

神经网络与深度学习

Neural Network and Deep Learning

2026版



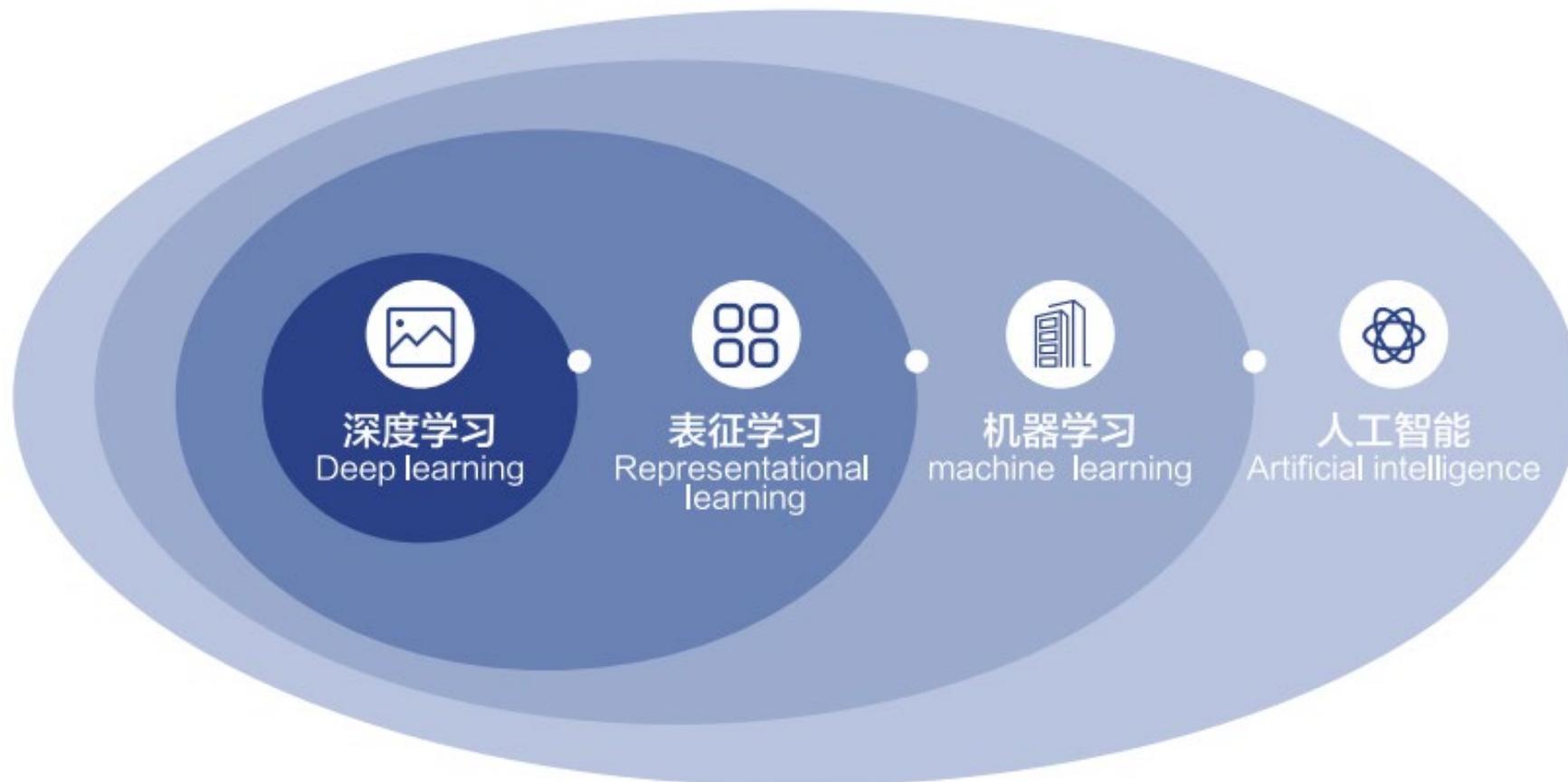
第1章 绪论

课程简介



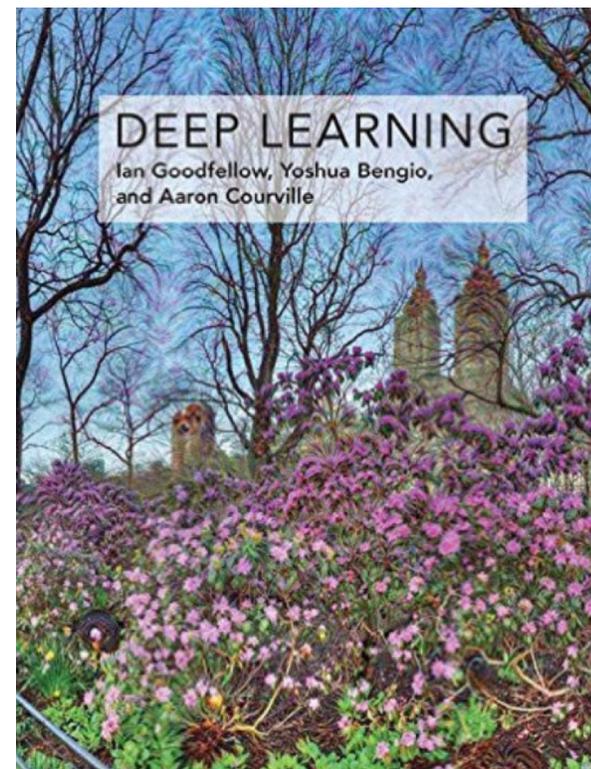
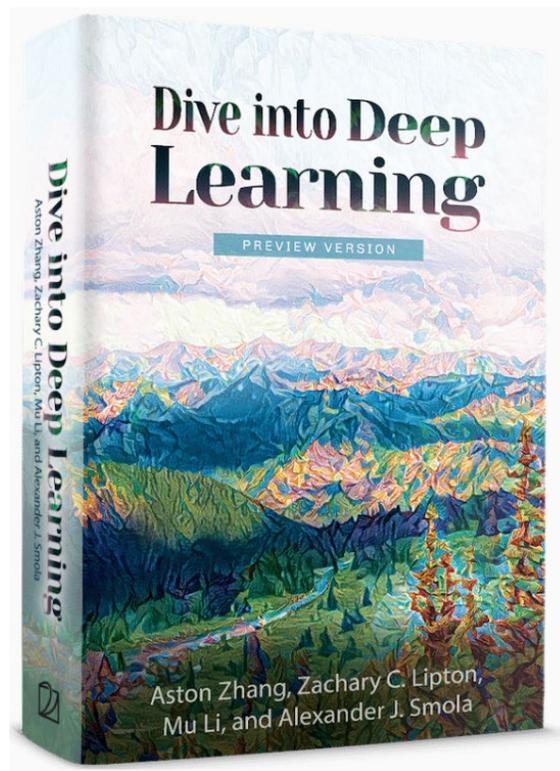
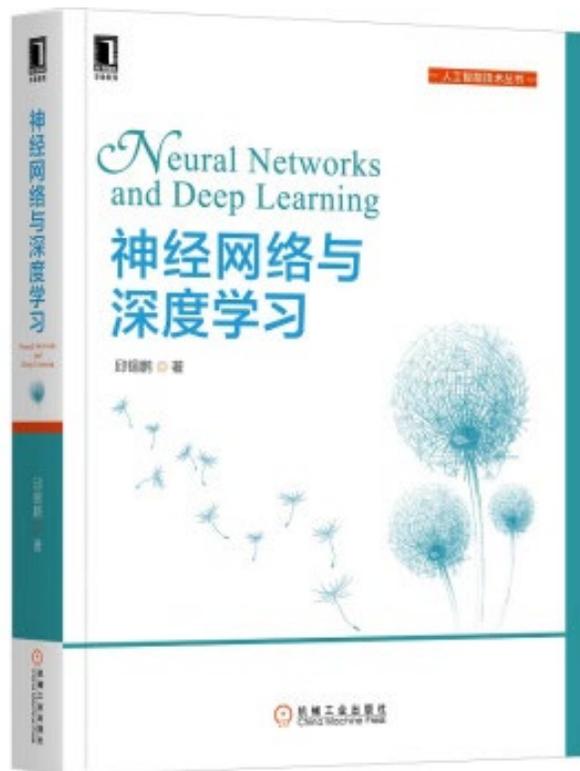
关于本课程

- **神经网络**: 一种以 (人工) 神经元为基本单元模型
- **深度学习**: 一类机器学习问题, 主要解决贡献度分配问题
- **人工智能的一个子领域**



参考教材

- 《神经网络与深度学习》，邱锡鹏， <https://nndl.github.io/>
- 《动手学深度学习》，李沐 等人 <https://d2l.ai/>
- 《Deep Learning》，Goodfellow et al. <https://www.deeplearningbook.org/>



课时分配

- 理论课：40学时 + 实验课：24学时
- 实验课包括：7~8次实验 + 知行会客厅 + 1次讨论课

考核方式

- 平时成绩：50%（出勤及课堂表现10% + 实验报告20% + 讨论课成绩20%）
- 期末成绩：50%（闭卷或开卷考试？）

课程资源

- 课程主页：<https://fwdeng.github.io/>
- 课程微信及QQ群

信息技术与计算机

AI科技评论

ID:aitechtalk

简介

雷锋网旗下人工智能垂直媒体，报道 AI 顶会前沿成果，关注学界业界大牛与大公司的 AI 实验室动态。



识别二维码, 了解公众号

信息技术与计算机

量子位

ID:QbitAI

简介

追踪人工智能新趋势，报道科技行业新突破。



识别二维码, 了解公众号

信息技术与计算机

机器之心

ID:almosthuman2014

简介

专业的人工智能媒体和产业服务平台。



识别二维码, 了解公众号

信息技术与计算机

机器人大讲堂

ID:RoboSpeak

简介

集前沿技术、产业资讯、创业服务、投融资于一体的机器人垂直领域服务平台。



识别二维码, 了解公众号

信息技术与计算机

新智元

ID:AI_era

简介

智能+ 中国主平台。关注人工智能、机器人前沿，及人机融合、人工智能和机器人革命对社会的影响。



识别二维码, 了解公众号

信息技术与计算机

大数据文摘

ID:BigDataDigest

简介

普及数据思维，传播数据文化。



识别二维码, 了解公众号

信息技术与计算机

MIT 科技评论
Technology
Review



识别二维码, 了解公众号

信息技术与计算机

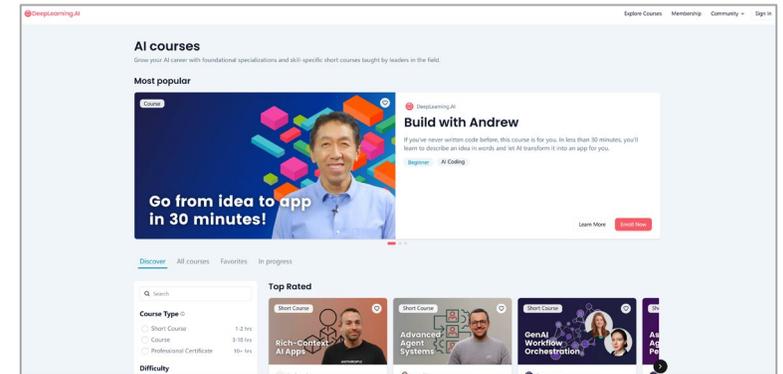


KnowingAI知智
更懂人工智能行业的知识传播者



识别二维码, 了解公众号

1. MIT 6.S191 Introduction to Deep Learning, <http://introtodeeplearning.com/>
2. 斯坦福大学CS231n: CNNs for Visual Recognition, <http://cs231n.stanford.edu/>, Feifei Li等人
3. 机器学习、深度学习, 李宏毅, B站
4. <https://learn.deeplearning.ai/> 吴恩达 等
5. 《智能计算系统》, 中国科学院计算技术研究所



- **人工智能综合会议**: NeurIPS、ICLR、ICML、AAAI...
- **自然语言处理**: ACL、EMNLP...
- **计算机视觉**: CVPR、ICCV...
- **交叉领域**: ACM MM、WWW、IROS...

