终身机器学习

Lifelong Machine Learning

1 绪论

主讲：梁国强
gqliang@nwpu.edu.cn

Outlines

1.1 引言

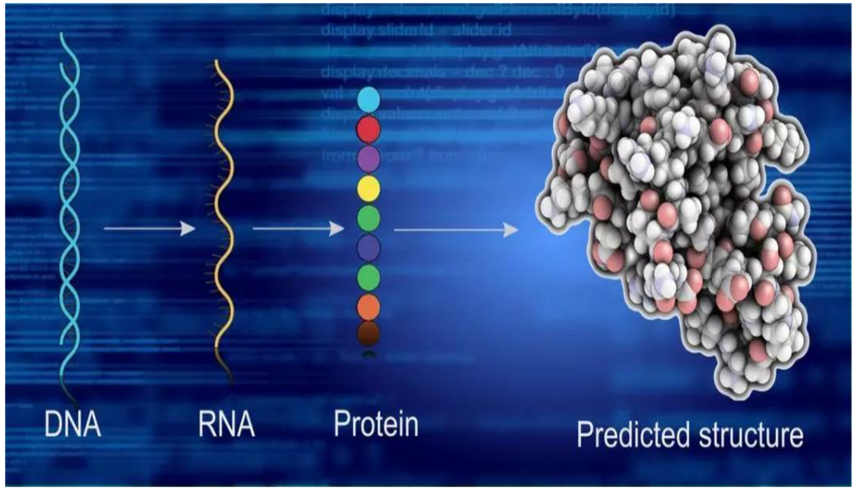
1.2 定义

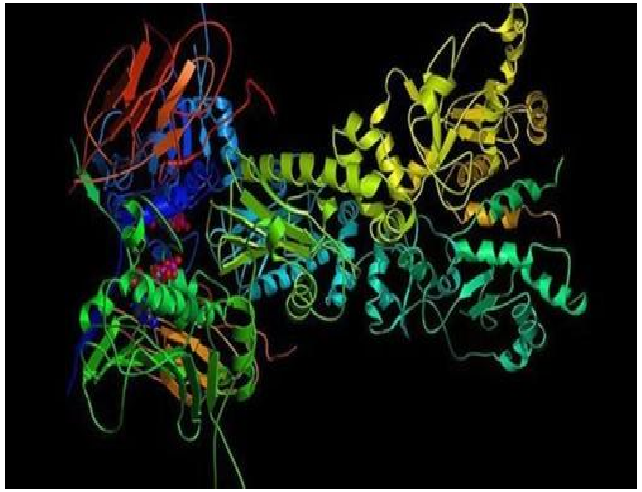
1.3 关键挑战

1.1 引言

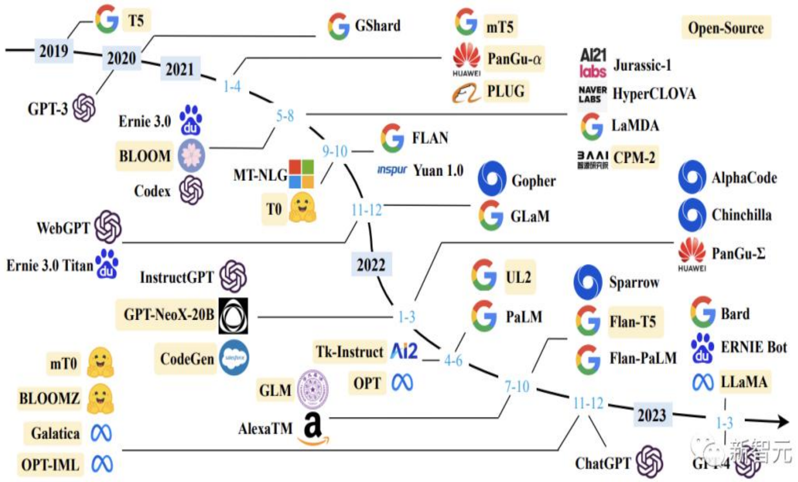
机器学习发展训练，尤其是大模型出现以来

AlphaZero

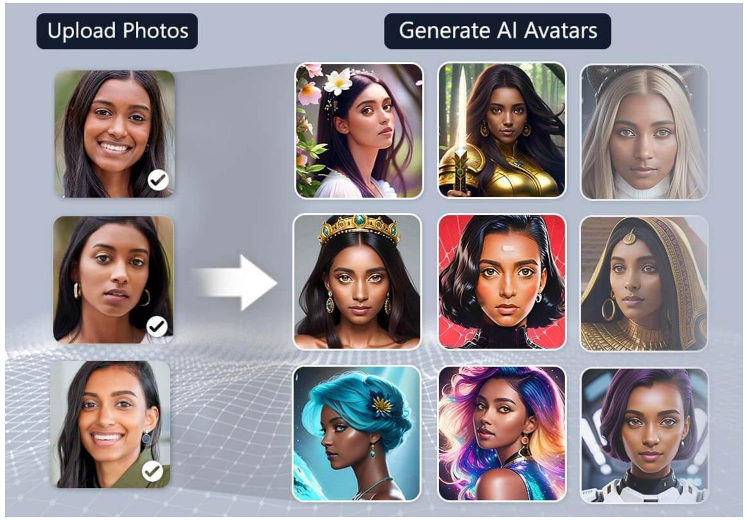


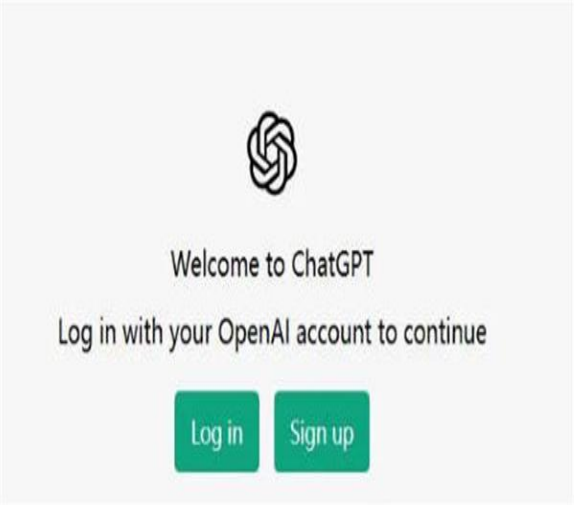


