



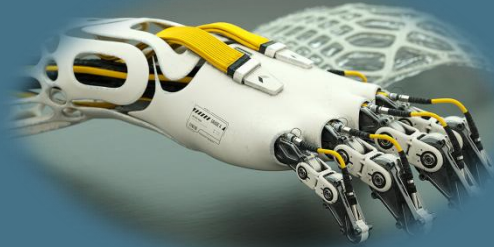
机器学习

授课人：王闻博

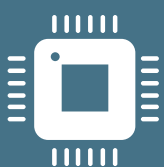
Email: wenbo_wang@kust.edu.cn

昆明理工大学 机电工程学院

2025年02月24日



机器学习 导论



1. 课程简介
2. 人工智能与机器学习的基本概念和发展历程
3. 机器学习任务分类和应用示例

课程简介



• 任课教师

- **王闻博** (wenbo_wang@kust.edu.cn, 机电楼428)
- **研究方向：博弈论视角下的工业物联网组网和资源分配，深度强化学习，面向无人系统的环境感知和群智能决策，基于工业物联网的生产线智能故障诊断、预测和维护决策。**

• 推荐教材

- [1] 姚期智, 人工智能. 9787302612797. 清华大学出版社, 2022.
- [2] 周志强, 缪玲娟. 人工智能技术. 9787576322259. 北京理工大学出版社, 2023. (北京理工大学“双一流”建设精品出版工程教材)

• 其他参考教材

- [1] 王万良, 人工智能导论. 9787040551532. 高等教育出版社, 2020. (首十二五普通高等教育本科国家级规划教材)
- [2] 孙平, 唐非, 张迪, 人工智能基础及应用 (微课版). 9787302614951. 清华大学出版社, 2022.
- [3] 斯图尔特·罗素 (Stuart Russell) [美], 人工智能：现代方法 (第4版). 9787115598103, 人民邮电出版社, 2022.
- [4] 刘远超, 深度学习基础, 9787300253916, 高等教育出版社. 2023.9.



本课程内容涵盖范围

• 本课程将要讲授什么

- 传统统计机器学习的一般性原理、要素、典型任务和基本模型等；
- 机器学习的数学原理概述，包括优化方法初步原理、正则化、概率估计等；
- 深度学习初步，包括神经网络的基本模型和训练原理、卷积神经网络、循环神经网络、注意力机制以及某些生成模型的初步原理等；
- 强化学习初步，包括马尔科夫决策过程、时间差分方法、Q-learning原理等。

• 本课程没有讲授/提供什么

- 可解释机器学习（如以PAC理论为代表的方法）原理；
- 传统的统计机器学习理论中的一大部分方法，如核方法（支持向量机）、EM（期望最大）模型、以聚类为代表的非监督学习等；
- 特征工程方法，如PCA（主元素分析）等；
- 人工智能范畴中的搜索方法、机器自动逻辑推理等；
- 论文级的前沿机器学习方法，如迁移学习、离线强化学习、大语言模型等。

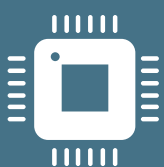
课程主要章节



- 第一章：绪论（人工智能和机器学习的基本概念）
- 第二章：机器学习的数学基础
 - 矩阵分析补缺
 - 条件概率和贝叶斯原理基础
 - 优化问题求解和正则化
- 第三章：统计机器学习中的分类和回归问题
 - 线性和核（广义线性）回归
 - 朴素贝叶斯分类器和多分类器
- 第四章：深度学习基础——前馈神经网络
- 第五章：卷积神经网络
 - 卷积的概念和特征
 - 多通道卷积、池化和残差网络
 - 卷积神经网络的应用：mnist手写数字分类
- 第六章：循环神经网络
- 第七章：注意力机制和Transformer模型
- 第八章：强化学习基础
 - Markov决策过程
 - 从Bellman方程到Q-learning
- 第九章：生成式神经网络
 - 生成对抗神经网络（GAN）



机器学习 导论



1. 课程简介
2. 人工智能与机器学习的基本概念和发展历程
3. 机器学习任务分类和应用示例



人工智能的基本概念

• 智能的特征

- **感知能力：**（人）通过视觉、听觉、触觉、嗅觉等感知器官感知外部世界的能力。
 - 感知是产生智能活动的前提。
 - 80%以上信息通过视觉得到，10%信息通过听觉得到。
- **记忆与思维能力**
 - 记忆：存储由感知器官感知到的外部信息以及由思维所产生的知识。
 - 思维：对记忆的信息进行处理。
- **学习能力**
 - 学习既可能是自觉的、有意识的，也可能是不自觉的、无意识的；既可以是有教师指导的（Supervised），也可以是通过自己实践（Unsupervised, Reinforced）进行的。
- **行为能力（表达能力）**
 - **基于感知能力（信息输入和处理），进行适当的**信息的输出，对外界信息做出反应（肢体、语言等）。



人工智能的基本概念（续）

- **人工智能**

- 用人工的方法在机器（计算机）上实现的智能；或者说是人们使机器具有类似于人的智能。

- **图灵测试（行为主义视角的人工智能定义）**

- 裁判通过终端（不直接见面或听到声音）与机器和人类双方进行多轮对话，机器需通过语言理解与生成能力让裁判误判其身份。如果一台机器能够与人类展开对话而不能被辨别出其机器身份，那么称这台机器具有智能。

- **人工智能学科**

- 一门研究如何构造智能机器（智能计算机）或智能系统，使它能模拟、延伸、扩展人类智能的学科。



人工智能的基本概念（续）

- 对图灵主义人工智能定义的争议
- 约翰·瑟尔（JOHN SEARLE）提出的中文房间问题（1980年）
 - 一个对中文一窍不通，只会英语的人被关在一个封闭房间中。
 - 房间里有一本英文手册，说明该如何处理收到的汉语信息。
 - 房间外的人向房间内输入中文问题，房间内的人便按照手册的说明查找到合适的指示，将相应的中文字符组合成对问题的解答，并将答案输出。
 - 房间外面的人看到自己递进去的中文输入能得到回答，很可能就会认为房间内的人有智能，懂中文，就像现在的聊天机器人那样，那么这是“智能”么？