学科背景



8 - 15 March 2016



AlphaGo


Lee Sedol


人工智能
AlphaGo

- 2019年3月：
OpenAI以2-0完胜
dota2 ti冠军og战队

- 象棋
- 围棋
- 德州扑克





Google™

self-driving car

无人驾驶汽车 深瞳

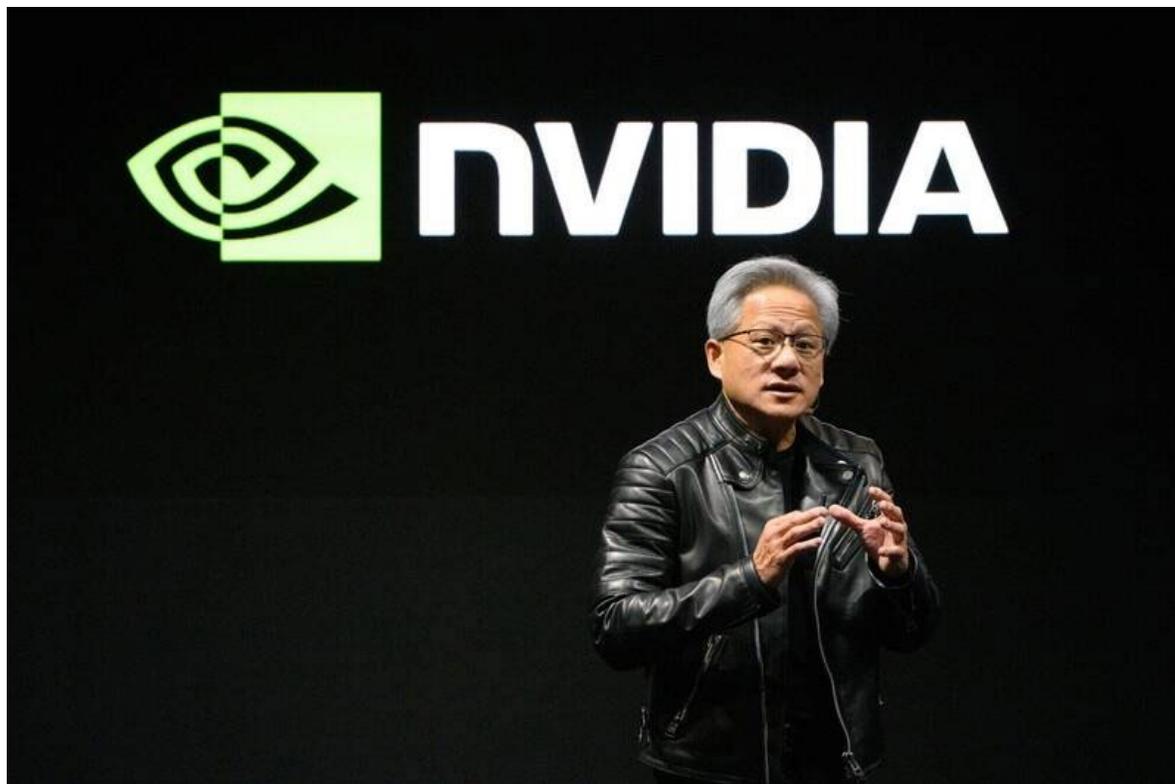
Google, 特斯拉, Uber



OpenAI

ChatGPT **5.0**

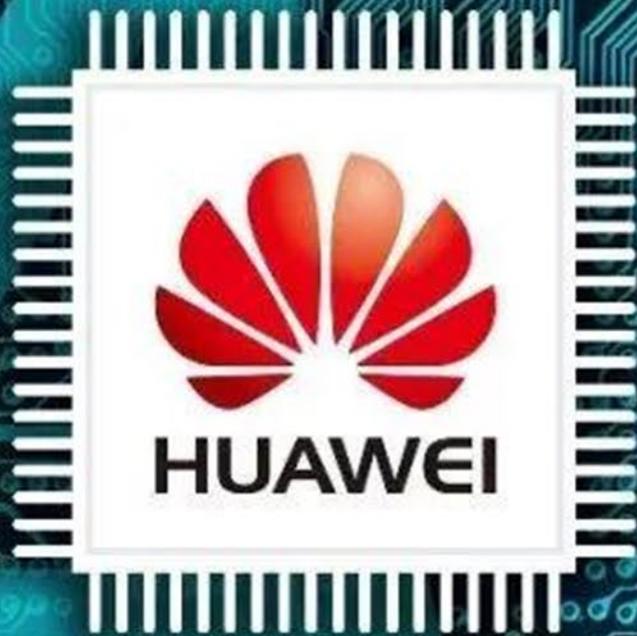
AI算力芯片



英伟达市值登顶



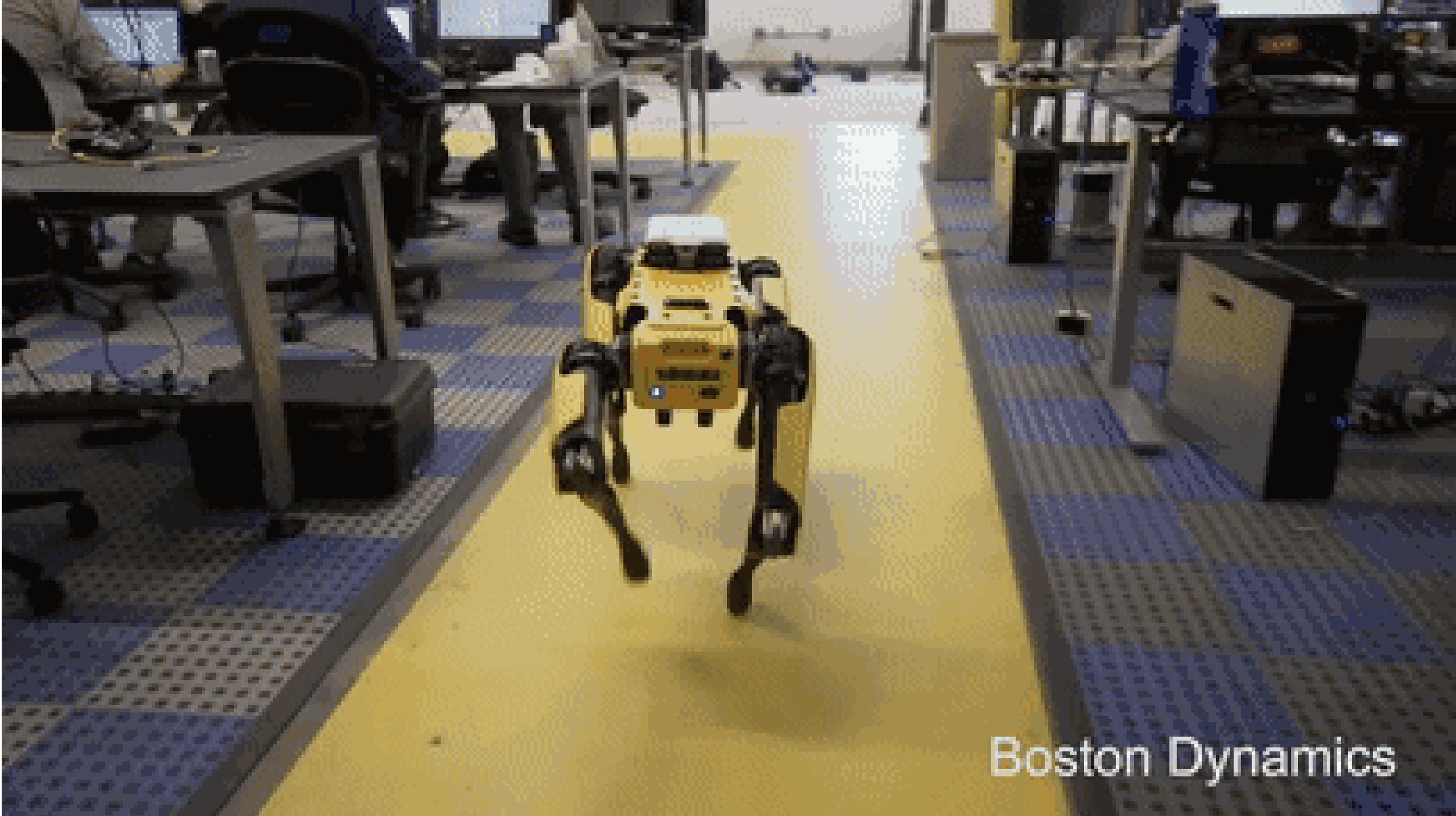
NVIDIA成为过去25年里表现最好的股票，自从启动首次公开招股以来，英伟达股票的总回报率达到了591078%