AlphaFold预测蛋白质结构——《自然》评选的十大科学突破之首



大模型

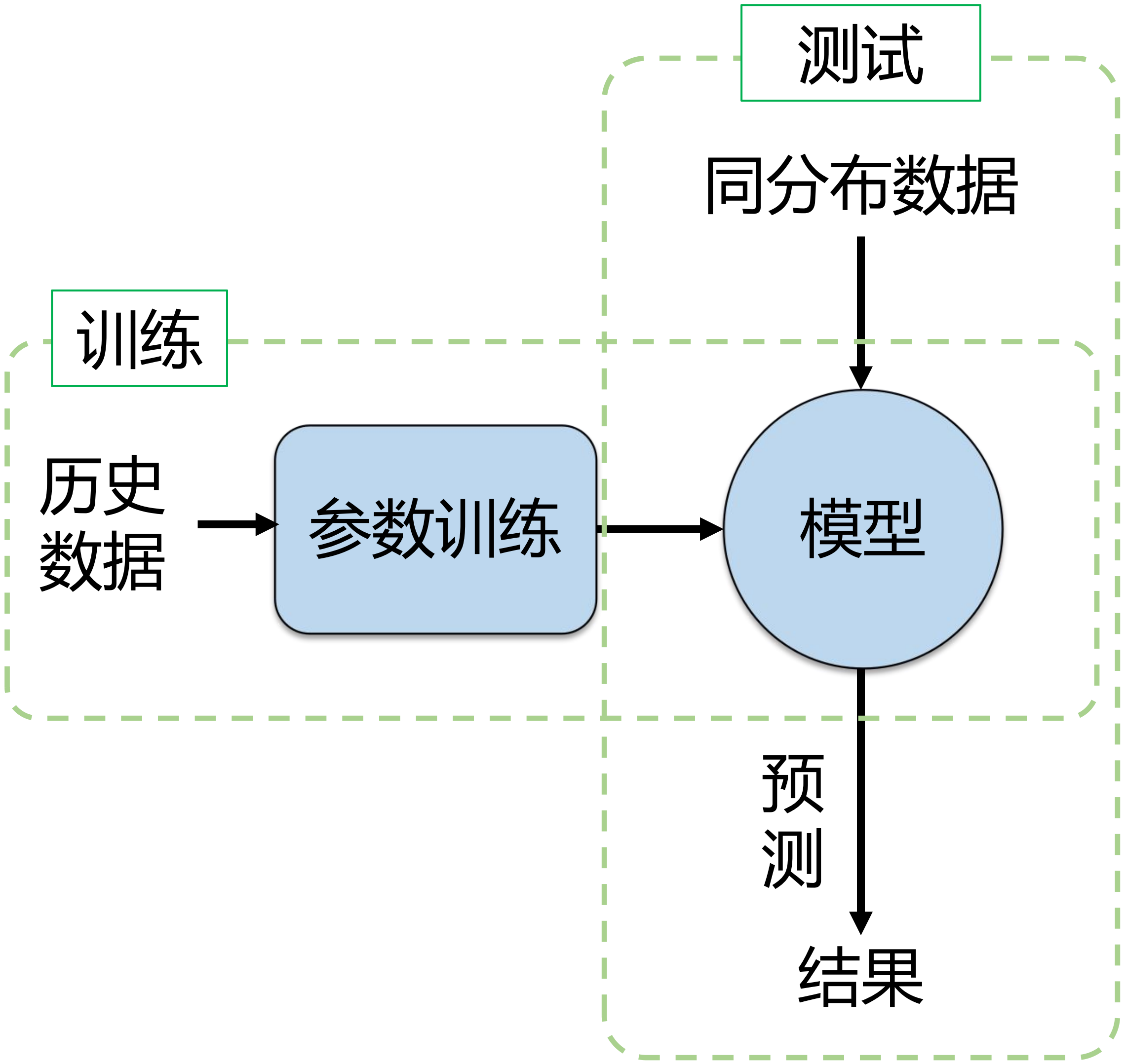




AIGC

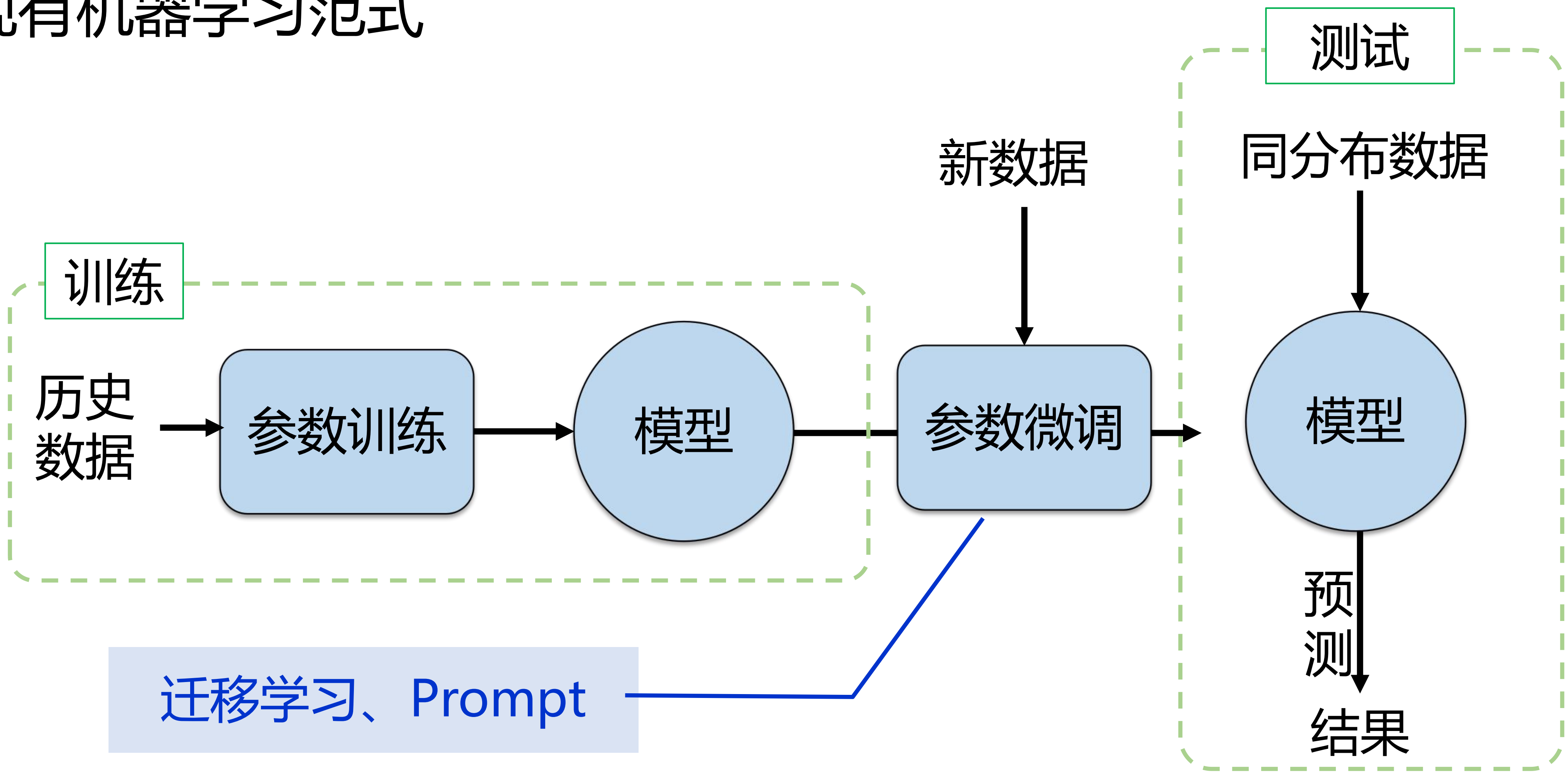
1.1 引言

现有机器学习范式



1.1 引言

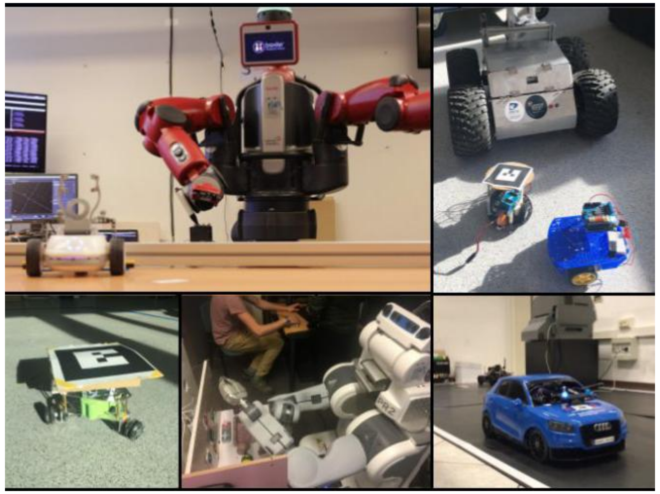
现有机器学习范式

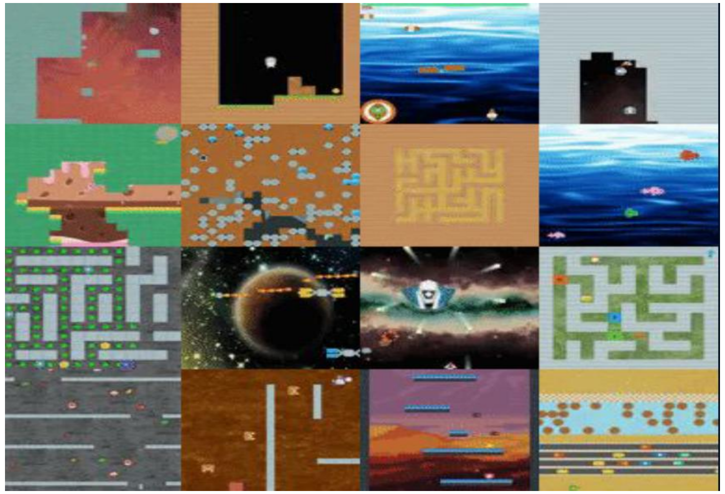


孤立学习范式存在很多不足

1.1 引言

实际AI面临挑战





- 开放复杂的环境

  高速、城市、乡村