人工智能研究的基本内容

1. 知识表示

- **知识表示：将人类知识形式化或者模型化。**
- **知识表示方法：符号表示法、连接机制表示法**
 - **符号表示法：**用各种包含具体含义的符号，以各种不同的方式和顺序组合起来表示知识的一类方法。例如，**一阶谓词逻辑、产生式**等。
 - **连接机制表示法：**把各种物理对象以不同的方式及顺序连接起来，并在其间互相传递及加工各种包含具体意义的信息，以此来表示相关的概念及知识。例如，**神经网络**等。



简介：基于符号表示法的谓词逻辑

- **知识库 (KB) 构建**：智能体所处的世界，如状态、感知到的信息、以及信息与事实之间的关联都属于可以利用的知识。知识库是由描述有关信息及其关联的语句遵循一定的语法（如基于命题逻辑的语法）构成，如 $x + y = 4$ 。
- **语义**：语句的含义，它定义了语句在每个“可能世界”的**真值**。
- **推理**：一般地，指**基于逻辑连接词（如与或非）和谓词（Predicate，如 Teacher(Wang)）**关系将基于KB的原子语句构造为复合句后，再根据原子语句的**真值和逻辑连接词的真值判定复合语句的真值（语义）**这一过程。



人工智能研究的基本内容（续）

2. 机器感知

- **机器感知**：使机器（计算机）具有类似于人的感知能力。以机器视觉(machine vision)与机器听觉为代表。

3. 机器思维

- **机器思维**：对通过感知得来的外部信息及机器内部的各种工作信息进行有目的的处理（推理和决策）。

4. 机器学习

- **机器学习**（machine learning）：研究如何使计算机具有类似于人的学习能力，使它能通过学习自动地获取知识。

5. 机器行为

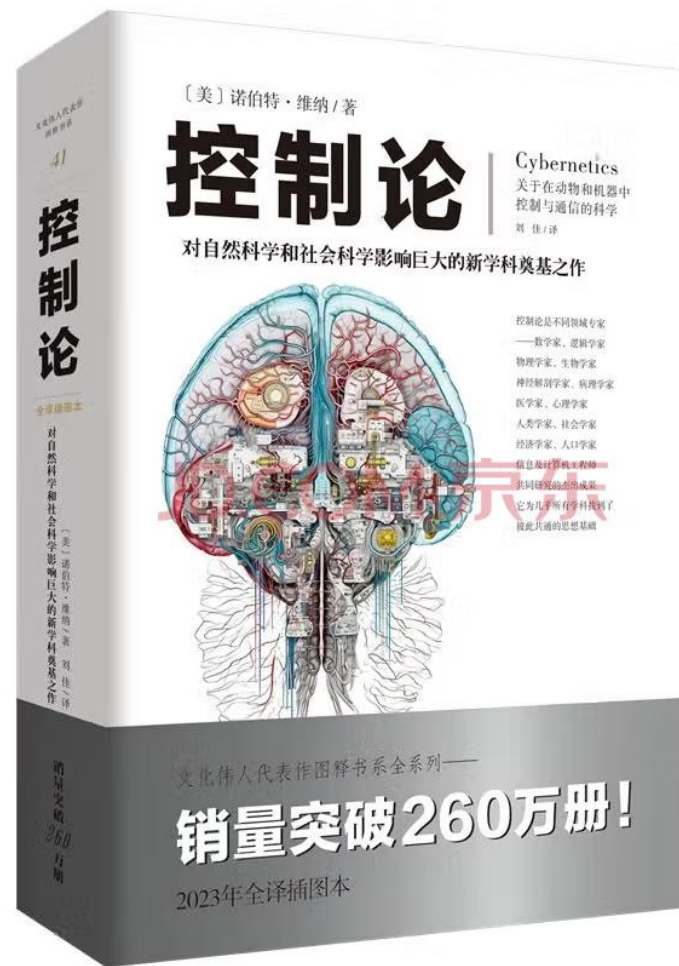
- **机器行为**：计算机的表达能力，即“说”、“写”、“画”等能力。

人工智能的发展简史



• 孕育（1956年之前）

- 1948年维纳（N. Wiener）发表了《控制论——关于在动物和机器中控制和通讯的科学》一书，创立了控制论（维纳当时起名叫Cybernetics）。它是一门研究和模拟自动控制的生物和人工系统的学科，标志着人们根据动物心理和行为科学进行计算机模拟研究和分析的基础已经形成。



人工智能的发展简史



- 形成 (1956年 - 1969年)
- 里程碑：1956年夏的达特茅斯会议
 - 当时美国的一群年轻学者在达特茅斯学院召开了两个月的封闭式学术研讨会，讨论如何用机器来模拟人的智能。
 - 在这个会议上提出了“人工智能”的概念，标志着人工智能作为一门新兴学科正式诞生。
 - 建议研究的领域：（1）自动机；（2）编程语言；（3）神经网络；（4）计算复杂性理论；（5）自我改进；（6）抽象；（7）随机性和创见性。