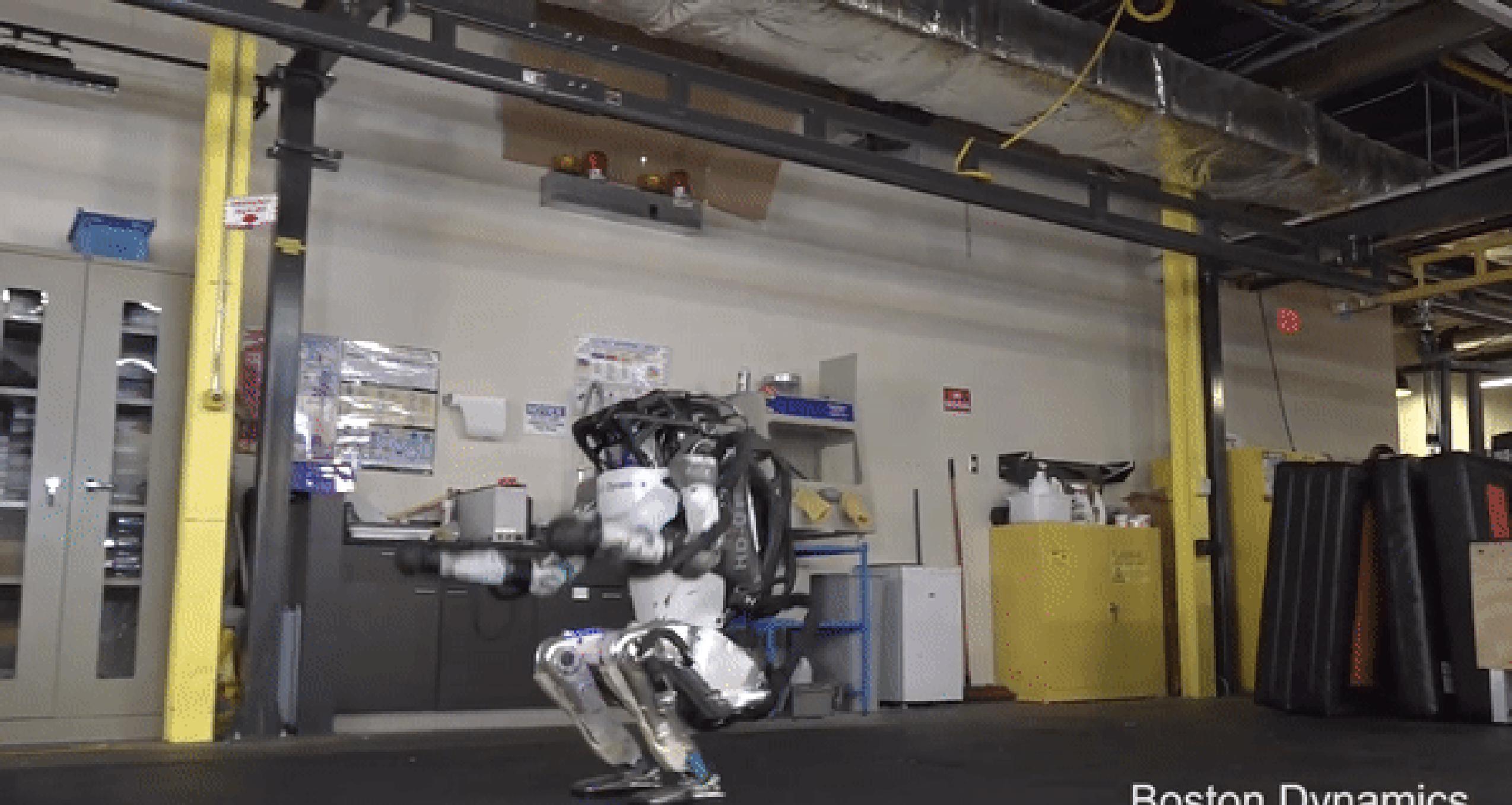
智能芯片
寒武纪、华为海思、地平线



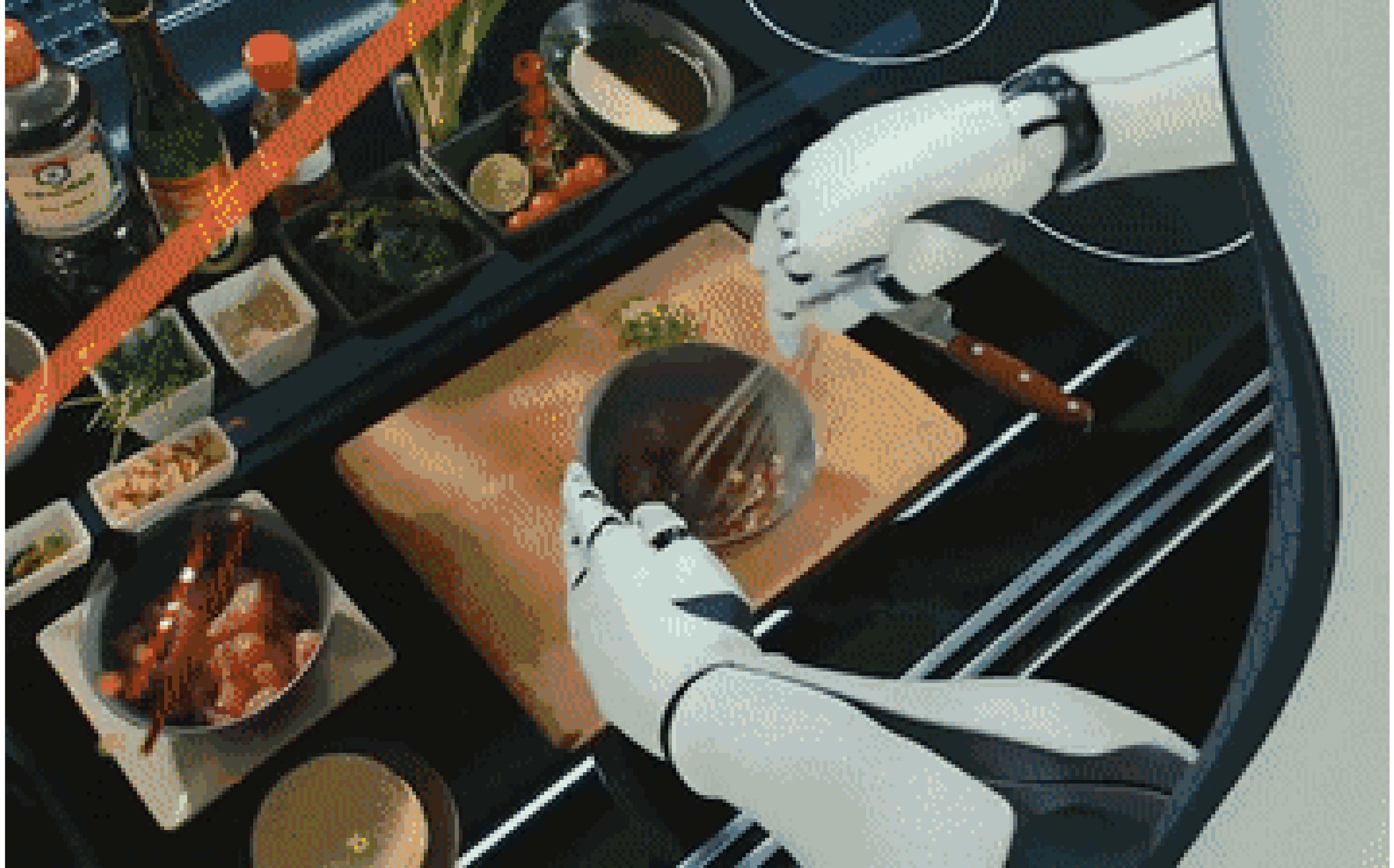
机器人
具身智能



Boston Dynamics



Boston Dynamics



工业4.0

智能工厂+智能生产

工业1.0: 机器制造, 机械化生产;
 工业2.0: 流水线, 批量生产, 标准化;
 工业3.0: 高度自动化, 无人/少人化生产;
 工业4.0: 网络化生产, 虚实融合。



1959年第一个可编程逻辑控制器PLC



第四次工业革命
 系统基于信息物理融合

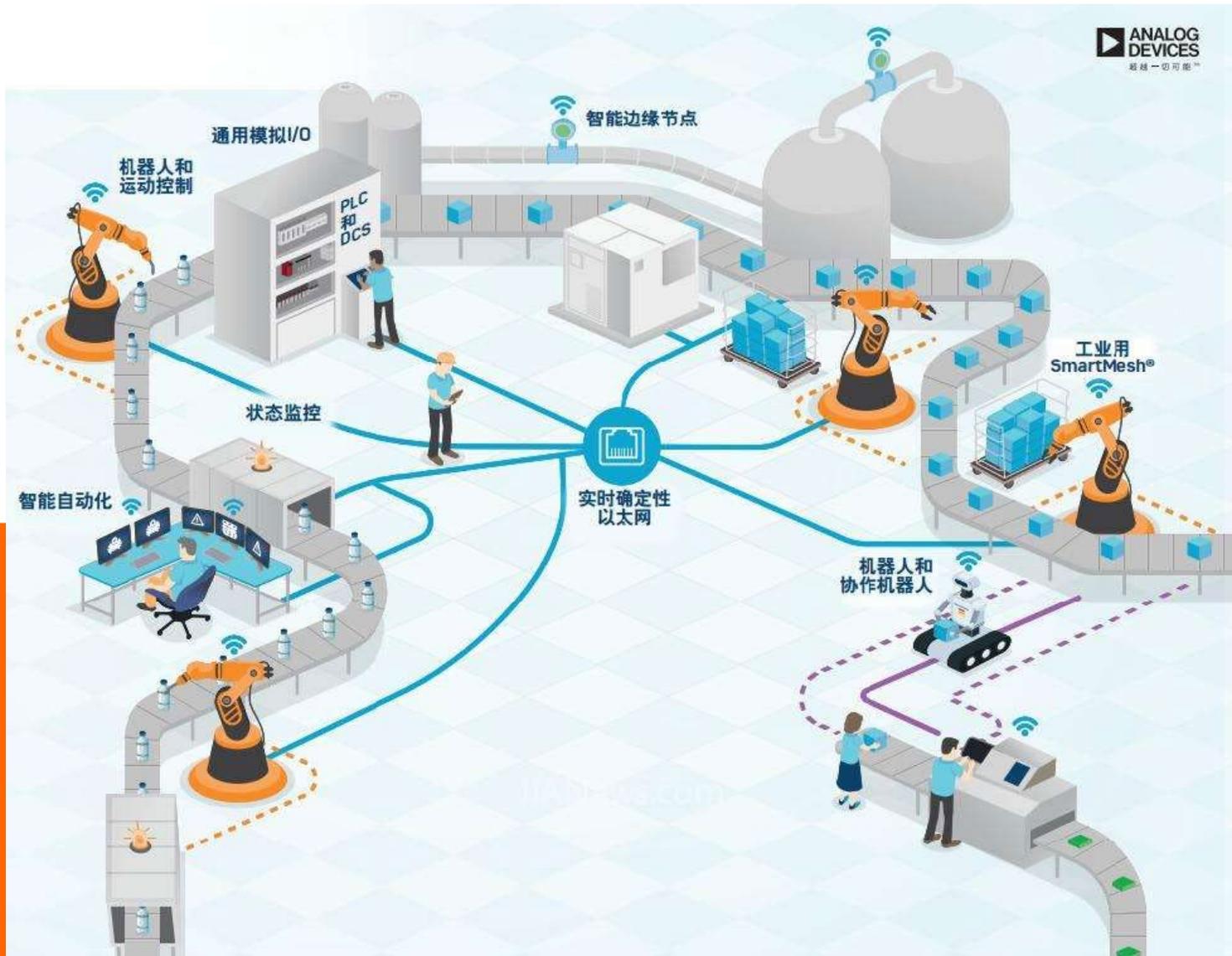


1870年第一条生产线
 美国辛辛那提屠宰场

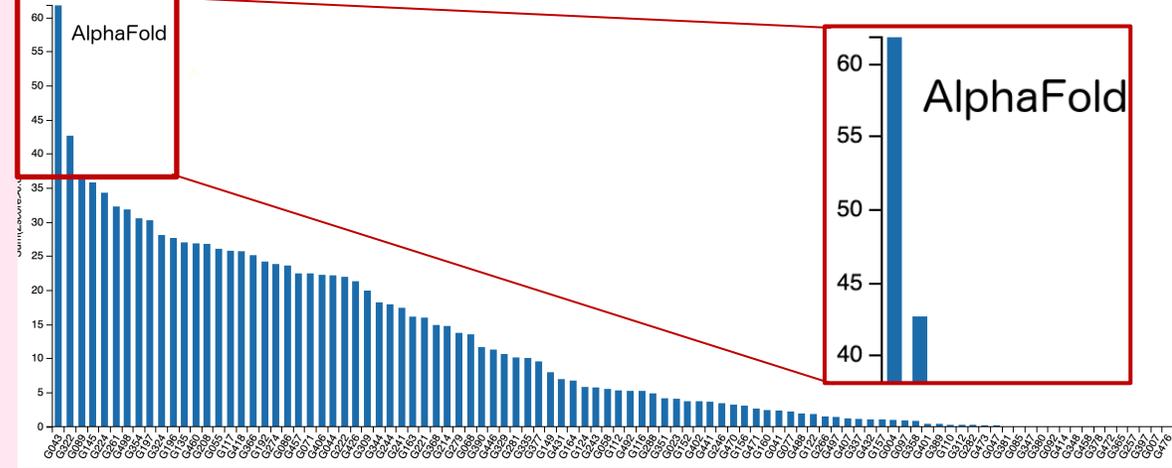
第二次工业革命
 随着基于劳动分工的、电力驱动的大规模生产的出现



1784年第一台纺织机
 第一次工业革命
 随着蒸汽驱动的机械制造设备的出现



AlphaFold: Using AI for scientific discovery

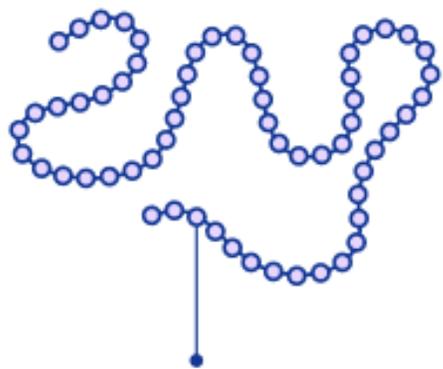


Every protein is made up of a sequence of amino acids bonded together

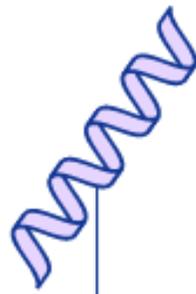
These amino acids interact locally to form shapes like helices and sheets

These shapes fold up on larger scales to form the full three-dimensional protein structure

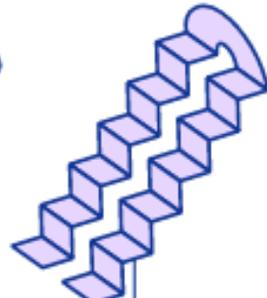
Proteins can interact with other proteins, performing functions such as signalling and transcribing DNA



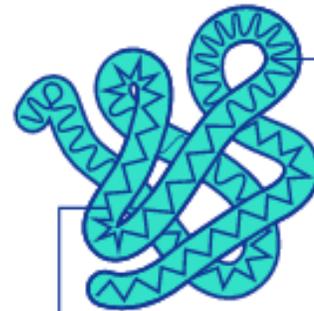
Amino acids



Alpha helix



Pleated sheet



Pleated sheet

Alpha helix

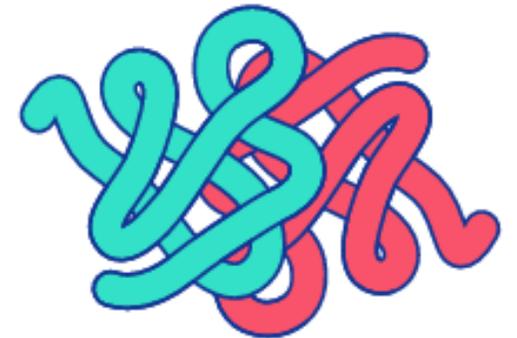


FIGURE 1: COMPLEX 3D SHAPES EMERGE FROM A STRING OF AMINO ACIDS.