- 多种任务

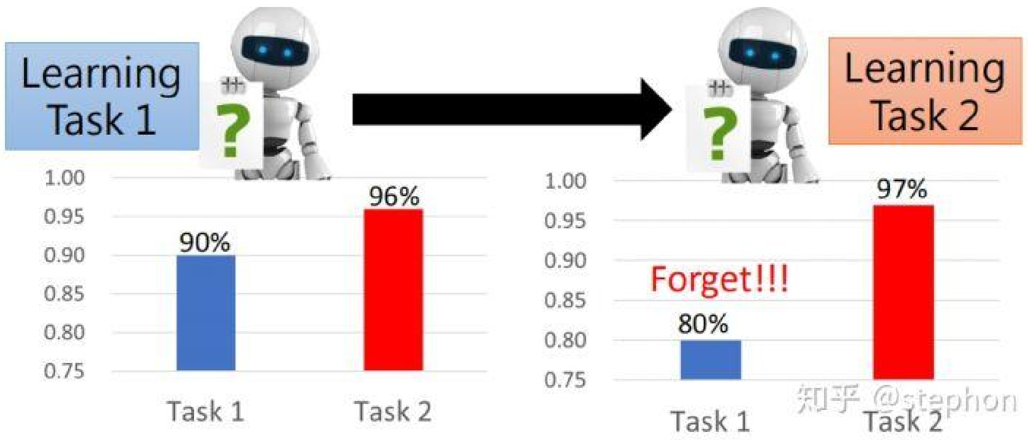
  识别、检测、抓取

- 不同用户

  孩子、老人、成人

1.1 引言

孤立学习范式

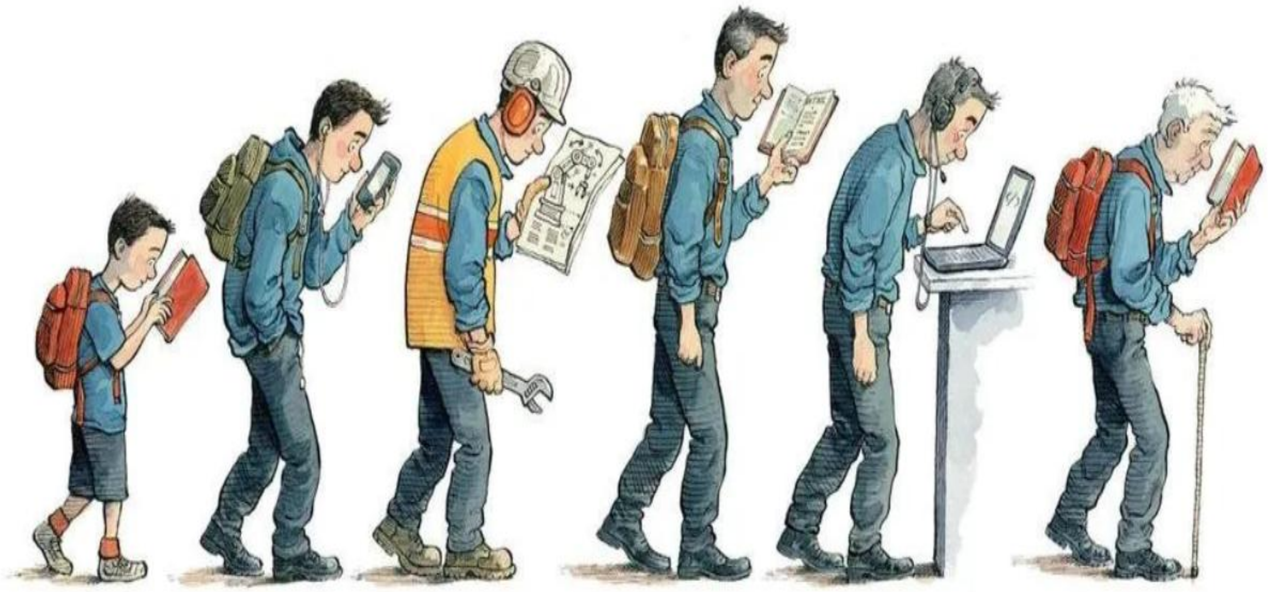


- 独立学习每个任务
- 很少存在积累、使用知识
- 训练需要大量标记数据，计算成本高
- 部署后较少发生模型参数变化
- 学习新任务后，旧任务性能显著降低

1.1 引言

人类智能

吾生也有涯　而知也无涯



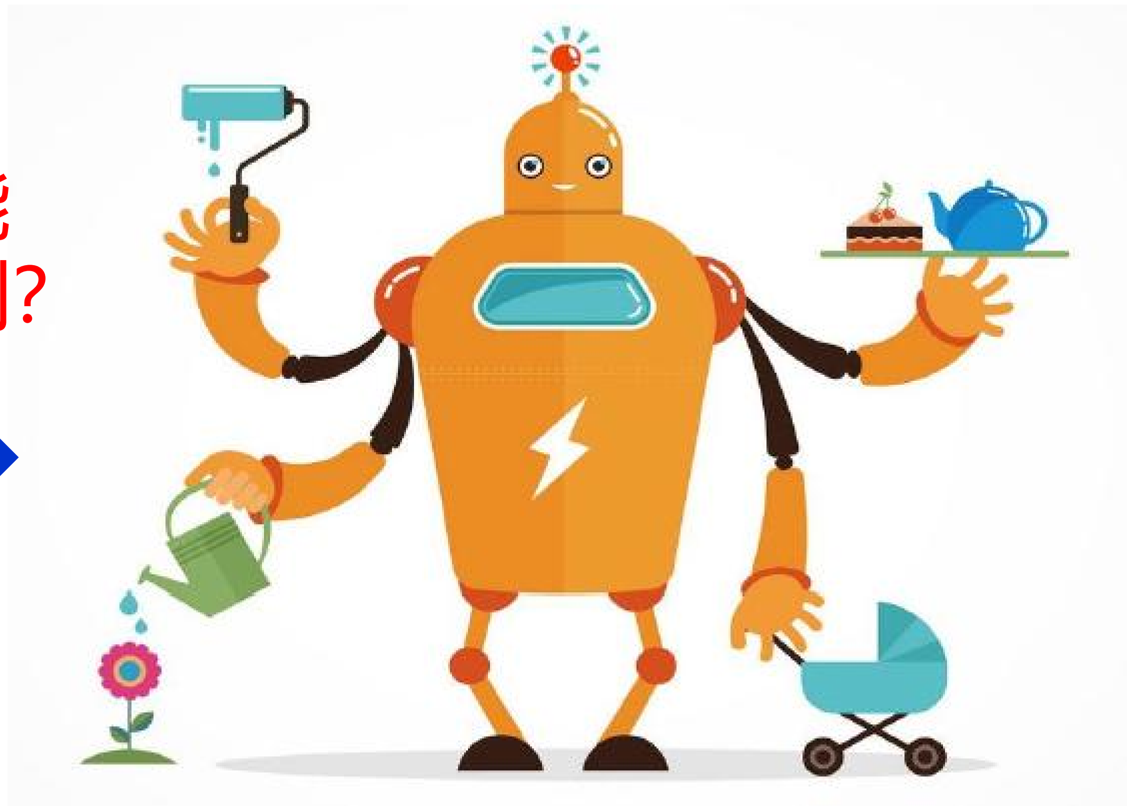
终身学习能力

各种能力切换自如

1.1 引言

人类智能

人工智能
如何做到?