2006年，会议50年后，几个重要当事人重聚达特茅斯

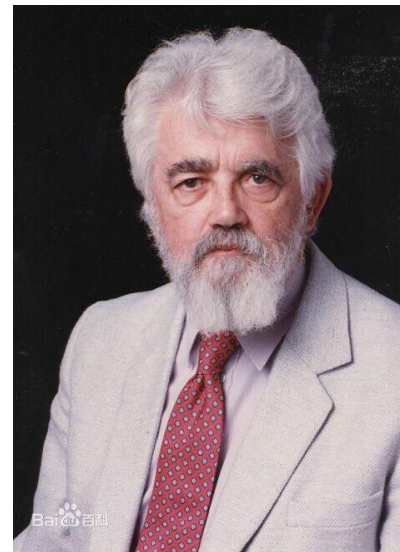
人工智能的发展简史



- **代表人物 (1956年 - 1969年)**

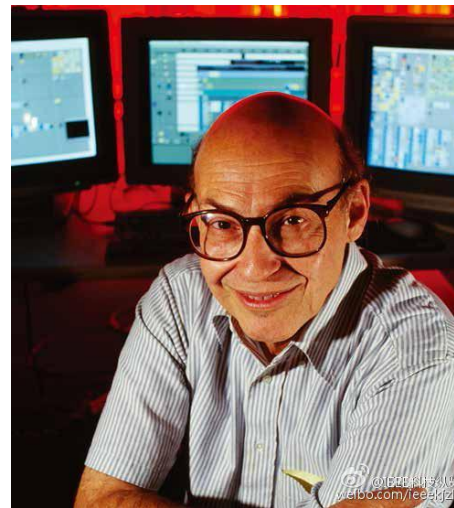
- **麦卡锡 (McCarthy, 1927-2011)**

- 达特茅斯会议的召集者，正式采用“人工智能”这一术语，被称为“人工智能之父”。后创建斯坦福人工智能实验室并任主任。是LISP语言的发明人。



- **明斯基 (Minsky, 1927-2016)**

- 与麦卡锡等一起创建了世界上第一个人工智能实验室：MIT人工智能实验室 (CSAIL)。
- 创立了框架理论，研究神经网络和脑模型，虚拟现实。
- 1969年，因“人工智能理论及软件”获得图灵奖。



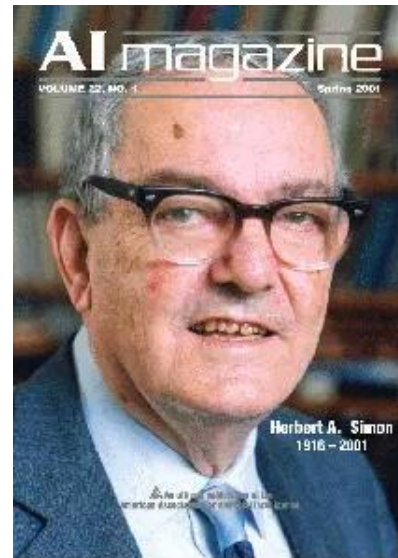
人工智能的发展简史



- **代表人物（1956年 - 1969年）：符号主义学派**

- **西蒙 (Simon, 1916-2001)**

- 1975年和纽厄尔一起被授予图灵奖。
- 1969年美国心理学会杰出科学贡献奖。
- 1978年因对决策理论的研究获“诺贝尔经济学奖”。



- **纽厄尔 (Newel, 1927-1992)**

- 他是信息处理语言（IPL）的发明人之一，并写了该语言最早的两个AI程序，和西蒙合作开发了早期人工智能软件逻辑理论家（Logic Theorist）和一般问题求解器（General Problem Solver, GPS）。



人工智能的发展简史



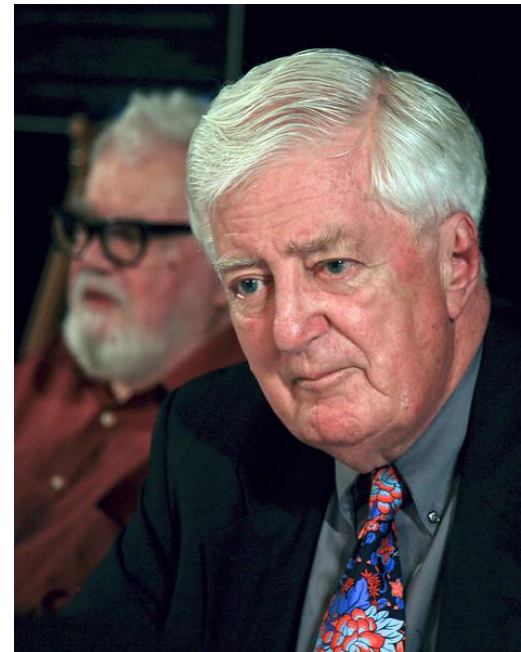
- **代表人物（1956年 - 1969年）：连接主义学派**

- **赛弗里奇(Selfridge, 1926-2008)**

- MIT林肯实验室等作为项目研究负责人和首席科学家1969年美国心理学会杰出科学贡献奖。
- 是模式识别的奠基人，被誉为“机器感知之父”。

- **人工智能“结构”与“功能”两派路线**

- 赛弗里奇和纽厄尔代表了两派人的观点：一派人企图模拟神经系统，而纽厄尔企图模拟心智。
- 经典人工智能主张人工智能应从功能模拟入手，将智能视为符号处理过程，采用形式逻辑实现智能，故称为“符号主义(Symbolism)”或“逻辑主义(Logicism)”。
- 符号主义学派初期过于乐观，赫伯特·西蒙1958年就曾预测计算机10年内就会成为国际象棋冠军。事实上，40年后“深蓝”才战胜国际象棋冠军。



人工智能理论的形成（1956年 - 1969年）



- 1956年纽厄尔和西蒙的“逻辑理论家”程序，该程序模拟了人们用数理逻辑证明定理时的思维规律。该程序证明了怀特海和罗素的《数学原理》一书中第二章中的38条定理；而到了1963年，该程序已经能证明该章中的全部52条定理。这一工作受到了人们高度的评价，被认为是计算机模拟人的高级思维活动的一个重大成果，是人工智能的真正开端。
- 1956年，赛弗里奇研制出第一个字符识别程序，开辟了模式识别这一新的领域。
- 1956年塞缪尔（Samuel）研制了一个跳棋程序，该程序具有自学习功能，能够从棋谱中学习，也能在比赛中总结经验，提高棋艺。它在1959年打败了塞缪尔本人，到了1962年该程序已经可以击败美国一个州的跳棋冠军。这是模拟人类学习过程的一次卓有成效的探索，是人工智能的一个重大突破。
- 1958年麦卡锡（J. McCarthy）研制出的表处理语言LISP，不仅可以处理数据，而且可以方便地处理符号，进行逻辑推理，成为人工智能程序设计语言的重要里程碑。目前LISP语言仍然是人工智能系统重要的程序设计语言和开发工具。

