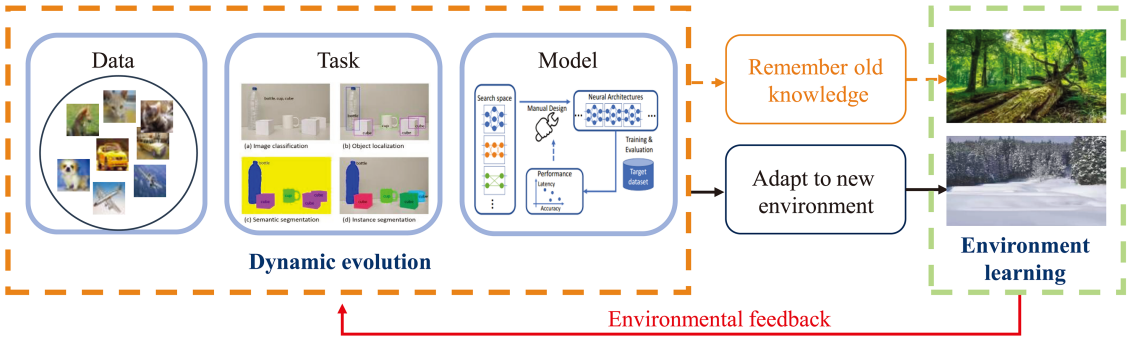

图 5　(网络版彩图) 自主演化学习框架.

Figure 5　(Color online) The framework of self-evolving learning.

构建能够模拟真实世界动态变化的测试平台, 用于验证不同策略下机器自主演化学习的有效性, 将是推进这项技术发展的重要步骤. 通过这样的努力, 我们有望见证更加智能和灵活的系统出现, 它们能够在复杂多变的真实世界环境中独立运行, 而不需要频繁的人工干预.

3　自主机器学习应用

首先, 通过 3 个案例研究来展示如何利用我们提出的自主学习框架来解决实际问题: (1) 自主救援机器人; (2) 无人驾驶; (3) 智能无人机. 其次, 介绍自主机器学习的其他应用.

3.1　自主救援机器人

在自主搜救等机器人学应用中, 对智能体来说, 能够基于已有知识推导出类似的解决方案至关重要, 例如, 根据学到的开门技能来打开窗户. 这要求智能体能够进行类比推理, 包括理解哪些工作 (如操作、运动、导航、装配等) 相似、将执行源工作的动作进行调整以适应与之类似的目标工作等. 现有的类比推理研究 [26,27] 严重依赖于人类对工作之间类比关系知识库和符号系统的人工构建, 从而在相似的工作之间调整动作, 这是耗费人力且昂贵的, 很难随时间的推移而发展, 并且不够鲁棒. 在现有的工作规划方法 [28,29] 中, 给定一个新的任务, 需要由人类专家编写一个在机器人系统上运行的软件程序 (如策略、工作计划、PDDL [30] 描述等) 来执行这项任务, 这既耗时又不可扩展. 已经有一些数据驱动的方法 [31,32] 致力于减少对人类的依赖. 然而, 它们需要大量的由专家提供的标注来进行模型训练, 而使用这些标注很难获得实用的机器人系统.

为了解决现有工作的局限性, 我们可以利用本文所提出的自主机器学习框架来建立类比推理系统, 使自主智能体能够在不严重依赖人类监督的情况下, 掌握范围更广的工作. 给定相似工作的一个大集合, 一旦智能体学会了解决其中的任何一个, 我们的系统就能够让其自动找出解决剩余问题的方法. 这将会使得自主智能体更具适应性、自主性、鲁棒性和智能性. 具体而言, 我们的解决方案通过将之前未见过的工作与曾见过的工作进行类比推理, 从而自动合成一个正确且高效的程序 (例如, 用规划域定义语言 [30] 编写的应用域定义, 其规定了完成工作所需的动作序列及其前提条件和效果) 来执行该工作. 例如, 可以通过自动构造类比工作图、程序 – 工作关联、程序合成等方式进行自主任务选择, 通过图 – 图 (graph-to-graph, G2G) 构建模型、跨模态散列模型、程序 – 程序 (program-to-program, P2P) 合成模型进行自主模型选择, 通过深度神经网络为每个训练样本学习权重, 从而决定哪些数据样本应该被选中用于训练进而实现自主数据选择. 在上述任务中, 其中的一些拥有可微的目标函数, 而另一些则没有. 对于可微的目标函数, 可以使用基于梯度的策略来优化; 对于不可微的那些, 可以向强化学习算法寻求帮助. 可选的优化器包括随机梯度下降、Adam、AdaGrad、RMSProp、策略梯度、深度确定性策略梯度、参与者 – 评价者算法、异步优势参与者 – 评价者算法、可信域策略优化、近端策略优化