终身学习旨在模拟人类学习过程和学习能力

1.1 引言

终身机器学习

nature machine intelligence

Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

nature > nature machine intelligence > perspectives > article

Perspective | Published: 23 March 2022

Biological underpinnings for lifelong learning machines

Article | Open Access | Published: 05 December 2022

Three types of incremental learning

Gido M. van de Ven, Tinne Tuytelaars & Andreas S. Tolias

Nature Machine Intelligence **4**, 1185–1197 (2022) | Cite this article

2118 Accesses | **13** Altmetric | Metrics

nature > nature communications > articles > article

Article | Open access | Published: 13 August 2020

Brain-inspired replay for continual learning with artificial neural networks

Gido M. van de Ven, Hava T. Siegelmann & Andreas S. Tolias

DARPA DEFENSE ADVANCED RESEARCH PROJECTS AGENCY

EXPLORE BY TAG

ABOUT US / OUR RESEARCH / NEWS / EVENTS / WORK WITH US /

› Defense Advanced Research Projects Agency › Our Research › Lifelong Learning Machines

Lifelong Learning Machines (L2M) (Archived)

Artificial intelligence (AI) and machine learning (ML) systems have advanced significantly in recent years. Despite a wide range of impressive results, current AI is not intelligent in the biological sense. These systems are limited to performing only those tasks for which they have been specifically programmed and trained, and are inherently subject to safety hazards when encountering situations outside them. The issue is further limiting to DoD applications, where situations can be unpredictable and the ability to react quickly and adapt to dynamic circumstances is of primary importance.

The Lifelong Learning Machines (L2M) program seeks to achieve paradigm-changing developments in AI architectures and ML techniques. The program seeks to develop systems that can learn continuously during execution and become increasingly expert while performing tasks, are subject to safety limits, and apply previous skills and knowledge to new situations - without forgetting previous learning.

L2M consists of two technical areas. The first concentrates on the development of complete systems and their components; the second brings together researchers with diverse expertise to explore biological mechanisms that underlie learning, which will be translated into a new generation of computational architectures, mechanisms, and algorithms. Discoveries in both technical areas are expected to generate new methodologies that will allow AI systems to learn and improve during tasks, apply previous skills and knowledge to new situations, incorporate innate system limits, and enhance safety in automated assignments.

美国DAPRA终身机器学习项目

专栏1 基础理论

1.大数据智能理论。研究数据驱动与知识引导相结合的人工智能新方法、以自然语言理解和图像图形为核心的认知计算理论和方法、综合深度推理与创意人工智能理论与方法、非完全信息下智能决策基础理论与框架、数据驱动的通用人工智能数学模型与理论等。

2.跨媒体感知计算理论。研究超越人类视觉能力的感知获取、面向真实世界的主动视觉感知及计算、自然声学场景的听知觉感知及计算、自然交互环境的言语感知及计算、面向异步序列的类人感知及计算、面向媒体智能感知的自主学习、城市全维度智能感知推理引擎。

3.混合增强智能理论。研究"人在回路"的混合增强智能、人机智能共生的行为增强与脑机协同、机器直觉推理与因果模型、联想记忆模型与知识演化方法、复杂数据和任务的混合增强智能学习方法、云机器人协同计算方法、真实世界环境下的情境理解及人机群组协同。

4.群体智能理论。研究群体智能结构理论与组织方法、群体智能激励机制与涌现机理、群体智能学习理论与方法、群体智能通用计算范式与模型。

5.自主协同控制与优化决策理论。研究面向自主无人系统的协同感知与交互，面向自主无人系统的协同控制与优化决策，知识驱动的人机物三元协同与互操作等理论。

6.高级机器学习理论。研究统计学习基础理论、不确定性推理与决策、分布式学习与交互、隐私保护学习、小样本学习、深度强化学习、无监督学习、半监督学习、主动学习等学习理论和高效模型。

7.类脑智能计算理论。研究类脑感知、类脑学习、类脑记忆机制与计算融合、类脑复杂系统、类脑控制等理论与方法。

8.量子智能计算理论。探索脑认知的量子模式与内在机制，研究高效的量子智能模型和算法、高性能高比特的量子人工智能处理器、可与外界环境交互信息的实时量子人工智能系统等。