人工智能理论的形成（1956年 - 1969年）



- 1960年，纽厄尔和西蒙等人研制了通用问题求解（GPS）程序，是对人们求解问题时的思维活动的总结。他们发现人们求解问题时的思维活动包括三个步骤：1) 首先想出大致的计划；2) 根据记忆中的公理、定理和解题计划，按计划实施解题过程；3) 在实施解题过程中，不断进行方法和目的的分析，修正计划。其中，他们首次提出了启发式搜索的概念。
- 1965年鲁宾逊（J.A. Robinson）提出**逻辑归结法**，被认为是一个重大的突破，也为定理证明的研究带来了又一次高潮。
- 1968年，Stanford大学费根鲍姆（Feigenbaum）等人研制成功了化学分析**专家系统**DENDRAL，被认为是专家系统的萌芽，是人工智能研究从一般思维探讨到专门知识应用的一次成功尝试。
- 知识表示采用了奎廉（Quillian）提出的特殊的结构：“语义网络”，将知识表示为图结构。明斯基在1968年从信息处理的角度对语义网络的使用做出了很大的贡献。

人工智能理论的发展（1970年 - 今）

• 60年代末，人工智能研究开始遇到瓶颈：

- 机器定理证明领域遇到了瓶颈，计算机推了数十万步也无法证明两个连续函数之和仍是连续函数。
- 塞缪尔的跳棋程序停留在了州冠军的层次，无法进一步战胜世界冠军。
- 在机器翻译领域，1966年美国顾问委员会的报告裁定：还不存在通用的科学文本机器翻译，也没有很近的实现前景。
- 当时，英国、美国中断了大部分机器翻译项目的资助。

The **spirit** is willing but the **flesh** is weak.
(心有余而力不足)



The **wine** is good but the **meat** is spoiled.
(酒是好的，但肉变质了)

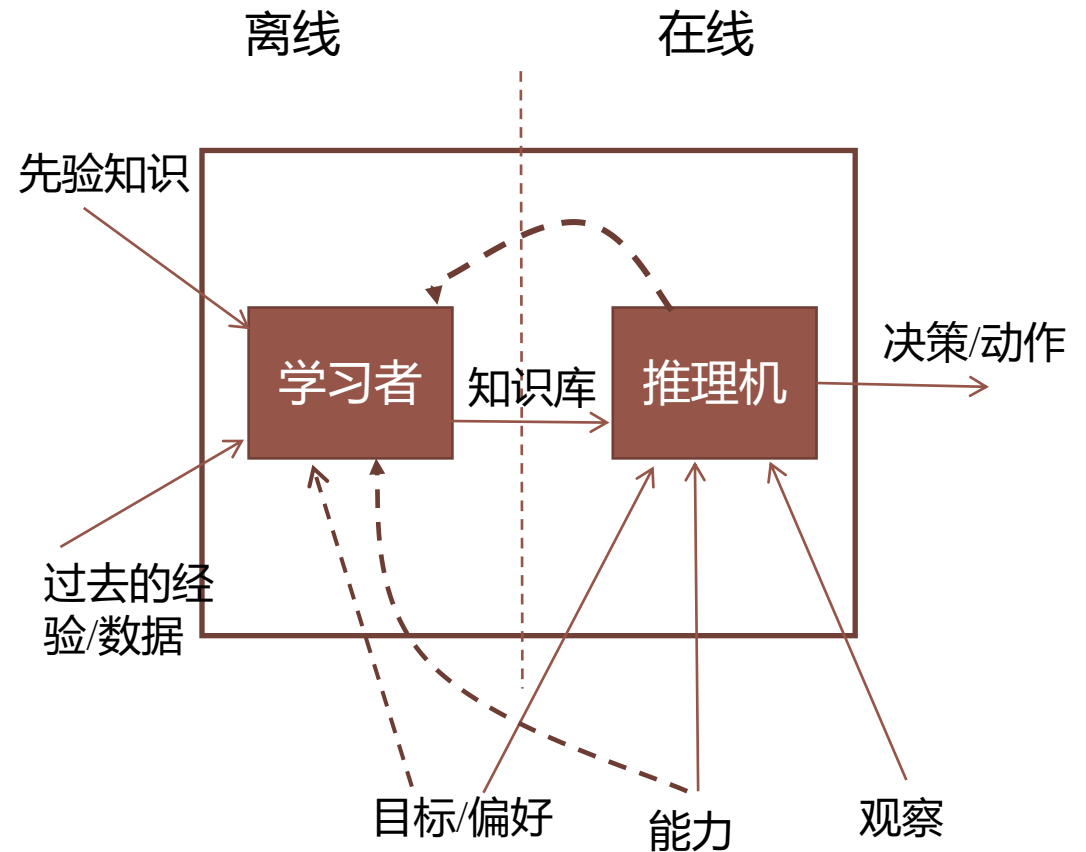
人工智能理论的发展：智能体主义

• 新希望：以知识为中心的时代

- 1977年，费根鲍姆在第五届国际人工智能联合会议上提出了“知识工程”概念，推动了知识为中心的研究。知识工程和专家系统成为人工智能研究的主流，并取得了很多可喜的成果。
- 在知识工程的刺激下，80年代左右，各国开始雄心勃勃地布局人工智能。日本的第五代计算机计划、英国的阿尔维计划、西欧的尤里卡计划、美国的星球大战计划和当时中国相应的863计划陆续推出。

• 知识和行动

- 知识是关于一个领域的信息。
- 基于知识的系统是使用领域知识来行动或解决问题的系统。



一般化的智能系统学习-推理框架



人工智能理论的发展：机器学习的兴起

• 机器学习大致可以分为下面三种类型

- **监督学习**：通过标注的数据来学习，例如，程序通过学习标注了正确答案的手写数字的图像数据，它就能认识其他的手写数字。
- **无监督学习**：通过没有标注的数据来学习。这种算法可以发现数据中自然形成的共同特性（聚类），可以用来发现不同数据之间的联系，例如，买了商品A的顾客往往也购买了商品B。
- **强化学习**：我们可以让程序选择和它的环境互动（例如玩一个游戏），环境给程序的反馈是一些“奖励”（例如游戏中获得高分），程序要学习到一个模型，能在这种环境中得到高的分数，不仅是当前局面要得到高分，而且最终的结果也要是高分才行。