和 Q 学习. 上述每个模型都可以从多个角度来被评估. 可选的评估策略包括精确率、召回率、F1 值、受试者工作特征曲线 (receiver operating characteristic curve, ROC) 下面积、准确率、双语评估替补 (bilingual evaluation understudy, BLEU)、美国国家标准与技术研究院 (National Institute of Standards and Technology, NIST) 困惑度等.

3.2　无人驾驶

在无人驾驶领域, 智能体需要能够根据已有的驾驶经验推导出应对新动态开放环境和未知任务的驾驶策略. 例如, 智能体可以通过观察人类驾驶员在复杂路况下的行为, 学习如何应对不同的道路环境、交通状况和紧急情况. 这要求无人驾驶系统不仅具备从示范中学习的能力, 还需要具备类比推理能力, 能够识别哪些驾驶任务和环境相似, 进而能够根据新的场景调整驾驶行为. 然而, 当前的无人驾驶系统往往依赖于大量手工设计的规则、模型和人工标注的数据, 这种方式虽然能够保证系统的安全性, 但在应对新的、未知的复杂环境时, 系统的适应性和扩展性受到限制.

为了解决这一问题, 本文提出了一种基于自主机器学习的模仿推理框架, 使得无人驾驶系统能够在较少人工干预的情况下, 基于观察和模仿人类驾驶员的行为, 通过自主学习和推理来适应各种新环境和新任务. 例如, 智能体可以根据在山区道路上行驶的经验, 推导出如何应对更复杂的城市交叉口驾驶. 每当系统学会了一个新场景的驾驶策略后, 它就能够自动推导出与其他类似场景的驾驶方案. 具体而言, 本文的解决方案包括以下几个关键步骤. 首先, 系统通过自动构建类比图谱进行自主任务选择, 识别不同任务和场景之间的相似性, 并将不同场景下的驾驶任务与相应的控制程序和决策模型关联. 然后, 通过自动机器学习方法进行自主模型选择, 自动生成适用于新驾驶任务的策略. 此外, 系统还通过深度神经网络自适应地为每个训练样本分配权重, 从而实现自主数据选择, 确保系统能够在面对未知环境时, 选择最合适的训练数据来进行学习. 为了解决任务中的优化问题, 系统针对不同目标函数使用了多种优化算法, 如随机梯度下降、Adam、策略梯度等. 上述每个模型都可以从多个角度来被评估. 可通过自主评估指标来选择指导模型优化的方向.

基于自主机器学习的无人驾驶优势在于, 它减少了对专家手工标注数据的依赖, 同时提高了无人驾驶系统在面对复杂、多变的道路环境时的自适应能力. 通过模仿推理, 系统能够将之前学到的经验应用到新的驾驶任务中, 从而提升其在未知环境中的鲁棒性和可靠性.

3.3　智能无人机

在智能无人机领域, 智能体能够通过主动探索感知与自主演化学习来应对动态开放环境的需求日益增长. 例如, 智能无人机需要能够在城市、山区、室内等多样化且未知的环境中, 通过主动探索和不断演化的能力, 感知环境变化, 并根据实时反馈进行调整飞行策略. 传统的无人机系统通常依赖于事先规划的飞行路径和固定的规则来执行任务, 这种方式虽然有效, 但对于应对复杂、动态和未知的环境来说, 缺乏足够的灵活性和适应性.