中国新一代人工智能规划

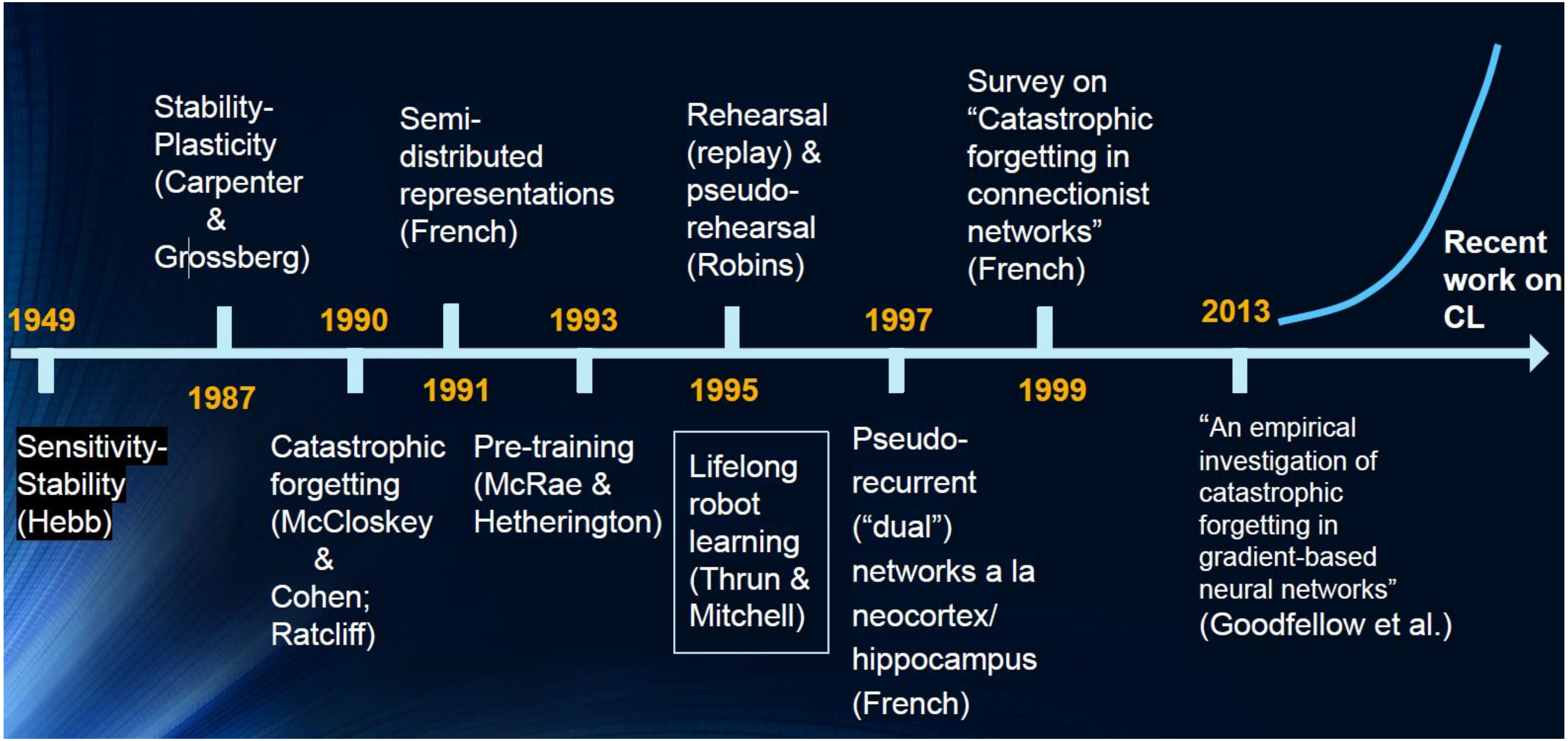
1.1 引言

2023

- Hierarchical Decomposition of Prompt-Based Continual Learning: Rethinking Obscured Sub-optimality (NeurIPS 2023)[paper]
- FeCAM: Exploiting the Heterogeneity of Class Distributions in Exemplar-Free Continual Learning (NeurIPS 2023)[paper]
- Continual Learners are Incremental Model Generalizers (ICML2023)[paper]
- Learnability and Algorithm for Continual Learning (ICML2023)[paper][code]
- Parameter-Level Soft-Masking for Continual Learning (ICML2023)[paper]
- Continual Learning in Linear Classification on Separable Data (ICML2023)[paper]
- DualHSIC: HSIC-Bottleneck and Alignment for Continual Learning (ICML2023)[paper]
- BiRT: Bio-inspired Replay in Vision Transformers for Continual Learning (ICML2023)[paper]
- DDGR: Continual Learning with Deep Diffusion-based Generative Replay (ICML2023)[paper]
- Neuro-Symbolic Continual Learning: Knowledge, Reasoning Shortcuts and Concept Rehearsal (ICML2023)[paper]
- Theory on Forgetting and Generalization of Continual Learning (ICML2023)[paper]
- Poisoning Generative Replay in Continual Learning to Promote Forgetting (ICML2023)[paper]
- Continual Vision-Language Representation Learning with Off-Diagonal Information (ICML2023)[paper]
- Prototype-Sample Relation Distillation: Towards Replay-Free Continual Learning (ICML2023)[paper]
- Does Continual Learning Equally Forget All Parameters? (ICML2023)[paper]
- Self-regulating Prompts: Foundational Model Adaptation without Forgetting (ICCV 2023)[paper][code]
- Tangent Model Composition for Ensembling and Continual Fine-tuning (ICCV 2023)[paper][code]
- CBA: Improving Online Continual Learning via Continual Bias Adaptor (ICCV 2023)[paper]
- CTP: Towards Vision-Language Continual Pretraining via Compatible Momentum Contrast and Topology Preservation (ICCV 2023)[paper][code]
- NAPA-VQ: Neighborhood Aware Prototype Augmentation with Vector Quantization for Continual Learning (ICCV 2023)[paper][code]
- Online Continual Learning on Hierarchical Label Expansion (ICCV 2023)[paper]
- Class-Incremental Grouping Network for Continual Audio-Visual Learning (ICCV 2023)[paper][code]
- Rapid Adaptation in Online Continual Learning: Are We Evaluating It Right? (ICCV 2023)[paper][code]
- When Prompt-based Incremental Learning Does Not Meet Strong Pretraining (ICCV 2023)[paper]
- Online Class Incremental Learning on Stochastic Blurry Task Boundary via Mask and Visual Prompt Tuning (ICCV 2023)[paper][code]
- Dynamic Residual Classifier for Class Incremental Learning (ICCV 2023)[paper]
- First Session Adaptation: A Strong Replay-Free Baseline for Class-Incremental Learning (ICCV 2023)[paper]
- Masked Autoencoders are Efficient Class Incremental Learners (ICCV 2023)[paper]
- Introducing Language Guidance in Prompt-based Continual Learning (ICCV 2023)[paper]
- CLNeRF: Continual Learning Meets NeRFs (ICCV 2023)[paper]
- Preventing Zero-Shot Transfer Degradation in Continual Learning of Vision-Language Models (ICCV 2023)[paper][code]
- LFS-GAN: Lifelong Few-Shot Image Generation (ICCV 2023)[paper]
- TARGET: Federated Class-Continual Learning via Exemplar-Free Distillation (ICCV 2023)[paper]
- Learning to Learn: How to Continuously Teach Humans and Machines (ICCV 2023)[paper]
- Audio-Visual Class-Incremental Learning (ICCV 2023)[paper][code]
- MetaGCD: Learning to Continually Learn in Generalized Category Discovery (ICCV 2023)[paper]
- Exemplar-Free Continual Transformer with Convolutions (ICCV 2023)[paper]
- A Unified Continual Learning Framework with General Parameter-Efficient Tuning (ICCV 2023)[paper]
- Incremental Generalized Category Discovery (ICCV 2023)[paper]
- Heterogeneous Forgetting Compensation for Class-Incremental Learning (ICCV 2023)[paper][code]
- Augmented Box Replay: Overcoming Foreground Shift for Incremental Object Detection (ICCV 2023)[paper][code]
- MRN: Multiplexed Routing Network for Incremental Multilingual Text Recognition (ICCV 2023)[paper][code]
- CLR: Channel-wise Lightweight Reprogramming for Continual Learning (ICCV 2023)[paper][code]
- ICICLE: Interpretable Class Incremental Continual Learning (ICCV 2023)[paper]
- Proxy Anchor-based Unsupervised Learning for Continuous Generalized Category Discovery (ICCV 2023)[paper]
- SLCA: Slow Learner with Classifier Alignment for Continual Learning on a Pre-trained Model (ICCV 2023)[paper][code]
- Online Prototype Learning for Online Continual Learning (ICCV 2023)[paper][code]
- Analyzing and Reducing the Performance Gap in Cross-Lingual Transfer with Fine-tuning Slow and Fast (ACL2023)[paper]
- Class-Incremental Learning based on Label Generation (ACL2023)[paper]
- Computationally Budgeted Continual Learning: What Does Matter? (CVPR2023)[paper][code]
- Real-Time Evaluation in Online Continual Learning: A New Hope (CVPR2023)[paper]
- Dealing With Cross-Task Class Discrimination in Online Continual Learning (CVPR2023)[paper][code]
- Decoupling Learning and Remembering: A Bilevel Memory Framework With Knowledge Projection for Task-Incremental Learning (CVPR2023)[paper][code]
- GKEAL: Gaussian Kernel Embedded Analytic Learning for Few-shot Class Incremental Task (CVPR2023)[paper]
- EcoTTA: Memory-Efficient Continual Test-time Adaptation via Self-distilled Regularization (CVPR2023)[paper]
- Endpoints Weight Fusion for Class Incremental Semantic Segmentation (CVPR2023)[paper]
- On the Stability-Plasticity Dilemma of Class-Incremental Learning (CVPR2023)[paper]
- Regularizing Second-Order Influences for Continual Learning (CVPR2023)[paper][code]
- Rebalancing Batch Normalization for Exemplar-based Class-Incremental Learning (CVPR2023)[paper]
- Task Difficulty Aware Parameter Allocation & Regularization for Lifelong Learning (CVPR2023)[paper]
- A Probabilistic Framework for Lifelong Test-Time Adaptation (CVPR2023)[paper][code]
- Continual Semantic Segmentation with Automatic Memory Sample Selection (CVPR2023)[paper]
- Exploring Data Geometry for Continual Learning (CVPR2023)[paper]
- PCR: Proxy-based Contrastive Replay for Online Class-Incremental Continual Learning (CVPR2023)[paper][code]
- Learning with Fantasy: Semantic-Aware Virtual Contrastive Constraint for Few-Shot Class-Incremental Learning (CVPR2023)[paper][code]
- Foundation Model Drives Weakly Incremental Learning for Semantic Segmentation (CVPR2023)[paper]
- Continual Detection Transformer for Incremental Object Detection (CVPR2023)[paper][code]
- PIVOT: Prompting for Video Continual Learning (CVPR2023)[paper]
- CODA-Prompt: COntinual Decomposed Attention-based Prompting for Rehearsal-Free Continual Learning (CVPR2023)[paper][code]
- Principles of Forgetting in Domain-Incremental Semantic Segmentation in Adverse Weather Conditions (CVPR2023)[paper]
- Class-Incremental Exemplar Compression for Class-Incremental Learning (CVPR2023)[paper][code]
- Dense Network Expansion for Class Incremental Learning (CVPR2023)[paper]
- Online Bias Correction for Task-Free Continual Learning (ICLR2023)[paper]