人工智能理论的发展（1970年 - 今）

• “人工智能”与“机器学习”的边界日益模糊

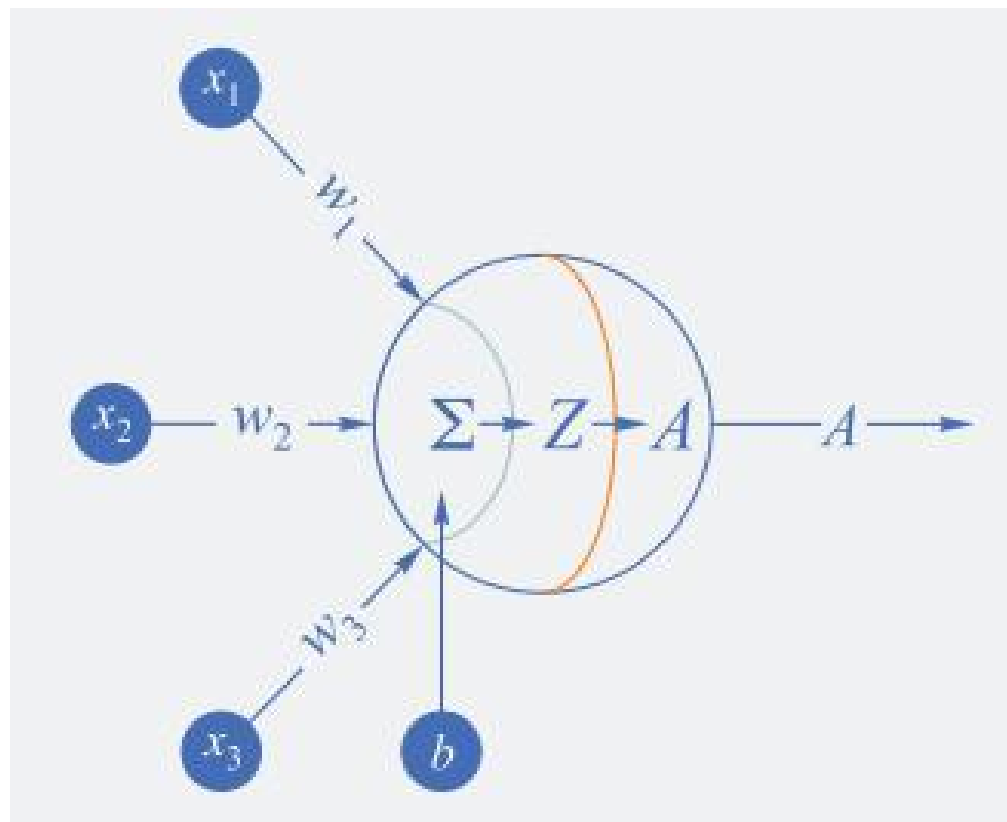
- **符号主义传统**：从“功能”上模拟人类的智能，认为智能机器是符号系统，通过符号操作来模拟人的认知过程；智能的基础依赖于知识，知识表示、知识推理、知识运用是人工智能的核心。又称为逻辑主义、心理学派。典型应用：Coq交互式定理证明。
- **连接主义**：从“结构”仿生上实现人工智能，研究神经网络及神经网络间的连接机制与学习算法。认为智能活动依赖的基本物理元素是神经元，认知过程是大量神经元的并行活动。典型应用：DeepSeek大语言模型。
- **行为主义**：认为人工智能源于控制论，这一学派的代表人物首推MIT的布鲁克斯，他开发了基于感知-动作模式模拟昆虫智能行为的控制系统。典型应用：基于强化学习的具身智能（如宇树机器人）。

以前馈神经网络为代表的连接主义模型



• 连接主义：当今显学“人工神经网络”

- 神经网络模型的原型在1943年就出现了。在生物神经网络中，每个神经元与其他神经元相连，当它兴奋时，就会像相邻的神经元发送化学物质，从而改变这些神经元内的电位；如果某神经元的电位超过了一个阈值，那么它就会被激活（兴奋），向其他神经元发送化学物质。把许多这样的神经元按照一定的层次结构连接起来，我们就构建了一个神经网络。
- 右图是 M-P 神经元模型的示意图。





人工神经网络的演进

• 当今显学“人工神经网络”

• 第一阶段（1940年代中～1960年代中）：**神经元和初级网络模型**

- M-P神经元，感知器模型等：典型学习问题如“XOR问题不能解决”

• 第二阶段（1980年代中～1990年代）：**神经网络的短暂复兴**

- BP（Back Propagation）算法，解决了多层前馈网络的学习训练问题
- 1986年，David E. Rumelhart, **Geoffrey E. Hinton** 和 Ronald J. Williams 三人在《Nature》发表论文：“**Learning representations by back-propagating errors**”。
- Hopfield网络，成功求解了旅行商问题。
- 存在问题：梯度消失/爆炸；训练速度慢，无法收敛；硬件计算能力所限，难以解决大规模网络训练问题。

人工神经网络的演进（续）

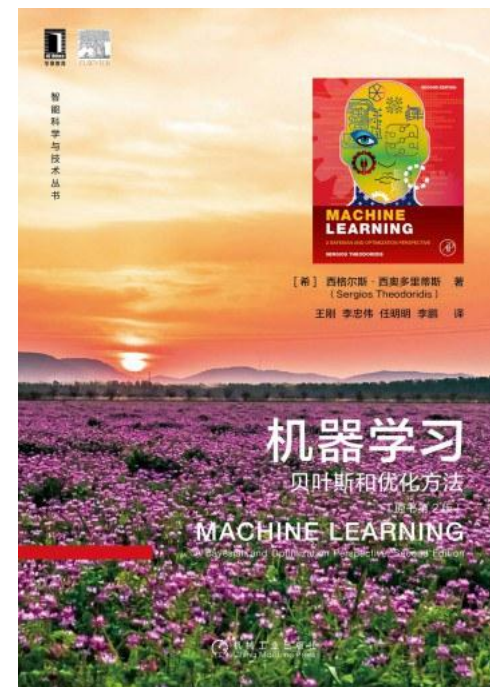


• 当今显学“人工神经网络”