为了克服这一挑战, 本文利用本文所提出的自主机器学习框架来建立探索演化系统, 通过主动探索感知和自主演化的机制, 使得无人机能够在未知环境中实时探索、感知、自我更新并优化其飞行策略. 这一过程不仅包括对环境的感知, 还包括通过自主演化逐步改进无人机的飞行策略, 例如, 不断优化其飞行路径、避障策略和任务执行能力, 使其能够适应更复杂的任务和环境. 具体而言, 本文的解决方案包括以下几个核心步骤. 首先, 给定多个任务, 智能无人机通过自主任务选择, 识别任务间的相似性, 例如, 通过解耦学习构建任务的原型空间或采用元学习架构, 根据已有任务的解决方案, 推导出适应新任务的解决方案; 在任务执行过程中, 智能无人机需要动态选择合适的机器学习模型来解决不同的问题 (如飞行控制、路径规划等); 智能体通过深度神经网络评估训练样本的有效性, 例如, 利用课程学习、知识蒸馏等进行数据重加权, 基于自主数据选择智能地挑选出对任务执行有益的数据进行训练; 为了不断提升智能无人机的飞行能力和任务执行效率, 系统通过自主优化策略选择, 自动选择最合适

的优化方法进行模型参数的调整和学习; 为了衡量任务执行的成功率和系统的性能, 智能无人机通过自主评估指标选择来评估模型和任务执行结果.

该系统能够根据环境和任务的变化, 动态调整策略, 提升任务执行效率, 同时减少对人工干预的依赖, 增强无人机在多种飞行场景中的适应性和智能性. 智能无人机不仅能完成常规任务, 还能够应对突发事件和新的任务需求, 如紧急救援、灾难响应等, 从而提升其在复杂环境中的应用潜力.

3.4　其他应用

除了上述的 3 个案例研究之外, 我们的自主机器学习框架还可以被广泛应用于自然语言处理 (natural language process, NLP)、计算机视觉 (computer vision, CV)、数据挖掘和多媒体等领域. 下面给出一些例子.

• 基于常识的可控故事写作, 即在给定情感、故事线等控制因素下自动编写故事. 为了生成有意义且信息丰富的故事, 需要故事写作系统具备自主探索感知数据环境的能力, 通过动态自我优化, 主动选择与任务最相关的常识知识源, 并实时适配模型结构和检索策略, 以高效融入外部常识. 例如, 给定 “晴天运动” 的故事线, 系统需自主优化知识检索, 理解 “远足是运动且适合晴天” 等常识, 才能生成如 “昨天晴天, 我们想运动. 加州流行远足, 这天气正合适, 所以我们去了” 的优质故事. 开发此系统面临的核心挑战在于如何实现: (1) 系统自主演化其知识库, 通过分析反馈和探索新来源持续更新常识图谱; (2) 系统自主优化数据选择机制, 在故事生成时高效精准调用相关知识; (3) 系统自主优化学习策略, 在有限标注故事语料下, 通过合成数据或迁移学习等自扩展方式提升性能. 我们利用本文所提出的自主机器学习框架, 通过对任务、数据、模型的自主选择和知识的自主演化可以被用来应对上述挑战.

• 可控视频描述生成致力于根据辅助信息的引导来控制视频描述的生成过程, 如文本标签 (浪漫的、幽默的) [33]、词性 (part-of-speech, POS) 标签序列 [34] 或一个例句 [35]. 最具挑战性的是利用例句, 即生成与例句拥有相同句法结构的对应视频描述. 例如, 当真实的描述是 “一群人在跳舞”, 例句是 “一串绿香蕉在香蕉树前面挂着” 时, 模型可能会输出 “一群年轻人在现场观众前面跳舞”. 因此, 模型需要提取例句的句法结构并将其合理地融入到描述生成的过程中, 并避免被额外的含噪声的语义信息所干扰. 其难点可以被概括为以下几个方面. (1) 如何从有限的例句中通过自主数据选择高效地提取语义信息, 以便进一步生成描述. (2) 如何在例句中存在干扰的情况下, 通过自主演化保持所生成描述的视频语义. 因此, 自主机器学习框架可以被用来应对上述挑战.