1.1 引言

终身机器学习



1949
Sensitivity-Stability (Hebb)
1987
Stability-Plasticity (Carpenter & Grossberg)
1990
Catastrophic forgetting (McCloskey & Cohen; Ratcliff)
1991
Semi-distributed representations (French)
1993
Pre-training (McRae & Hetherington)
1995
Rehearsal (replay) & pseudo-rehearsal (Robins)
Lifelong robot learning (Thrun & Mitchell)
1997
Pseudo-recurrent ("dual") networks a la neocortex/ hippocampus (French)
1999
Survey on "Catastrophic forgetting in connectionist networks" (French)
2013
"An empirical investigation of catastrophic forgetting in gradient-based neural networks" (Goodfellow et al.)
Recent work on CL

1.2 定义

定义1

(Thrun1996, Silver et al 2013; Ruvolo and Eaton, 2013; Chen and Liu, 2014, Chen and Liu, 2016)

- 系统已经学习了N个任务
- 当遇到第N+1个任务时，系统能够利用知识库中的知识，帮助学习第N+1个任务
- 在学习完第N+1个任务时，使用从N+1个任务学习的结果更新知识库

Thrun S. Is learning the n-th thing any easier than learning the first?. In NIPS. 1995.

1.2 定义

定义2

(Chen and Liu, 2014, 2018)

终身机器学习是一个持续学习过程，其中学习器已经执行一个包含N个任务的任务序列T_1，T_2，...，T_N。在面对第$N+1$个任务T_{N+1}和对应的数据D_{N+1}时，学习器可以利用其知识库中的先验知识来帮助学习T_{N+1}任务。知识库中存储和维护过去N个任务中学习和累积到的知识。在学习了任务T_{N+1}后，知识库会根据T_{N+1}任务中学习到的中间或最终结果进行更新。

理想终身学习还应具备：

在开放环境中学习和运作，需要具备发现要学习新任务的能力

在应用和测试中，学会优化模型性能

Liu B. Lifelong machine learning: a paradigm for continuous learning. Frontiers of Computer Science. 2017 Jun;11:359-61.

1.2 定义

定义2

(Chen and Liu, 2014, 2018)