生物制药

HUAWEI CLOUD Pangu drug molecule model: Faster drug research and development

华为云盘古药物分子大模型，大幅提升药物研发效率

🕒 Drug design takes a long time
药物设计周期长

📈 Experiments are costly and time-consuming
人工实验成本高、耗时长

ZINC
DrugSpaceX
UniChem
~1.7 billion chemicals
~17亿化合物

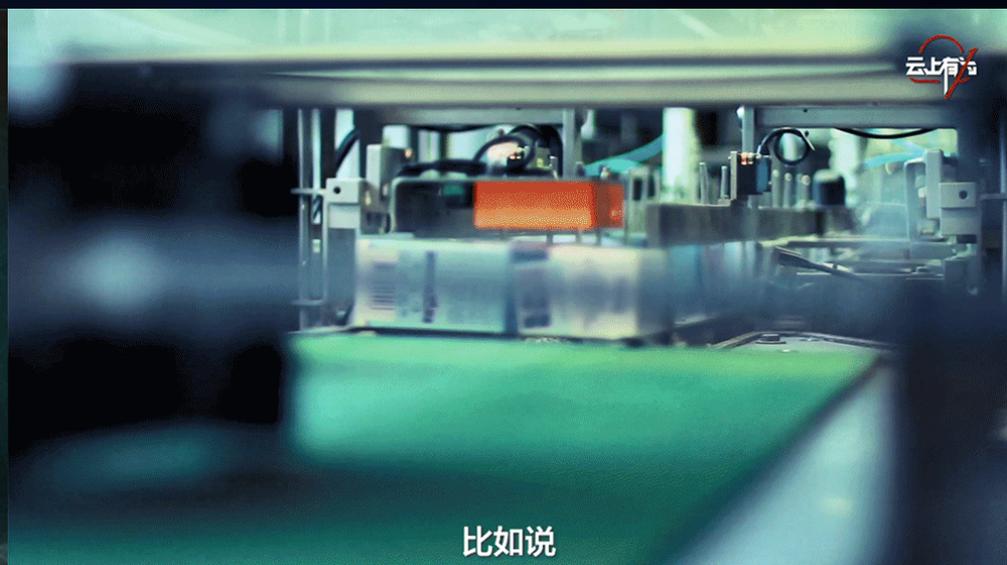
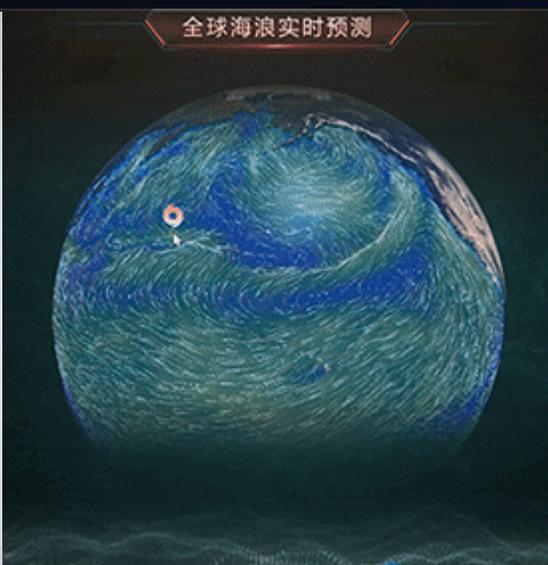


- Novel compound generation
新颖化合物生成
- Target-protein matching
蛋白靶点匹配
- Compound property prediction
化合物属性预测
- Compound structure optimization
化合物结构优化

New antimicrobial discovery
抗菌药发现 (联合西交大一附院)



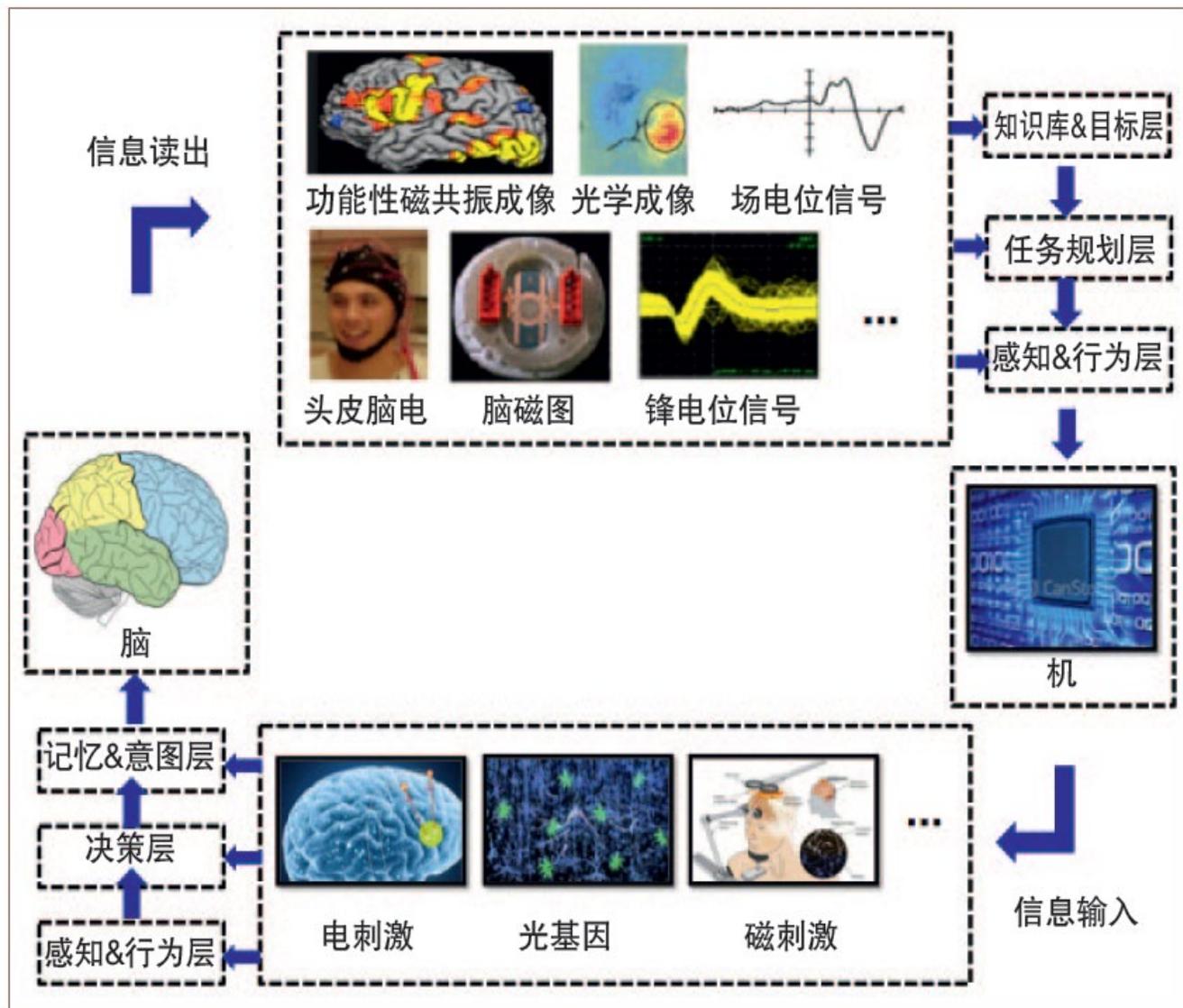
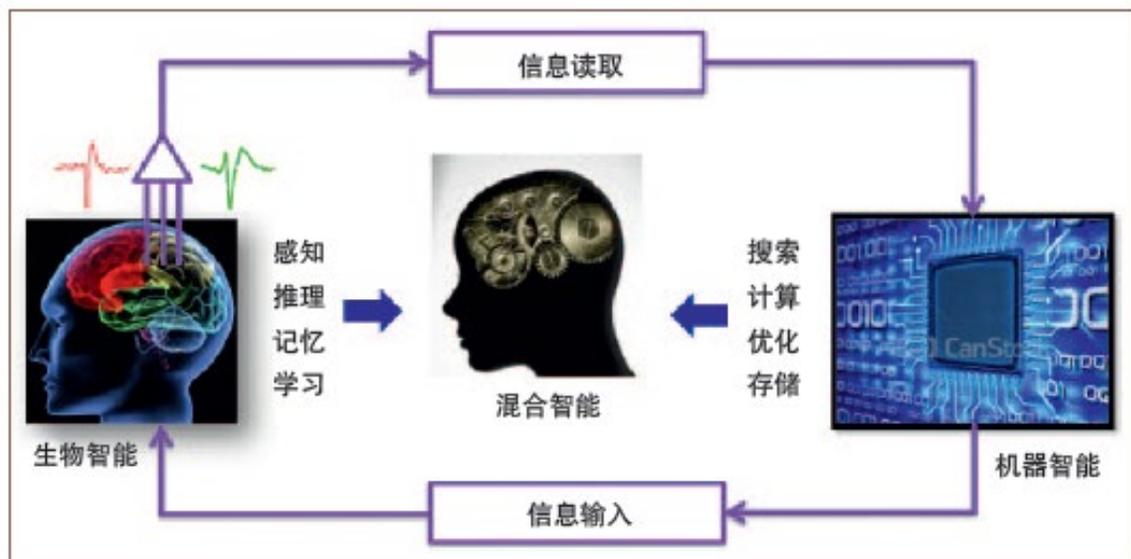
Years
数年 → 1 month
1个月
Lead compound discovery cycle
先导药研发周期



混合智能

Human-in-the-Loop

分类方式	混合智能形态		
智能混合方式	增强型混合智能	替代型混合智能	补偿型混合智能
功能增强方式	感知增强混合智能	认知增强混合智能	行为增强混合智能
信息耦合方式	穿戴人机协同混合智能	脑机融合混合智能	脑机一体化的混合智能



应用背景



应用场景



关键技术

深度学习/机器学习、计算机视觉、语音技术/自然语言处理、人机交互

基础支撑

传感器、芯片、大数据、云计算、服务商、生物识别

计算机视觉



智能交互技术



语音技术



自然语言处理



自动驾驶



激光雷达



智能机器人



无人机



机器学习深度学习技术平台



处理器芯片计算平台



数据平台



2020

全球100家最具潜力 人工智能初创企业

医疗健康



金融 & 保险



交通



建筑



零售 & 仓储



政府 & 城市规划



媒体 & 娱乐



教育



制造



法律



数据挖掘



能源



通信



食品 & 农业



房产



跨领域

AI处理器



NLP, NLG & CV



销售 & CRM



AI模型开发



网络安全

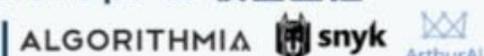


业务 & 运营智能



其他

DevOps & 模型监控



Created by You. Powered by CBINSIGHTS

北京、上海、深圳、杭州、西安、广州是拥有计算机视觉领域人才最多的TOP6城市且这前六名城市拥有的该领域人才总和超过全国的50%。同时，作为新一线城市的代表，杭州和西安在计算机视觉领域人才的储备量上超越了广州，分别排名全国第四和第五。