• 能够感知语义的聊天机器人. 在开放世界对话系统中, 为了给出有信息量的、正确且有用的回答, 理解对话历史记录的语义并对语义进行推理是十分重要的. 例如, 给定一句来自人类的话 “我想喝点咖啡. 咖啡店有多远?” 在不理解这句话语义的情况下, 聊天机器人往往会给出一个不含信息量且枯燥的回答, 如 “听起来很酷”. 相反, 如果使用语义解析器来将 “咖啡店有多远?” 的查询解析为表示查询语义的逻辑形式, 并与外部知识库一起对逻辑形式进行推理, 聊天机器人就能够给出一个有用且有信息量的回答, 如 “大约有半英里远.” 为了开发出能够感知语义的聊天机器人, 有几个技术性挑战需要被解决. (1) 给定有限的经过标注的 (话语, 逻辑形式) 对, 如何通过数据自主优化训练出高准确度的语义解析器? (2) 给定有限的 (逻辑形式, 回答) 对, 如何通过模型自主优化训练出能够感知语义并且能够对抗过拟合的回复生成模型? (3) 如何通过模型自主优化训练出用来推断深层语义的推理系统? 利用自主机器学习进行任务、数据、模型自主选择的能力, 可以应对前面上述挑战.

• 视频对话. 该任务也被称为视听场景感知对话 (audio-visual scene-aware dialog, AVSD) [36,37] 任务, 它要求智能体以对话框的形式, 用自然的对话语言与人类进行关于视频和音频内容的对话. 在被应用于儿童早教等日常使用场景的真实世界对话系统中, 理解对话历史的语义以及视频、音频和文字的多模态表征很重要, 只有这样才能给出具有高度相关性和足够信息量的回答. 例如, 给定一段包含音频的视频, 视频中一个人路过一个包并留下一本书, 然后用自然的人类语言给出问题 “她走得快还是慢?”, 在不理解这个视频的情况下, 回答可能是随机的 “快” 或 “慢”. 但一个能够完全理解视频内容的

智能体则可以给出 “在放下书之前, 她相当慢地来回走着.” 的回答, 这显然在更委婉的描述中提供了更多的信息 [37,38]. 为了开发出视听场景感知对话智能体, 有几个技术性挑战需要被解决. (1) 如何通过自我优化机制训练出高准确度的语义解析器? 因为所有的问题和回答都通过文本形式给出, 准确地理解问题成为关键的部分. (2) 如何通过自主演化机制训练出能够融合和理解包括动态场景、音频和对话历史 (之前轮次的对话) 等多模态信息的系统? 利用自主机器学习可以轻松应对上述挑战.

4　结论及未来方向

从人类的自主学习中汲取灵感, 我们引出了自主机器学习的基础概念, 并提出了一个它的框架. 自主机器学习的目标是实现高度自主, 包括自主任务选择、自主数据选择、自主模型选择, 而不需要人类手动进行这些选择, 同时根据环境反馈和自我状态通过动态自主演化来提升自身性能. 自主机器学习框架由自我优化和自主演化两个关键过程组成, 我们提出了基于多级优化和强化反馈的框架来表述自主机器学习. 在包括自主救援机器人、无人驾驶和智能无人机在内的研究中, 我们展示了如何利用自主机器学习去自主地解决复杂的实际问题.

自主机器学习与自动机器学习 (AutoML) 及强化学习存在本质区别. 自动机器学习的核心是自动化, 它聚焦于使用算法自动执行传统机器学习流程 (如超参调优、特征工程、模型选择), 目标是减少人工干预, 但其优化目标和流程通常是预先设定且相对静态的. 强化学习的核心是序列决策, 智能体通过与环境交互、基于奖励信号学习最优策略以完成特定任务, 其学习是目标驱动的, 但环境模型和奖励函数通常由外部设计. 相比之下, 自主机器学习的核心是系统自身的 “自主性” 与 “演化性”. 它不仅强调对数据、模型、任务等核心学习过程的自我优化, 更关键的是能根据环境反馈和自我状态, 动态自主演化其学习架构、知识库或能力边界 (如自动发现新任务、定义新目标、合成数据、更新知识结构), 表现出更强的内在驱动力、环境适应性和持续演进能力, 旨在构建能自我指导、自我提升的学习系统.

值得注意的是, 我们提出的自主机器学习框架与特定的机器学习应用无关, 并且能够被广泛应用于改进各种机器学习任务, 包括但不限于: 分类、回归、聚类、文本生成、对话系统、机器翻译、文档摘要、目标检测、语义分割、视觉问答、时间序列预测以及图的链接和节点预测等.