- 以增量的方式学习N个任务，每个任务有一个对应的训练集D_t
- 目标：增量式学习第N+1个任务

  1. 不发生灾难性遗忘（**catastrophic forgetting**）

  2. 知识迁移：帮助下次学习

- 假设：在任务学习完成后，其数据（至少大部分数据）不可用

Chen and Liu. Lifelong machine learning. Morgan & Claypool. 2018

1.2 定义

终身机器学习的五个关键特征

- 持续学习过程（不发生灾难性遗忘）
- 明确的知识积累和保存（长期记忆）
- 使用已学知识帮助学习

  Forward Transfer：旧任务知识帮助未来任务学习

  Backward Trasnfer: 未来任务提高旧任务模型
- 发现新任务能力
- 边工作边学习能力

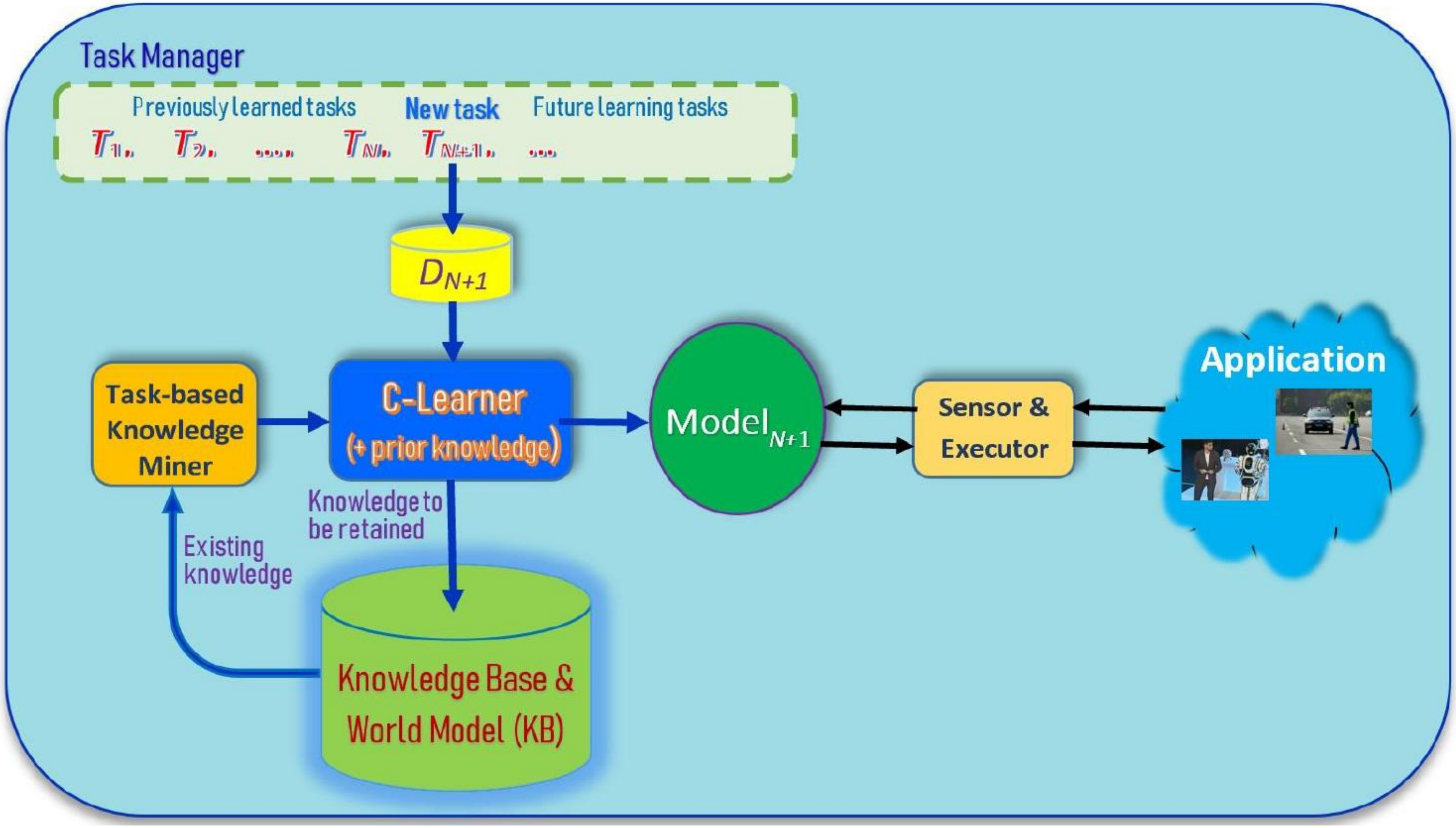
知识表示、获取、推理和存储

任务：不同域、可能存在关联

在交互过程中，自主找到学习任务和数据

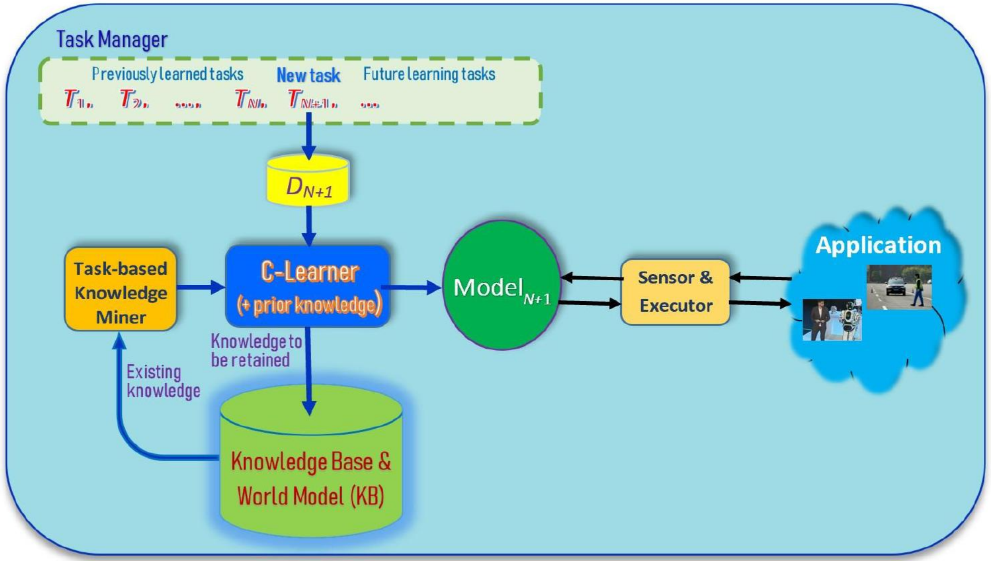
1.2 定义

终身机器学习系统架构



Chen and Liu. Lifelong machine learning. Morgan & Claypool. 2018

1.2 定义



终身机器学习系统架构

- 知识库：存储和推理知识
  - 历史信息库(Past Information Store)

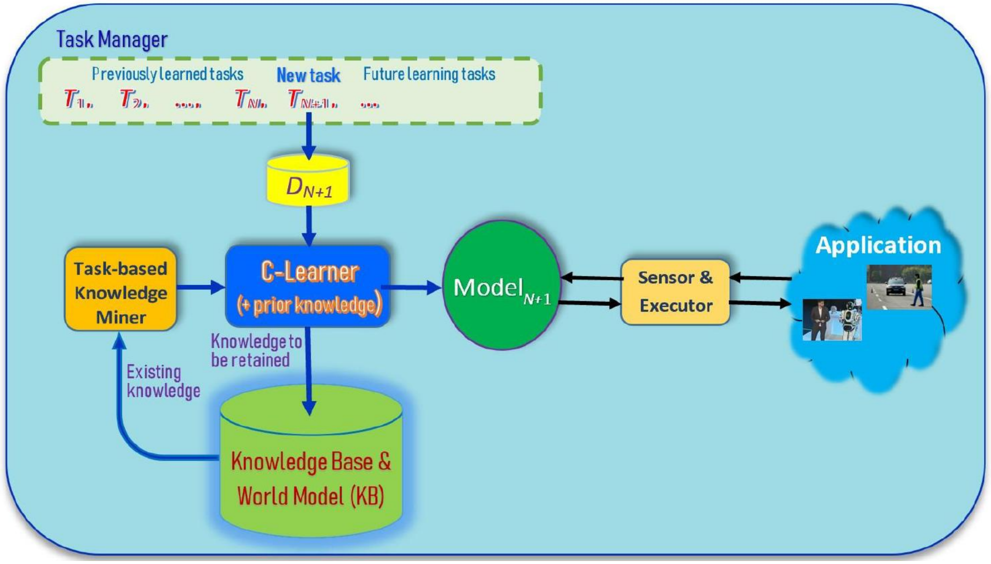
    存储之前学习产生的信息，包括结果模型、模式等
  - 元知识挖掘器(Meta-Knowledge Miner)

    执行知识的元挖掘
  - 元知识库(Meta-Knowledge Store)

    存储从PIS和MKS本身挖掘或合并的知识
  - 知识推理器(Knowledge Reasoner)