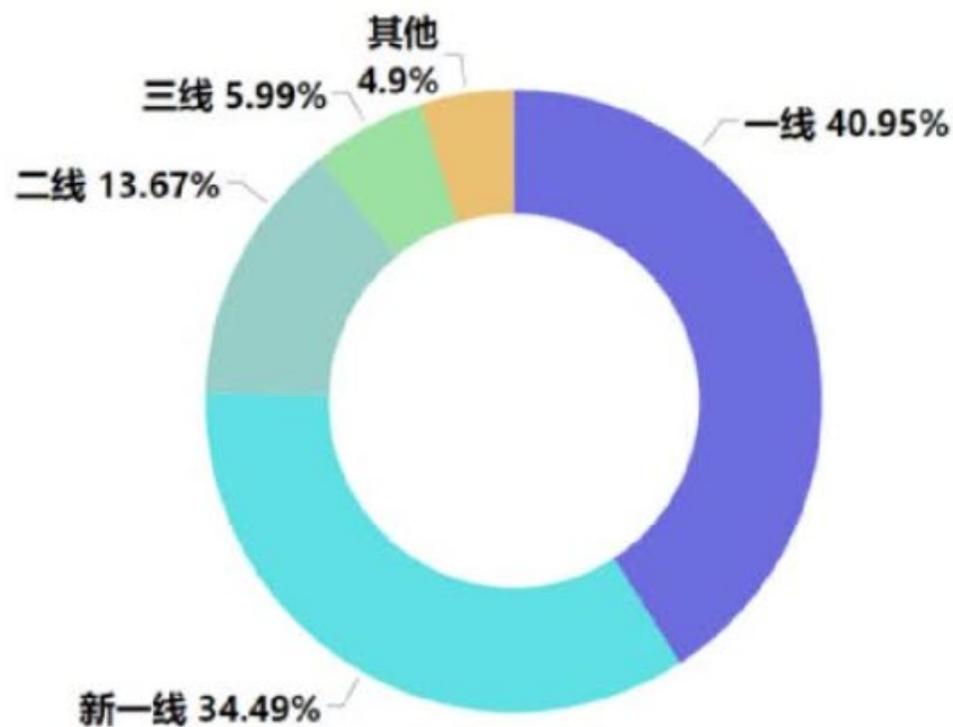


图10 中国计算机视觉人才城市分布情况

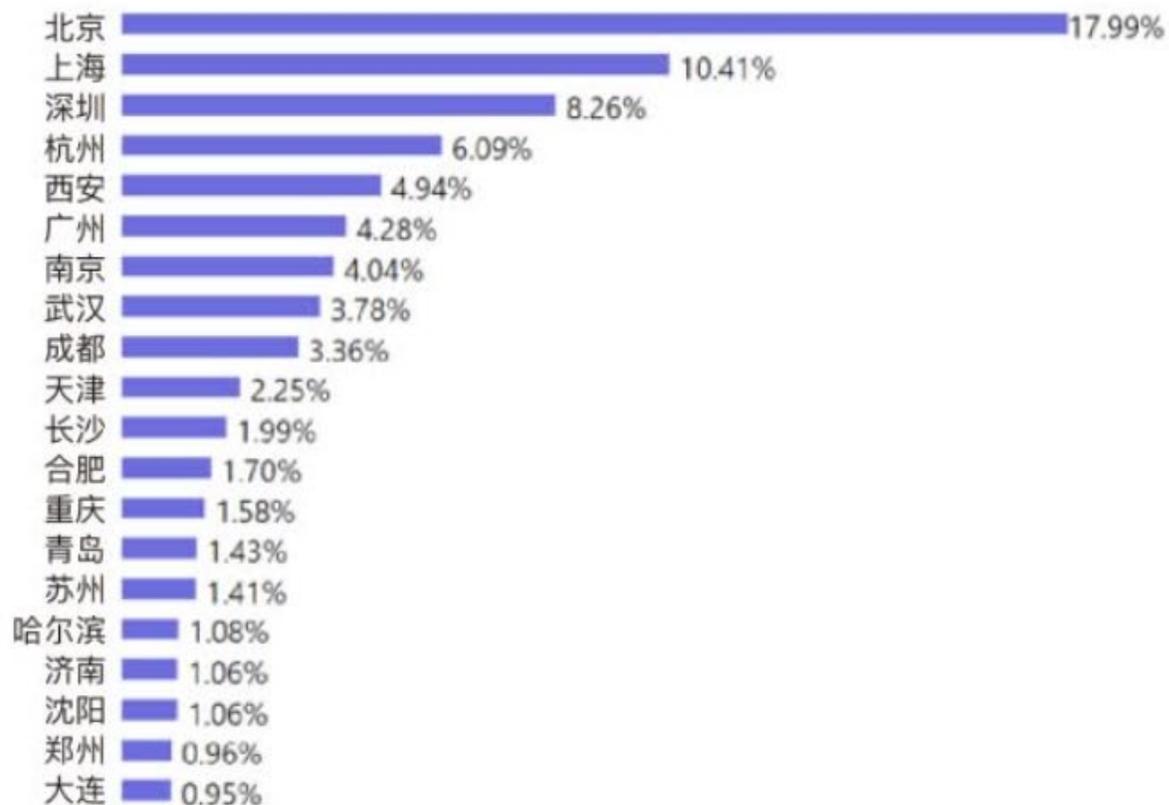


图11 中国计算机视觉人才城市分布TOP20



图5 2021年中国计算机视觉人才主要岗位薪资情况



Task List

- Control Task 1
- local-video-1
- surveillance task 1
- surveillance task 2
- surveillance task 3
- surveillance task 4
- surveillance task 5
- surveillance task 6
- surveillance task 7
- surveillance task 8
- surveillance task 9
- surveillance task 10
- surveillance task 11
- surveillance task 12
- surveillance task 13
- surveillance task 14
- surveillance task 15
- surveillance task 16
- surveillance task 17
- surveillance task 18



Captured

Today 1000 Month 1000 More 30



Target

Target Number 250 More 30

2016-04-03 11:57:53 local-video-1

Captured



Target

chenchuan...

2016-04-03 11:58:41 local-video-1

Captured



Target

shechuan...

2016-04-03 11:58:53 local-video-1

Captured



Target

shechuan

车辆抓拍



2016-05-05 08:30:24 丽岭大道与批发市场

显示牌位



行人抓拍



智能交通
商汤科技



地平线
Horizon Robotics

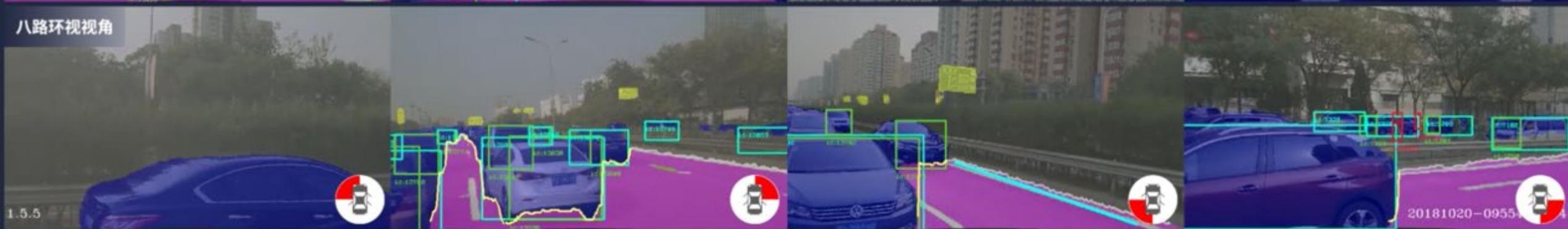
Matrix360°视觉感知方案

Matrix自动驾驶计算平台-- 基于征程2.0人工智能处理器架构

八路环视视角



八路环视视角



四路鱼眼视角



赋能万物，让人们的生活更安全，更美好



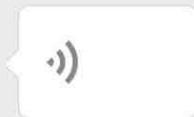
OpenClaw

THE AI THAT ACTUALLY DOES THINGS.

< 微信



11:51 AM



2"



10"

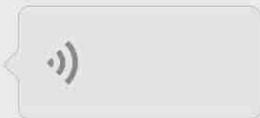


听筒播放

收藏

转文字

多选



5"



< 微信



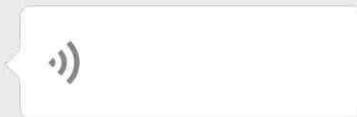
11:51 AM



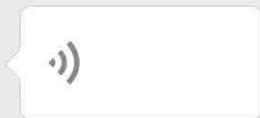
2"



10"



9"



5"

记得把明天的开会资料打印一下



✔ 转换完成



十四届全国人大四次会议
全国政协十四届四次会议

2026年

政府工作报告解读

打造智能经济新形态。深化拓展“人工智能+”，促进新一代智能终端和智能体加快推广，推动重点行业领域人工智能商业化规模化应用，培育智能原生新业态新模式。支持人工智能开源社区建设，促进开源生态繁荣。实施超大规模智算集群、算电协同等新基建工程，加强全国一体化算力监测调度，支持公共云发展。加快发展卫星互联网。打造“5G+工业互联网”升级版。深化数据资源开发利用，健全数据要素基础制度，建设高质量数据集。完善人工智能治理。