为实现自主机器学习, 仍需攻克以下核心难题. (1) 环境学习与环境意识构建: 如何通过传感器感知, 识别环境中的功能性对象、对象间关系、对象所属场景的类别及状态, 监测环境的动态变化, 动态建模客观环境并赋予环境 “可操作性” 意义 (环境意识), 并基于环境意识预测环境变化趋势, 支撑演化决策. (2) 自我意识构建及演化: 如何通过元认知模块量化自身能力边界, 建立动态能力图谱, 通过与环境意识融合, 实现记忆驱动的意识自主演化. (3) 数据 – 模型 – 任务自主演化: 针对动态开放环境, 如何自主扩展数据侧知识库, 实时重构模型侧架构, 自主定义任务侧新子目标, 实现数据 – 模型 – 任务自主演化. (4) 安全可信的应用部署: 如何无监督对齐法律法规和社会道德, 设计演化行为边界检测, 阻断高风险自主决策.

对于未来工作, 我们认为以下研究方向十分值得探索.

可解释自主机器学习. 可靠性几乎成为了机器学习模型能够被人类愿意使用的必要条件. 为了获得人类的信任, 我们将开发可解释的自主机器学习方法, 它能够生成可解释且透明的预测. 之前大多数可解释机器学习的关注点都是从输入数据 (如文本中的短语和图像中的区域) 中找出与作出预测最相关的关键证据, 然后用这些证据来证明预测的意义. 然而, 在很多情况下, 归因出的证据对人类来说没有意义. 根本原因是机器学习的推理过程与人类的推理过程并不一致, 尽管它们得出了相同的预测结果. 为了解决这个问题, 可以尝试使用自然语言处理的方法来分析文本, 并从中自动提取人类的决策过程, 然后将这些结构化过程作为归纳偏差输入到自主机器学习框架中, 以达到人机同步的效果. 此外, 与认知科学家合作, 深入理解人类如何解释现象和作出决策的基本机制, 并使用这些机制来指导自主机器学习框架的设计也是一个值得尝试的研究内容.

鲁棒自主机器学习. 在医疗保健、金融等许多应用场景中, 决策是至关重要的. 机器学习辅助决策支持软件需要具备安全性和鲁棒性以抵御恶意攻击. 现有的临床机器学习模型被证明容易受到对抗样本的影响. 例如, 给定一张被卷积神经网络预测含有肺炎的胸部 X 光片, 为其添加 (无法被人类察觉的) 微小扰动, 就会让模型认为该图片不含肺炎. 大多数之前的防御方法都是针对特定攻击高度定制的, 因此它们很容易在攻击发生改变时失效. 如何实现能够以统一形式表示攻击并进行防御的自主机器学习框架, 并相应地设计出能够应对各种形式的攻击并对攻击变化具备鲁棒性的防御技术是一个值得深挖的研究课题. 此外, 作为一个长期目标, 考虑与密码学家合作, 开发针对机器学习的同态加密 (homomorphic encryption, HE) 算法, 使得自主机器学习能够在密文上进行训练与推断也会是一个有前景的研究方向.

可信自主机器学习. 由于数据隐私问题和法律法规约束, 可信自主机器学习研究核心挑战在于如何保证能力可信和行为可信. 传统小样本学习、元学习、迁移学习虽能缓解数据不足, 却缺乏神经符号推理对高阶变量关系的解析能力, 如图神经网络驱动的逻辑规则自动发现与因果推理. 如何通过自主机器学习突破统计学习的局限性, 指导模型在有限数据下自主优化实现能力可信是一个值得探索的方向. 法律法规的约束进一步强化 AI 行为可信, 首先, 需要人类价值观建模, 借鉴宪法人工智能 (constitutional AI) 思路, 将伦理准则转化为可计算的符号约束, 动态对齐自主决策与人类价值; 其次, 通过安全推理验证, 构建多级验证架构, 通过形式化方法或可解释性模块对推理结果进行实时安全检测, 例如, 在医疗诊断中验证因果链是否符合医学伦理. 如何通过自主机器学习实现 AI 行为可信也是一个值得探索的研究方向.