    进行推理产生更多知识

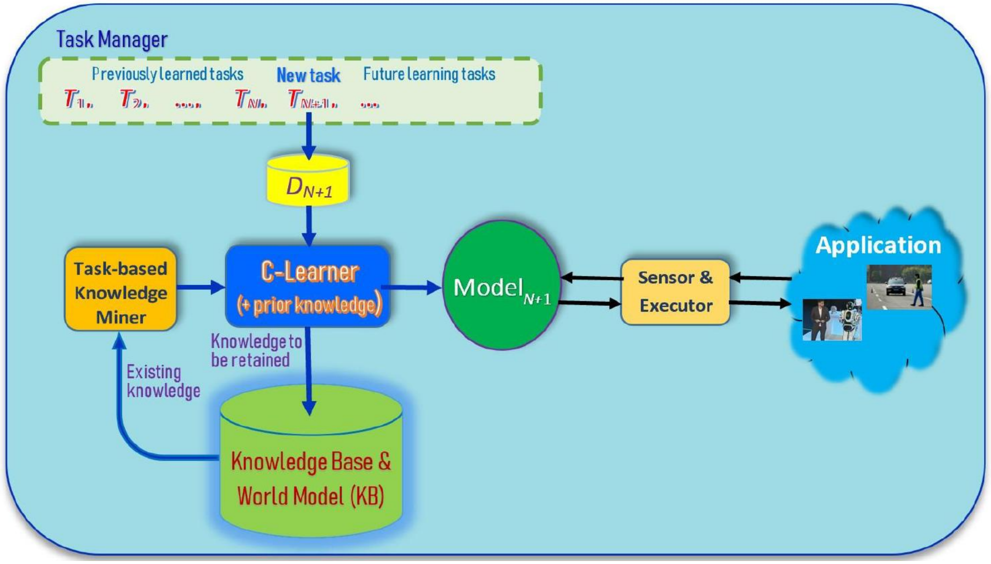
1.2 定义



终身机器学习系统架构

- 知识库：存储和推理知识
- 基于知识的学习器：使用知识进行学习
- 基于任务的知识挖掘器：为新任务挖掘知识库中的知识
- 模型：机器学习中的学习模型

1.2 定义



终身机器学习系统架构

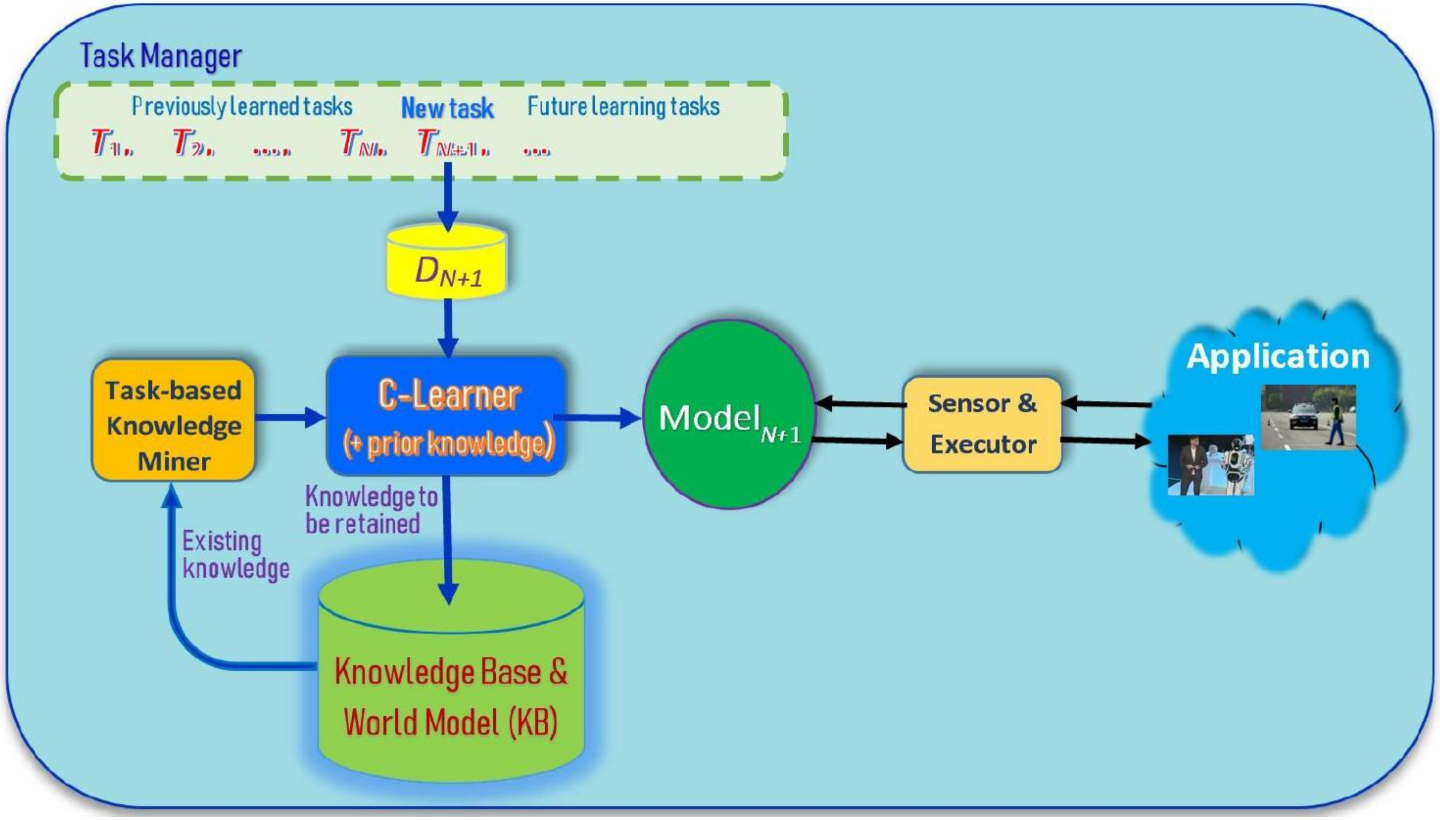
- 知识库：存储和推理知识
- 基于知识的学习器：使用知识进行学习
- 基于任务的知识挖掘器：为新任务挖掘知识库中的知识
- 模型：分类、预测模型及结果
- 应用：实际应用
- 任务管理器：接收和管理到达系统的任务

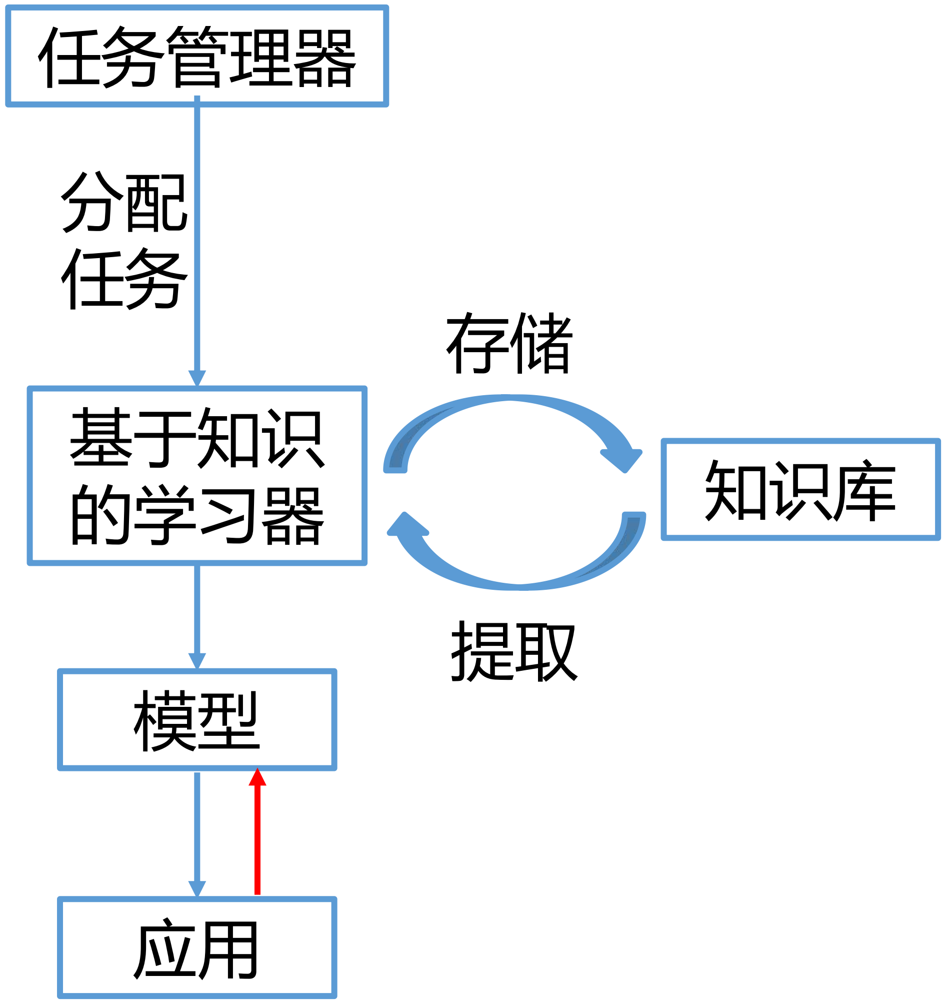
✓ 系统从实际应用中学习知识，并发现新任务

✓ 应用向学习期提供反馈，进行模型优化

1.2 定义

终身机器学习过程





1.3 关键挑战

- 知识

➢缺少通用的定义及表示形式

共享潜在参数、模型参数、模型结果、特征词

➢知识正确性

如何检测错误的知识

发展的过程
受限于任务、场景

➢知识适用性

是否适用于特定任务

1.3 关键挑战

- 知识
- 学习器

  可塑性和稳定性困境

- 开放环境

  应用反馈知识的正确性

  对抗数据问题

谢　谢！