- 第三阶段（1990年代中～2010年代）：**神经网络的沉寂和统计机器学习的兴起**
 - 主要由俄罗斯统计学家、数学家万普尼克（Vapnik）建立。
 - 支持向量机，一般的“核（Kernel）方法”。
 - 贝叶斯方法。
 - 为连接学派提供理论支持。
 - 应用：数据挖掘和分析，计算机视觉，自然语言处理。



Vladimir N. Vapnik



人工神经网络的演进（续）



• 人工神经网络

- 第四阶段（2010年代~今）：**深度学习与大数据时代**
 - 硬件计算能力大幅度提升。
 - 大幅更新的训练和正则化策略。
 - 传感器和互联网提供海量数据。



Geoffrey E. Hinton

面向问题：大规模神经网络训练时的梯度消失/爆炸问题；训练加速和收敛问题；解决大规模网络训练问题的硬件计算能力。

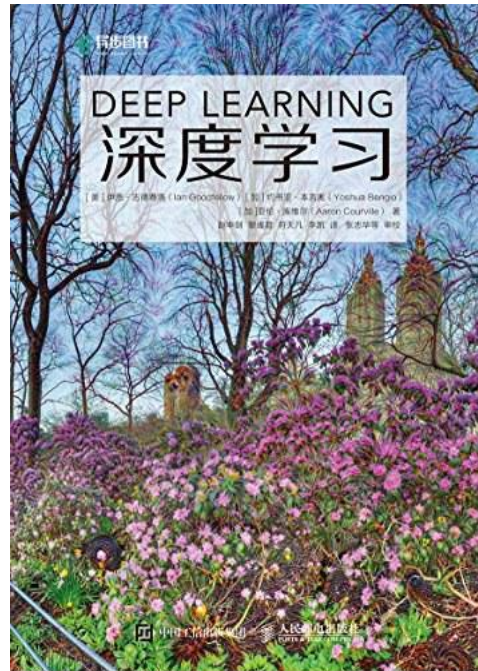
人工神经网络的演进（续）

• 连接主义：当今显学“人工神经网络”

• 第四阶段（2010年代~今）：

深度学习与大数据时代

- 2004年，执着于神经网络研究的辛顿获得加拿大高级研究院(CIFAR)每年50万加元的经费支持，召集为数不多的同道，启动了“神经计算与自适应感知”项目。



Yann LeCun
(杨立昆)