意识驱动的自主机器学习. 在教育科学中, 自主学习将学习者视为其自身学习过程的负责人和管理者 [15]. 内部意识和外部意识被视为改善和提升学习者对自身和他人看法的两个重要组成部分. 从这些事实中汲取灵感, 我们认为将自主机器学习设计为一个由自我意识指导的自主学习过程, 其中自我意识包括内部意识和外部意识, 是一个可行的研究方案. 基于该方案, 在学习过程中, 自主机器学习以自主的方式来选择学习任务、数据、模型、优化算法和评估指标, 学习结果又反过来为如何改进自我意识提供反馈. 这有助于为人工智能的研究奠定理论基础.

自主具身智能. 传统的具身智能系统往往需要人工干预来设置学习任务、选择学习数据和调节模型架构, 而自主机器学习通过智能体主动探索感知动态开放环境, 基于自我优化和自主演化, 实现环境学习与环境 – 自我具身, 构建具有自主演化能力的持续自适应智能. 与之对比, 大语言模型 (LLM) 驱动的自主智能体在方法论上主要依赖预训练的海量知识库和符号推理能力, 通过提示工程或微调实现任务执行. 在任务泛化性方面, 自主机器学习可以通过与环境持续交互, 基于自主演化实现全域泛化, 而 LLM 智能体由于体型庞大, 其可以实现任务逻辑推理, 但难以应对高动态开放环境下的全域具身任务执行. LLM 可以作为自主机器学习的知识库, 通过 LLM 高层任务推理, 提升自主机器学习应对复杂、动态和不确定现实世界的自主演化能力, 增强复杂场景下的自主决策鲁棒性, 这是一个值得研究的方向.

参考文献

1 Grira N, Crucianu M, Boujemaa N. Unsupervised and semi-supervised clustering: a brief survey. Rev Mach Learn Tech Process Multi Content, 2004, 1: 9–16

2 Jaiswal A, Babu A R, Zadeh M Z, et al. A survey on contrastive self-supervised learning. Technologies, 2020, 9: 2

3 Liu X, Zhang F, Hou Z, et al. Self-supervised learning: generative or contrastive. 2020. ArXiv:2006.08218

4 Qi G J, Luo J. Small data challenges in big data era: a survey of recent progress on unsupervised and semi-supervised methods. IEEE Trans Pattern Anal Mach Intell, 2022, 44: 2168–2187

5 Schmarje L, Santarossa M, Schroder S M, et al. A survey on semi-, self- and unsupervised learning for image classification. IEEE Access, 2020, 9: 82146–82168