课程大纲



- 概述

- 模型

- 前馈神经网络
- 卷积神经网络
- 循环神经网络
- 注意力机制、Transformer
- 深度生成模型
- 深度强化学习

- 网络优化与正则化

- 深度学习计算系统

- 深度学习应用

- 计算机视觉
- 自然语言处理

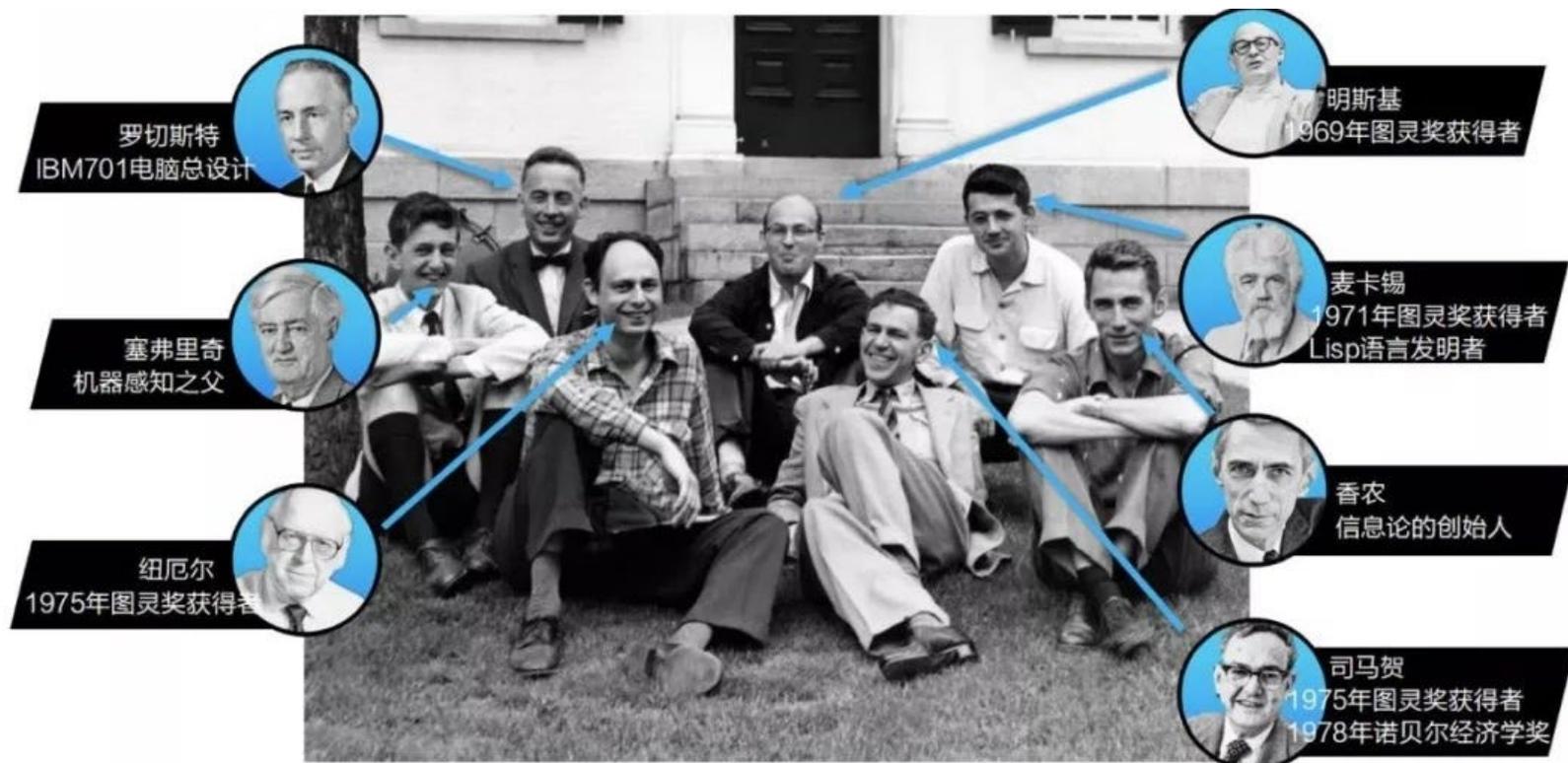
- 大语言模型

- 大语言模型基本原理
- 训练与微调
- 应用

人工智能与机器学习

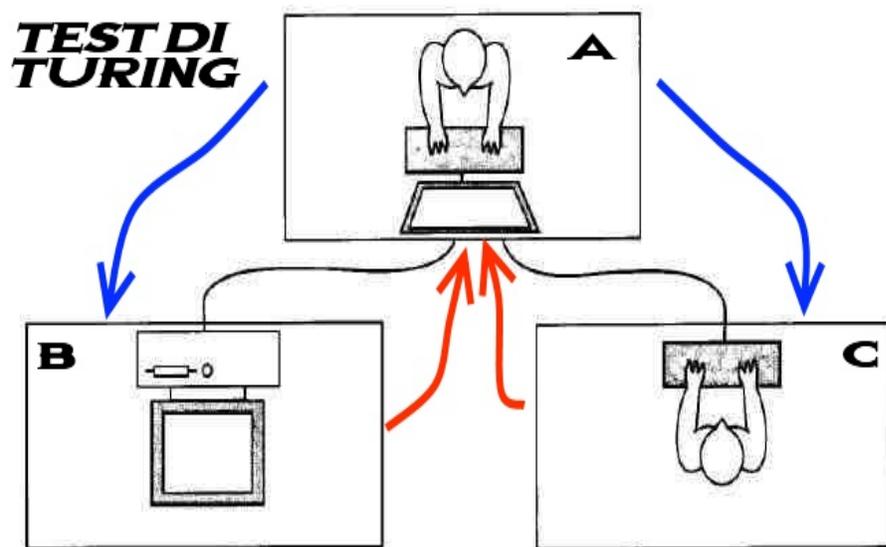


- **人工智能 (Artificial Intelligence, AI) 就是让机器具有人类的智能。**
 - “计算机控制” + “智能行为”
 - “人工智能就是要让机器的行为看起来就像是人所表现出的智能行为一样。” ——John McCarthy
- **人工智能学科诞生的是1956年的达特茅斯 (Dartmouth) 会议。**



• 图灵测试

- “一个人在不接触对方的情况下，通过一种特殊的方式，和对方进行一系列的问答。如果在相当长时间内，他无法根据这些问题判断对方是人还是计算机，那么就可以认为这个计算机是智能的。” ——Alan Turing (1950)

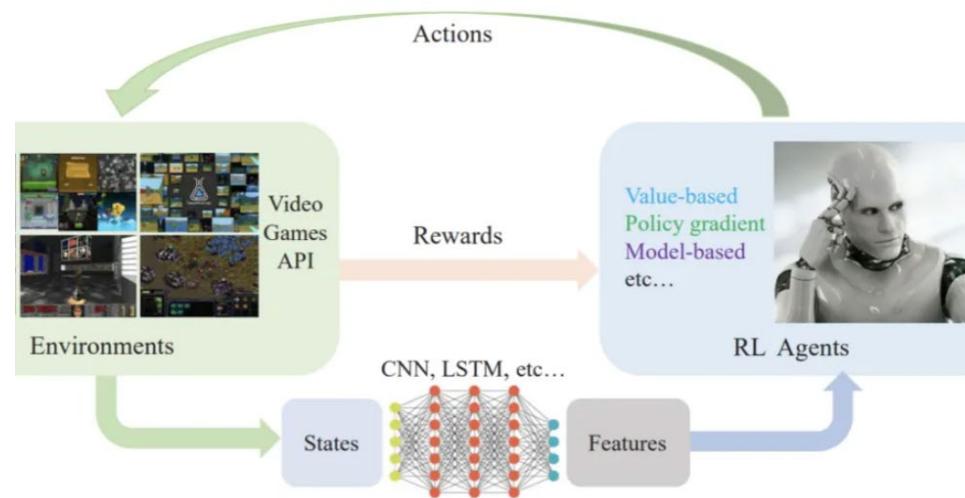
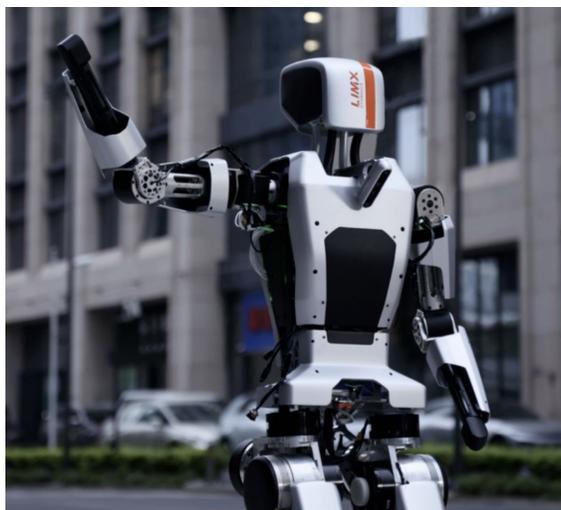


Alan Turing

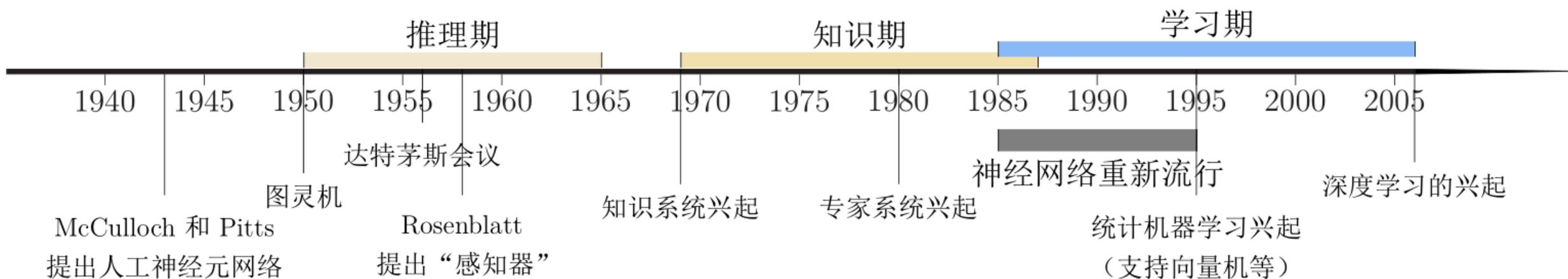
人工智能的研究领域

● 让机器具有人类的智能

- 机器感知（计算机视觉、语音信息处理）
- 学习（模式识别、机器学习、强化学习）
- 语言（自然语言处理）
- 记忆（知识表示）
- 决策（规划、数据挖掘）

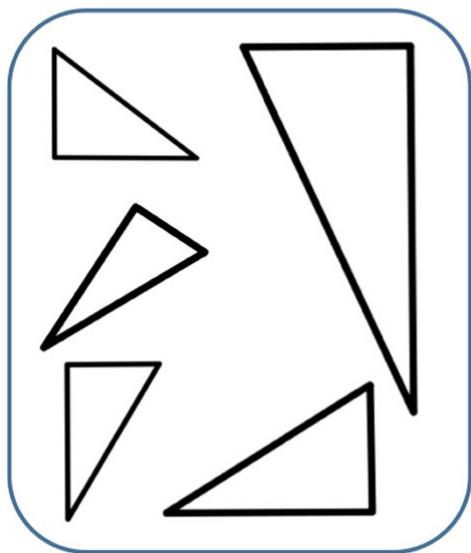


人工智能的发展历史

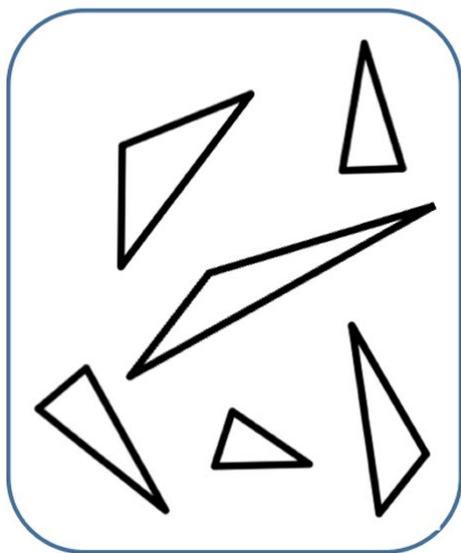


如何开发一个人工智能系统?

What's My Rule?