Geoffrey Hinton
杰弗里·辛顿

Yoshua Bengio
约舒瓦·本吉奥

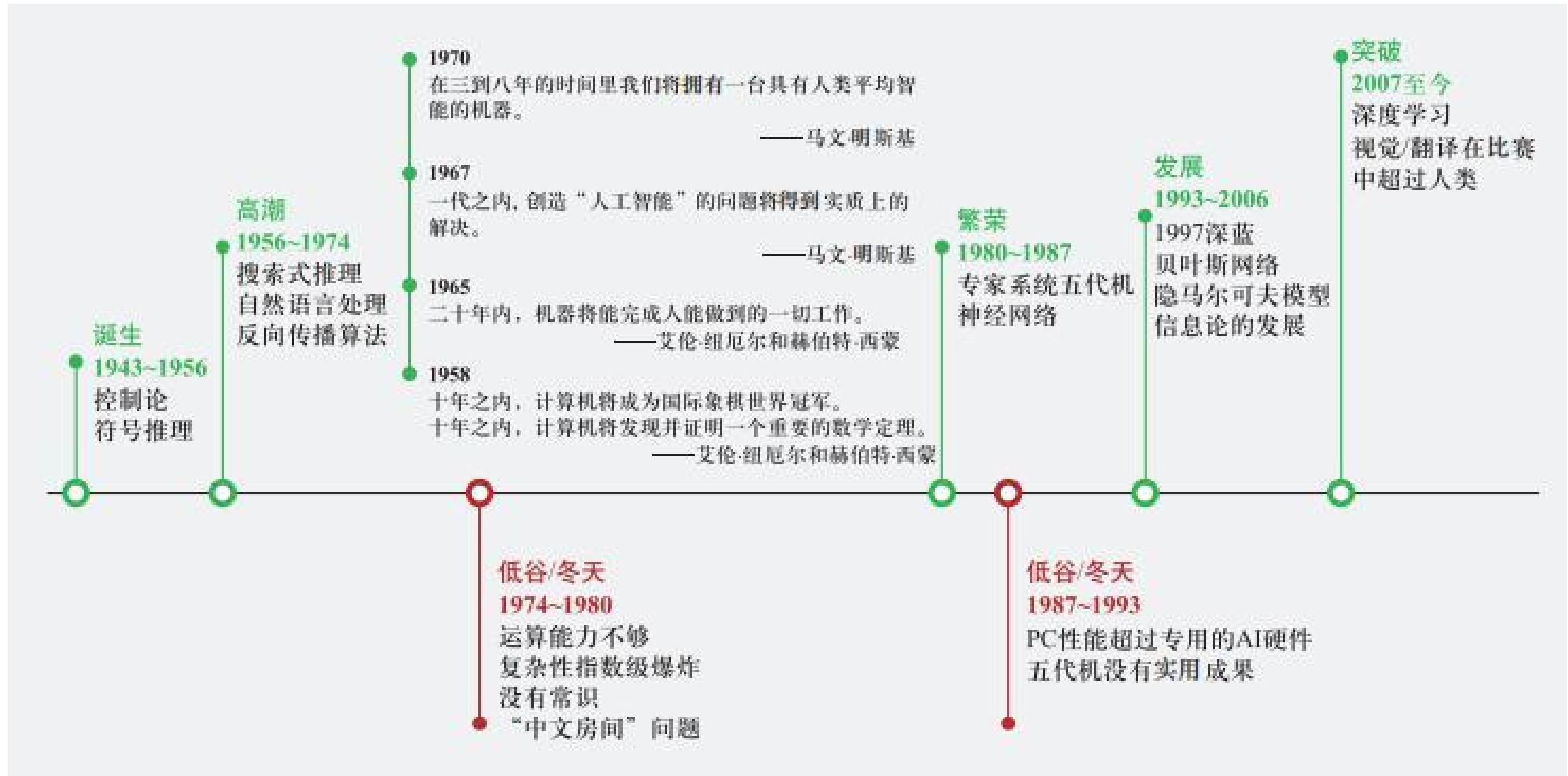
人工智能理论的发展（1970年 - 今）



• 深度学习兴起的里程碑事件

- 2006年，辛顿在《科学》发表论文，提出深度信念网络(Deep Belief Networks, DBNs)，开始掀起神经网络第三次浪潮。
- 2012年10月，辛顿团队把深度学习用于图像识别，将ImageNet视觉对象分类错误率从26%降低到15%，引发深度学习的全球热潮。
- 2017年5月，在[中国乌镇围棋峰会](#)上，AlphaGo 与排名世界第一的世界围棋冠军[柯洁](#)对战，以3比0的总比分获胜。2017年10月18日，DeepMind团队公布了最强版阿尔法围棋，代号AlphaGo Zero。
- 2022年11月，OpenAI推出了大型语言模型ChatGPT，它是在GPT-3基础上构建，能非常逼真地与人类对话，还能完成写代码、修bug、写小说和文案、做作业等一系列常见文字输出型任务。

人工智能理论的发展回顾

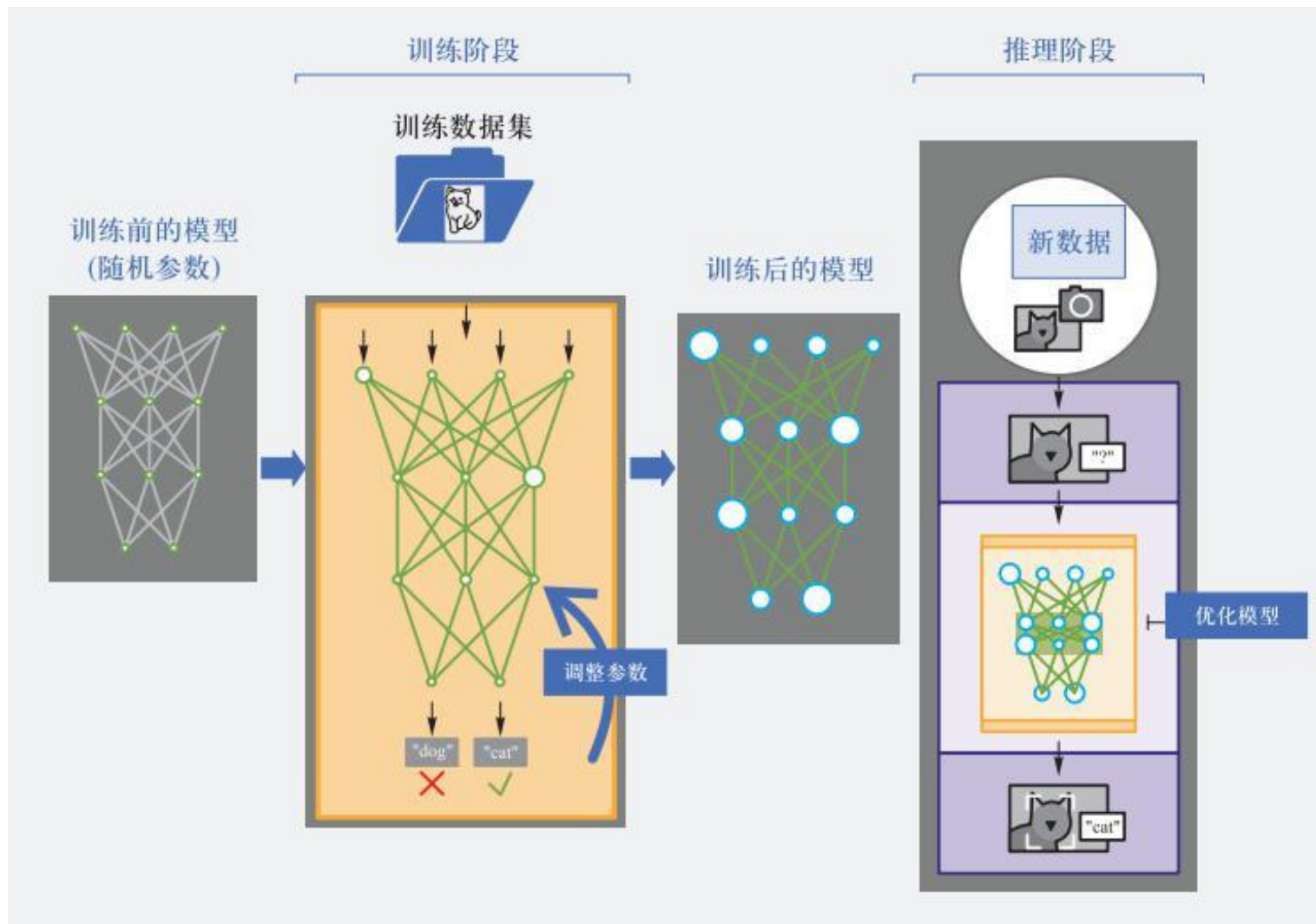
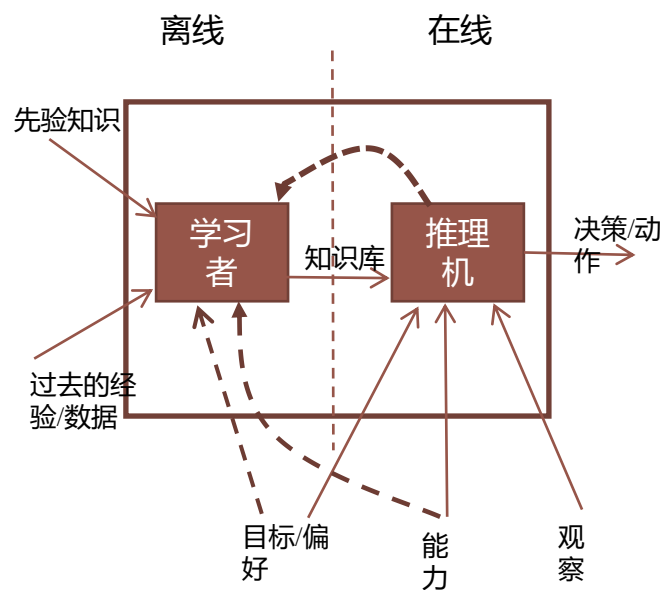