6 Kumar P, Gupta A. Active learning query strategies for classification, regression, and clustering: a survey. J Comput

Sci Technol, 2020, 35: 913–945
7 Ren P, Xiao Y, Chang X, et al. A survey of deep active learning. ACM Comput Surv, 2022, 54: 1–40
8 Wang M, Hua X S. Active learning in multimedia annotation and retrieval. ACM Trans Intell Syst Technol, 2011, 2: 1–21
9 He X, Zhao K, Chu X. AutoML: a survey of the state-of-the-art. Knowledge-Based Syst, 2021, 212: 106622
10 Yao Q, Wang M, Chen Y, et al. Taking human out of learning applications: a survey on automated machine learning. 2018. ArXiv:1810.13306
11 Zöller M A, Huber M F. Survey on automated machine learning. 2019. ArXiv:1904.12054
12 Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adaptation of deep networks. In: Proceedings of the International Conference on Machine Learning, Sydney, 2017. 1126–1135
13 Hospedales T, Antoniou A, Micaelli P, et al. Meta-learning in neural networks: a survey. 2020. ArXiv:2004.05439
14 Vanschoren J. Meta-learning: a survey. 2018. ArXiv:1810.03548
15 Caffarella R S. Self-directed learning. New Direct Adult Contin Ed, 1993, 57: 25–35
16 Garrison D R. Self-directed learning: toward a comprehensive model. Adult Education Q, 1997, 48: 18–33
17 Brookfield S. Self-directed learning, political clarity, and the critical practice of adult education. Adult Education Q, 1993, 43: 227–242
18 Loeng S. Self-directed learning: a core concept in adult education. Education Res Int, 2020, 2020: 1–12
19 LaTour K A, Noel H N. Self-directed learning online: an opportunity to binge. J Market Ed, 2021, 43: 174–188
20 Song L, Hill Y R. A conceptual model for understanding self-directed learning in online environments. J Interact Online Learn, 2007, 6: 27–42
21 Knowles M S. Self-Directed Learning: a Guide for Learners and Teachers. New York: Association Press, 1975
22 Zhang Z, Wang X, Guan C, et al. AutoGT: automated graph transformer architecture search. In: Proceedings of the ICLR, 2023
23 Chen H, Wang X, Zhang Z, et al. AutoGFM: automated graph foundation model with adaptive architecture customization. In: Proceedings of the ICML, 2025
24 Chen H, Wang X, Guan C, et al. Auxiliary learning with joint task and data scheduling. In: Proceedings of Machine Learning Research, 2022. 3634–3647
25 Chen H, Wang X, Zhou Y, et al. Joint data-task generation for auxiliary learning. In: Proceedings of the Advances in Neural Information Processing Systems 36 (NeurIPS 2023), 2023
26 Gentner D, Smith L, Ramachandran V S. Analogical reasoning. In: Encyclopedia of Human Behavior. 2nd ed. Cambridge: Academic Press, 2012, 130–136
27 Sunstein C R. On analogical reasoning. Harvard Law Rev, 1993, 106: 741–791
28 Galindo C, Fernández-Madrigal J A, González J, et al. Robot task planning using semantic maps. Robotics Auto Syst, 2008, 56: 955–966
29 Cambon S, Alami R, Gravot F. A hybrid approach to intricate motion, manipulation and task planning. Int J Robotics Res, 2009, 28: 104–126
30 Aeronautiques C, Howe A, Knoblock C, et al. Pddl—the planning domain definition language. Technical Report CVC TR-98-003/DCS TR-1165. New Haven: Yale Center for Computational Vision and Control, 1998
31 Dantam N T, Kingston Z K, Chaudhuri S, et al. An incremental constraint-based framework for task and motion planning. Int J Robotics Res, 2018, 37: 1134–1151
32 Grover S, Sengupta S, Chakraborti T, et al. RADAR: automated task planning for proactive decision support. Hum-Comput Interact, 2020, 35: 387–412
33 Gan C, Gan Z, He X, et al. Stylenet: generating attractive visual captions with styles. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017. 3137–3146
34 Deshpande A, Aneja J, Wang L, et al. Fast, diverse and accurate image captioning guided by part-of-speech. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019. 10695–10704
35 Yuan Y, Ma L, Wang J, et al. Controllable video captioning with an exemplar sentence. In: Proceedings of the 28th ACM International Conference on Multimedia, 2020. 1085–1093
36 Das A, Kottur S, Gupta K, et al. Visual dialog. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017. 326–335
37 Alamri H, Cartillier V, Das A, et al. Audio visual scene-aware dialog. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019. 7558–7567
38 Liu J, Chen W, Cheng Y, et al. Violin: a large-scale dataset for video-and-language inference. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020. 10900–10910

Autonomous machine learning

Wenwu ZHU[1], Xin WANG[1*] & Zongben XU[2]

1. *Department of Computer Science and Technology, Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China*
2. *School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China*
* Corresponding author. E-mail: xin_wang@tsinghua.edu.cn

Abstract Conventional machine learning (ML) is capable of training on data and constructing models for given tasks, enabling these models to acquire predictive and decision-making abilities for the corresponding tasks. This can be summarized as task-driven ML. Existing methods rely heavily on external human guidance and empirical specification of data and tasks, model configuration, and parameter learning during the learning process. In real world, dynamic, and open environments, these methods cannot learn autonomously as humans do. This paper introduces a new concept of autonomous ML. Specifically, we define autonomous ML as a self-driven, dynamically self-evolving learning process that encompasses self-optimization and self-evolution. Firstly, it can actively explore and perceive the environment for autonomous data selection, autonomous model adaptation, and task switching without human intervention. Simultaneously, autonomous ML is able to dynamically self-evolve based on feedback from environmental learning and the current machine state. Additionally, the paper presents several application case studies of autonomous ML and discusses future research directions. We anticipate that autonomous ML can enable machines to learn autonomously like humans and provide a new perspective for general artificial intelligence.

Keywords autonomous machine learning, self-optimization, self-evolution