Yes/True



No/False



2	6	8	9	3	4	7	5	6
3	4	7	9	5	5	6	7	2
5	8	7	0	9	4	3	5	4
5	2	3	4	9	5	6	7	8

机器学习 \approx 构建一个映射函数

- 语音识别

$$f(\text{语音波形}) = \text{"你好"}$$

- 图像识别

$$f(\text{猫咪照片}) = \text{"猫"}$$

- 围棋

$$f(\text{围棋棋盘}) = \text{"5-5" (落子位置)}$$

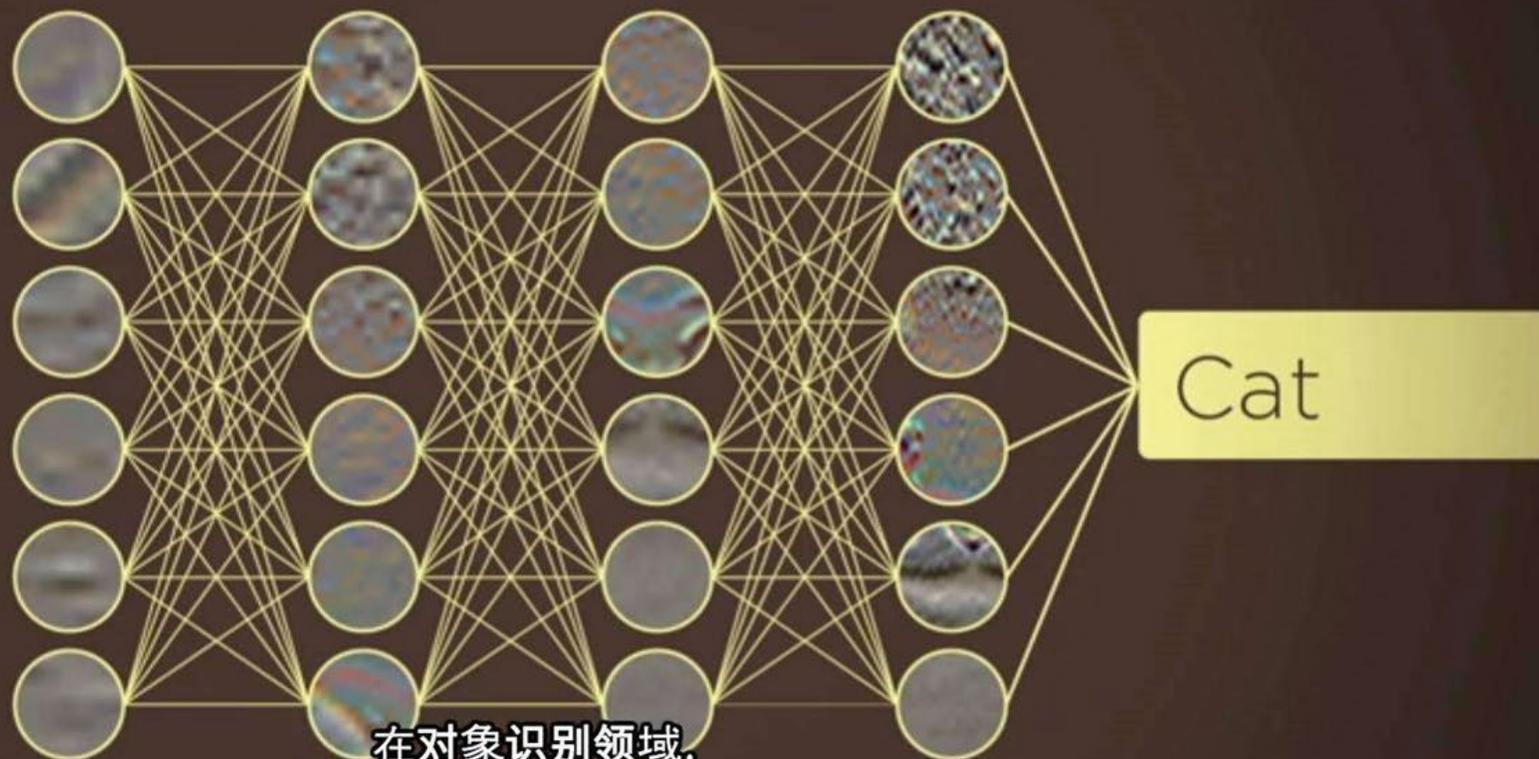
- 对话系统

$$f(\text{"你好"}) = \text{"今天天气真不错"}$$

用户输入

机器输出

机器学习 \approx 构建一个映射函数



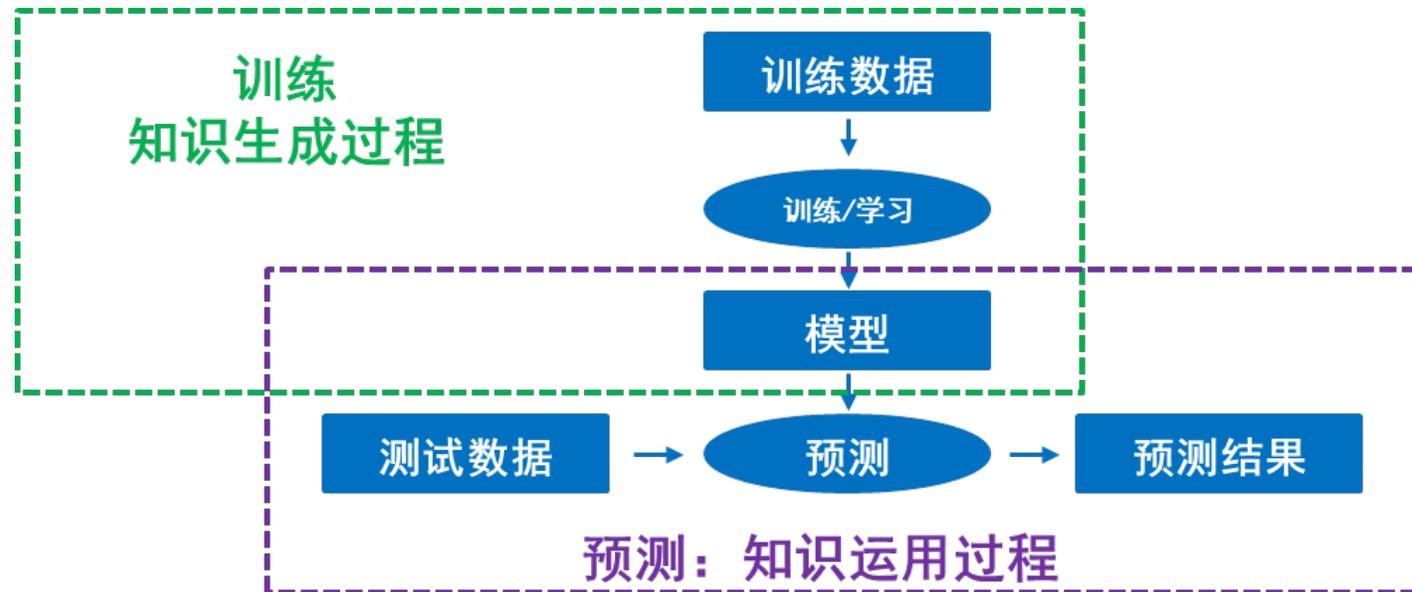
在对象识别领域，
产生了激动人心的新成果。

机器学习的定义

- **Tom Mitchell (1998) well-posed Learning Problem:**

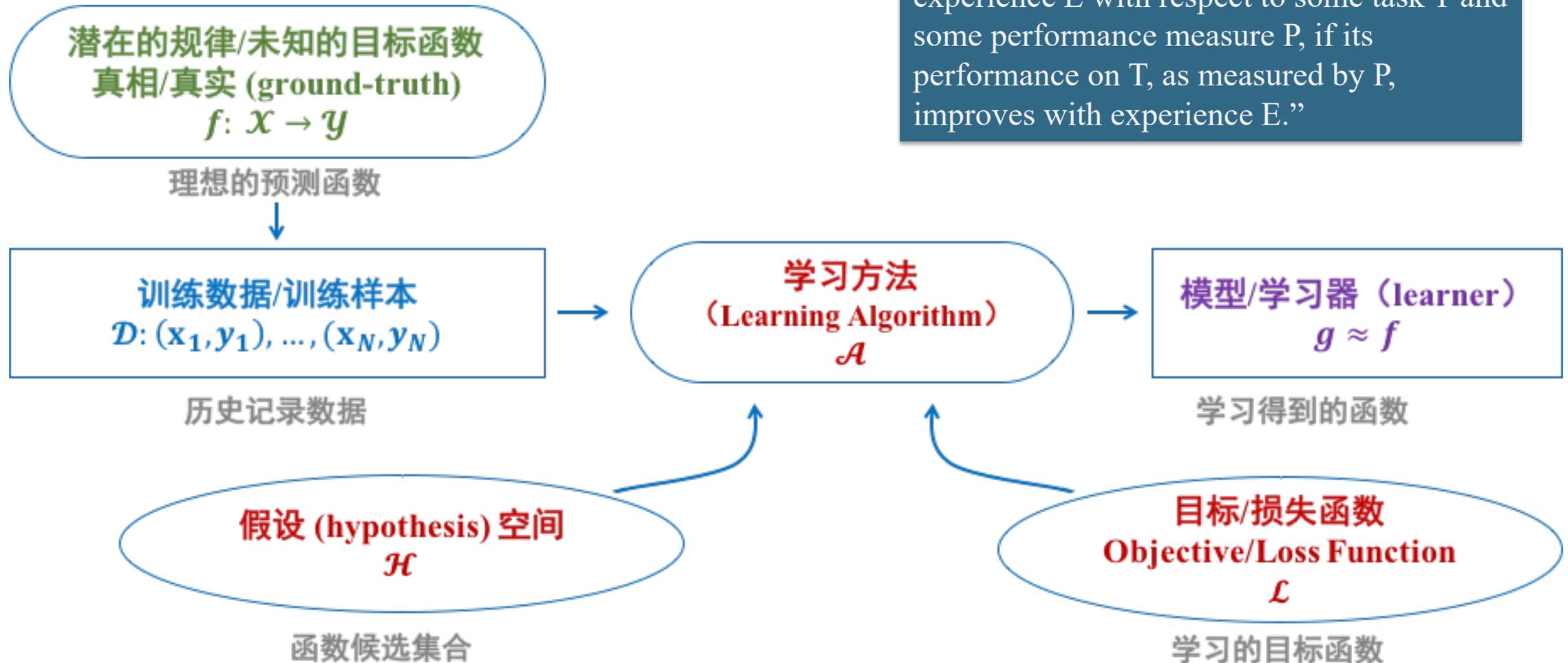
- A computer program is said to learn from experience E with respect to some task T and some performance measure P , if its performance on T , as measured by P , improves with experience E .

- **机器学习是一种通过利用数据，训练出模型，然后使用模型预测的方法。**



机器学习的定义

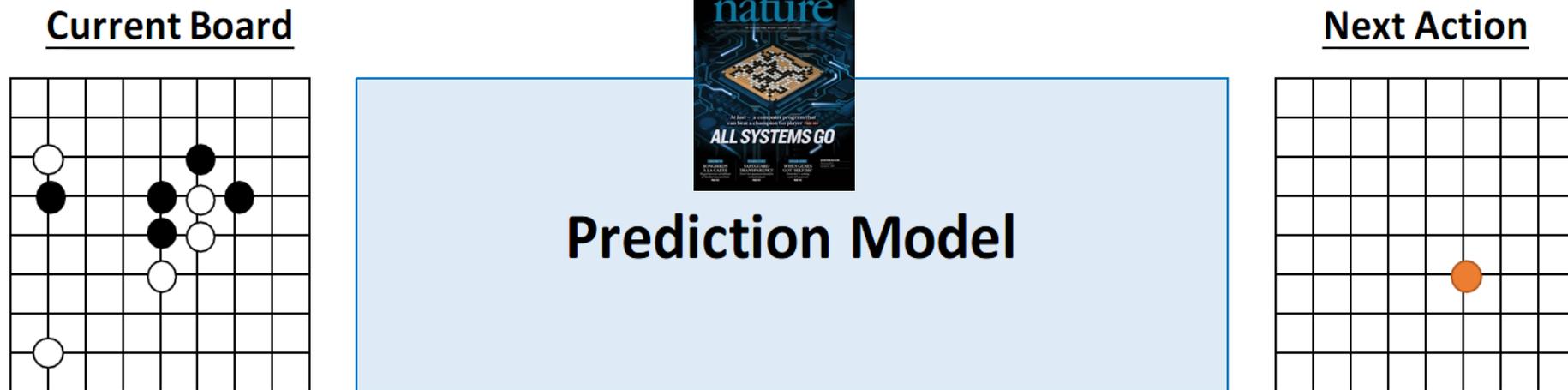
“A computer program is said to *learn* from experience E with respect to some task T and some performance measure P , if its performance on T , as measured by P , improves with experience E .”