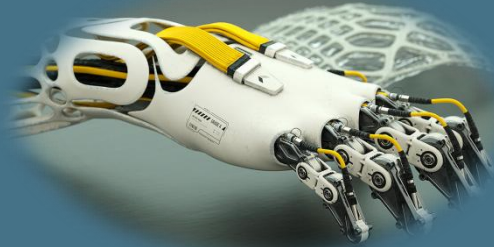
以神经网络为代表的统计机器学习方法



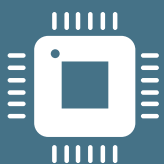
• 人工神经网络的工作范式

- 一个典型的有监督神经网络机器学习模型是怎么得来的，又是怎么在应用中使用的？





机器学习 导论



1. 课程简介
2. 人工智能与机器学习的基本概念和发展历程
3. 机器学习任务分类和应用示例



机器学习任务分类

- **监督学习 (Supervised Learning)**

- **定义：**通过带标签数据（输入特征 x + 对应标签 y ）学习映射关系，模型可预测新数据标签。
- **三大任务：**

类型	标签特点	典型应用
回归	连续值（如房价、温度）	线性回归、股票预测
分类	离散类别（二分类/多分类）	图像识别、垃圾邮件过滤
结构化学习	输出结构化对象（序列/图/树）	机器翻译、语音识别
结构化学习解码：通过动态规划优化联合特征空间映射（如序列标注）。		

机器学习任务分类 (续)



- **无监督学习 (Unsupervised Learning)**

- 定义：从无标签数据中挖掘潜在模式，无需明确输出指导。
- 典型任务：聚类分析（用户细分）、降维（PCA）、特征提取（自编码器）。

- **强化学习 (Reinforcement Learning)**

- 定义：智能体通过环境交互获得延迟奖励，动态调整策略以最大化长期收益。
- 特点：试错（Trial-and-Error）机制（如游戏AI）、策略优化（如机器人控制）、基于深度神经网络的函数近似（如Q-value近似）等。

机器学习的前沿应用举例



• 模式识别

- 机器视觉。
- 自然语言处理。
- 智能信息检索与数据挖掘。
- 多媒体内容生成。



我是 DeepSeek, 很高兴见到你!

我可以帮你写代码、读文件、写作各种创意内容, 请把你的任务交给我吧~

给 DeepSeek 发送消息

深度思考 (R1)

联网搜索



以DeepSeek和Kimi为代表的llm涵盖了自然语言处理、图像生成模型, 多模态学习等多个子领域



基于无人机视觉平台的目标识别 (YOLO算法)

内容生成: 国外社交网站上使用生成式AI (Generative AI) 生成的“特朗普被捕新闻图片”

机器学习的前沿应用举例



• 自动驾驶和具身智能

- 实时互动：实时感知环境变化并作出相应反应。
- 自主学习：通过与环境交互和学习，不断优化自身行为。
- 多模态感知：利用多种传感器（如摄像头、激光雷达等）获取环境信息。
- 跨学科融合：涉及计算机视觉、自然语言处理、认知和推理等多个学科子领域。



人民解放军战士和机器狗



25年春晚的《秧BOT》



在武汉试运营的“萝卜快跑”无人出租车

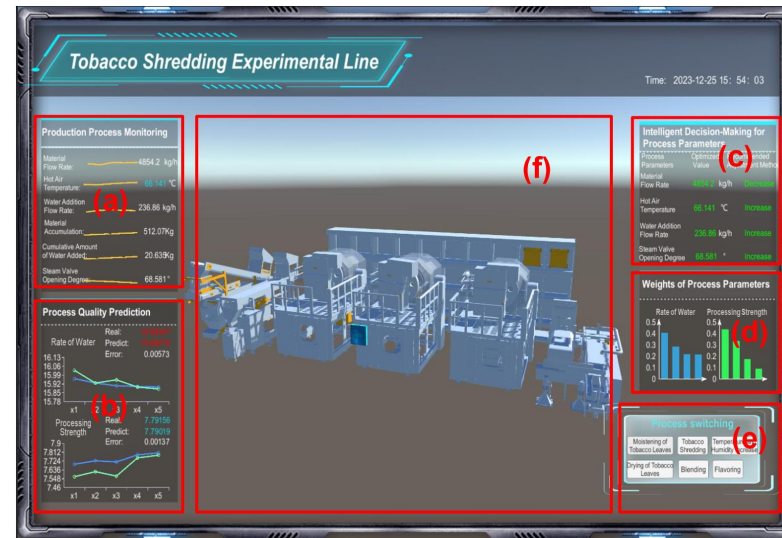
机器学习的前沿应用举例



- 机器博弈
- 自动程序设计
- 分布式机器学习 (如联邦学习与多智能体)
- 智能控制和仿真
- 智能管理与智能决策
- etc, etc...



智能控制驱动的机械外骨骼



智能决策驱动的烟草产线数字孪生



人形机器人编队的
机器人足球对抗

ML前沿应用举例：国家电网的人形机器人





本章作业：开放式问题2选1

• 选做作业1

- 请搜集并总结关于基于DeekSeek、Kimi、Qwen、豆包等大语言模型的科技论文写作的提示词，必须包括（但不限于）课题和文献归纳、研究思路梳理、文本逻辑订正、代码框架生成，以及图表生成几个部分。
- 基于现有免费大语言模型工具，将总结后的上述内容生成报告文档。

• 选做作业2

- 以“世界模型（World Model）”为课题，通过网络、文献搜索，完成一篇不少于6页的双栏英文文献综述（如有可能，请参考科技论文写作课上的内容），要求是用Latex工具（如overleaf），以IEEE期刊论文形式完成。
- 建议：充分利用现有llm工具，完成对相关论文的检索、阅读理解和内容综述。