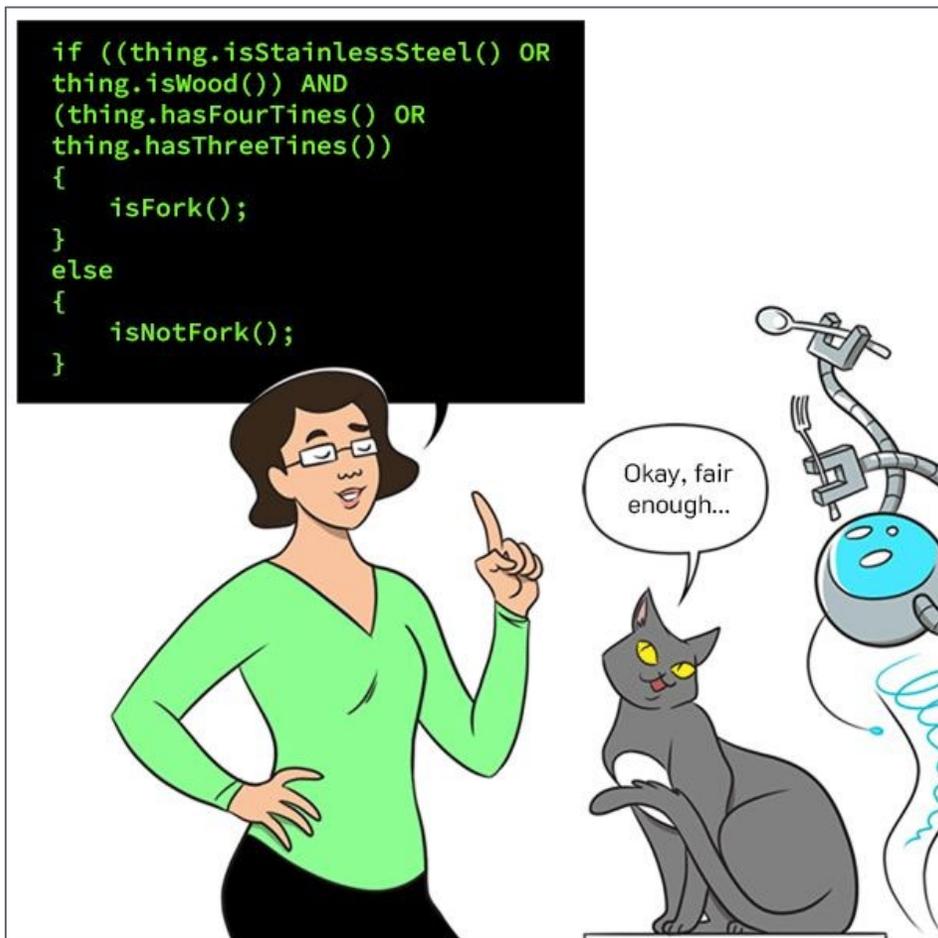
机器学习的定义

- **Arthur Samuel (1959). Machine Learning:**

- Field of study that gives computers the ability to learn without being explicitly programmed.
- 一种能够赋予机器学习的能力以此让它完成**直接编程无法完成**的功能的方法。



机器学习的定义

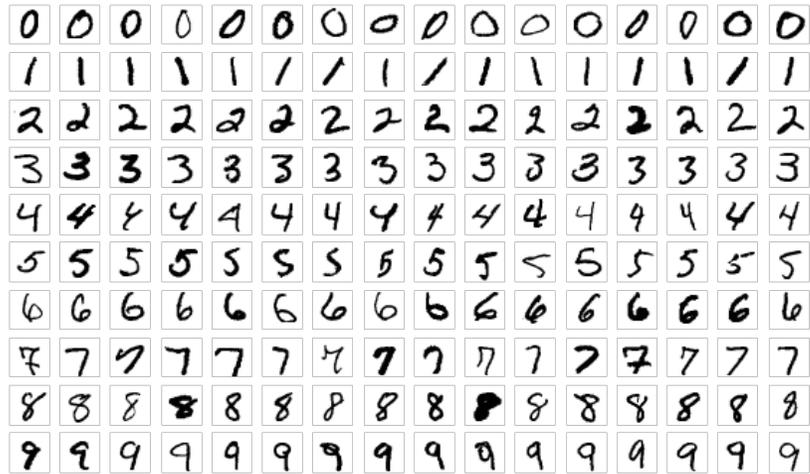


学习任务有哪些类型？

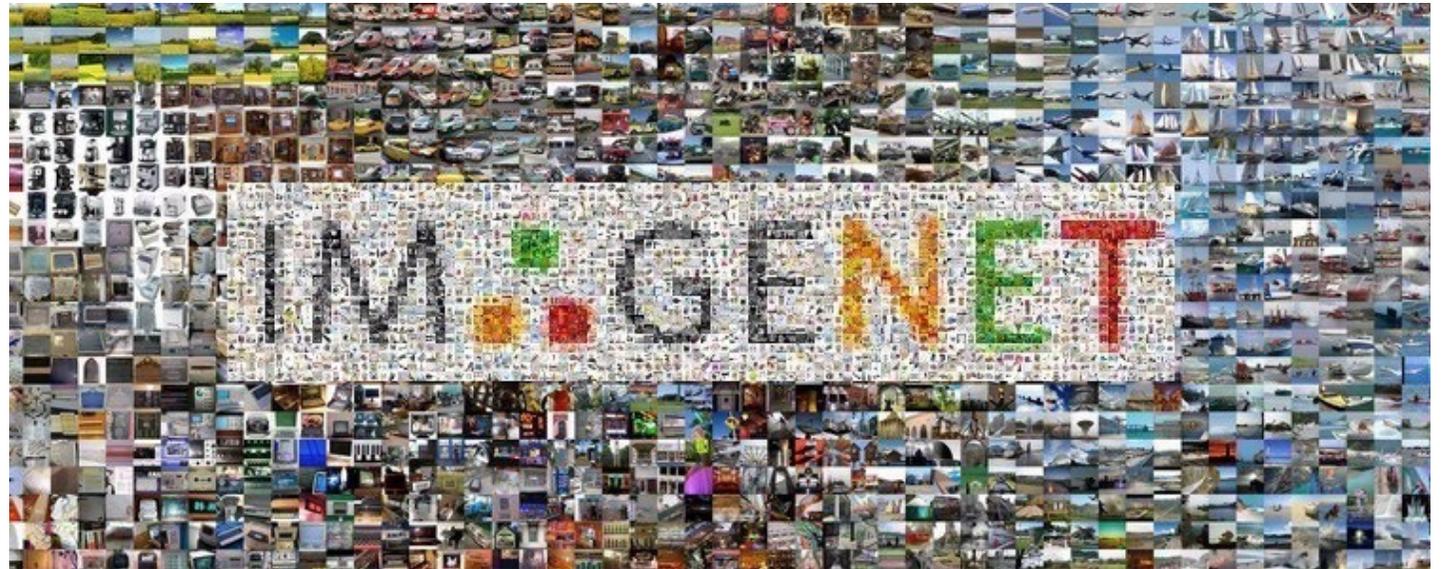


回归与分类

- 回归：估计连续值
- 分类：预测离散类别



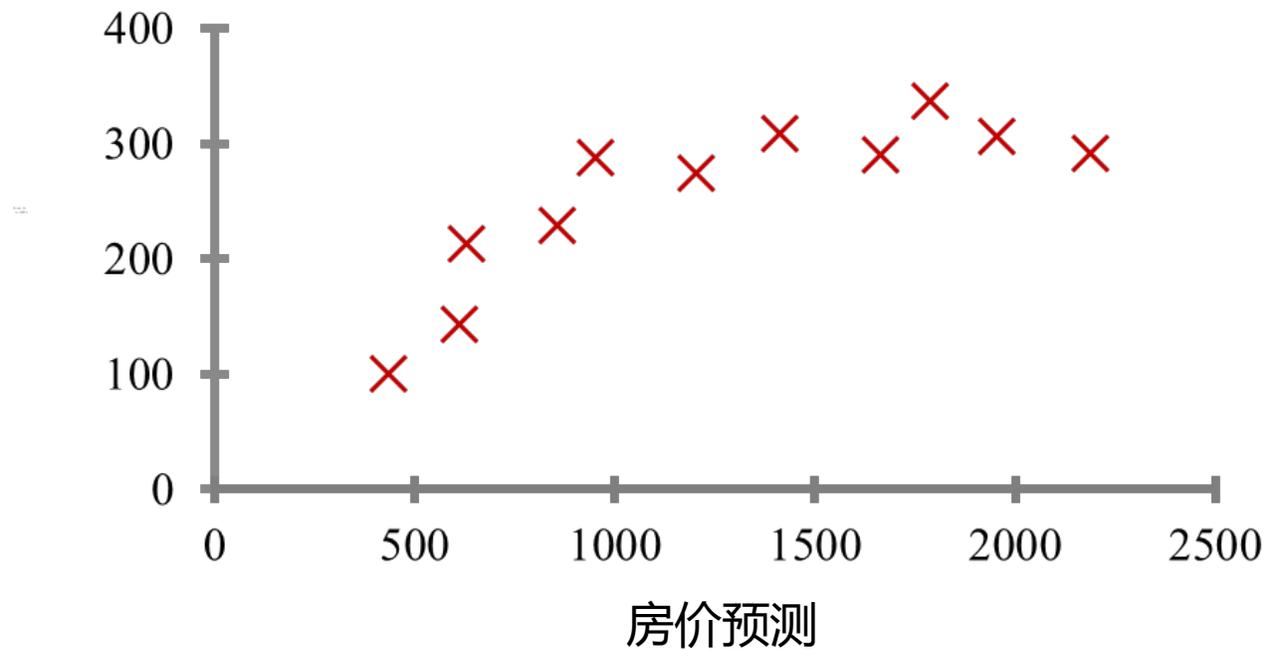
MNIST：对手写数字进行分类（10 类）



ImageNet：对自然对象进行分类（1000 类）

回归与分类

- 回归：估计连续值
- 分类：预测离散类别



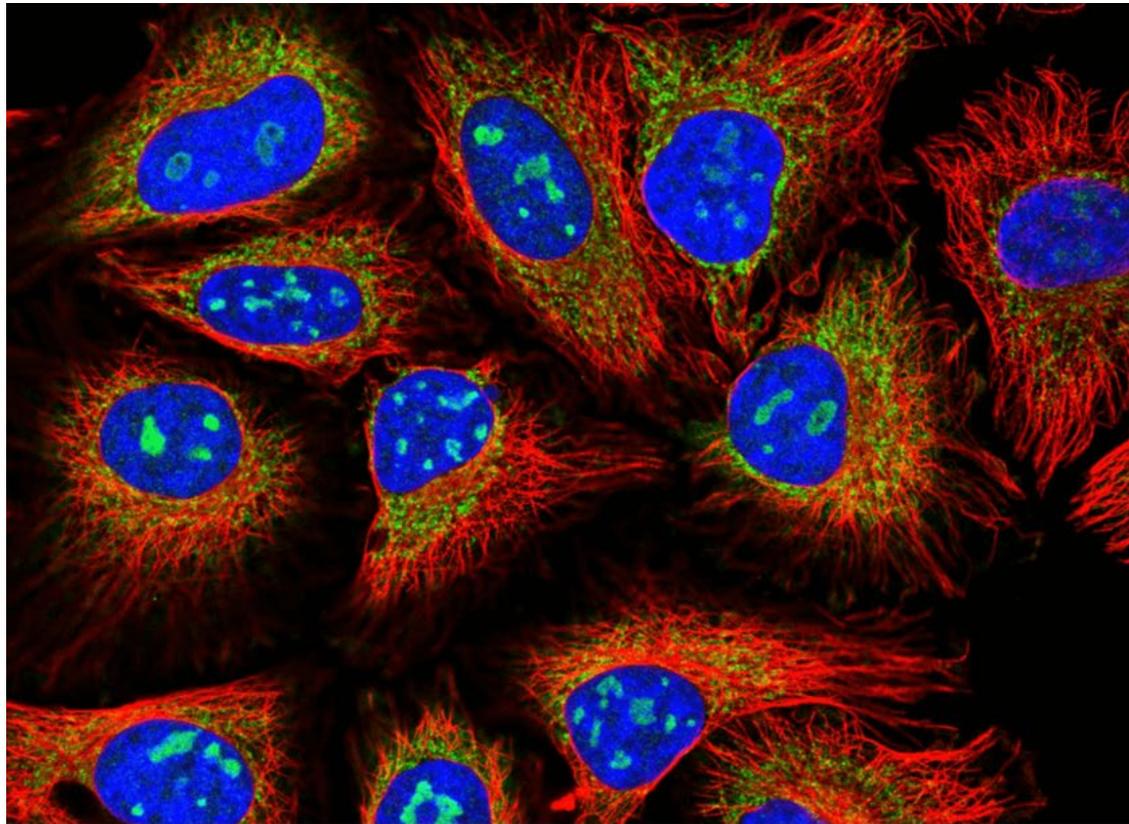
Supervised Learning

- “right answers” given

Regression:

- Predict continuous valued output (price)

- 将人类蛋白质显微镜图像分为28类



0. Nucleoplasm
1. Nuclear membrane
2. Nucleoli
3. Nucleoli fibrillar
4. Nuclear speckles
5. Nuclear bodies
6. Endoplasmic reticu
7. Golgi apparatus
8. Peroxisomes
9. Endosomes
10. Lysosomes
11. Intermediate fila
12. Actin filaments
13. Focal adhesion si
14. Microtubules
15. Microtubule ends
16. Cytokinetic brida

<https://www.kaggle.com/c/human-protein-atlas-image-classification>

- 将维基百科上的恶语评论分为7类

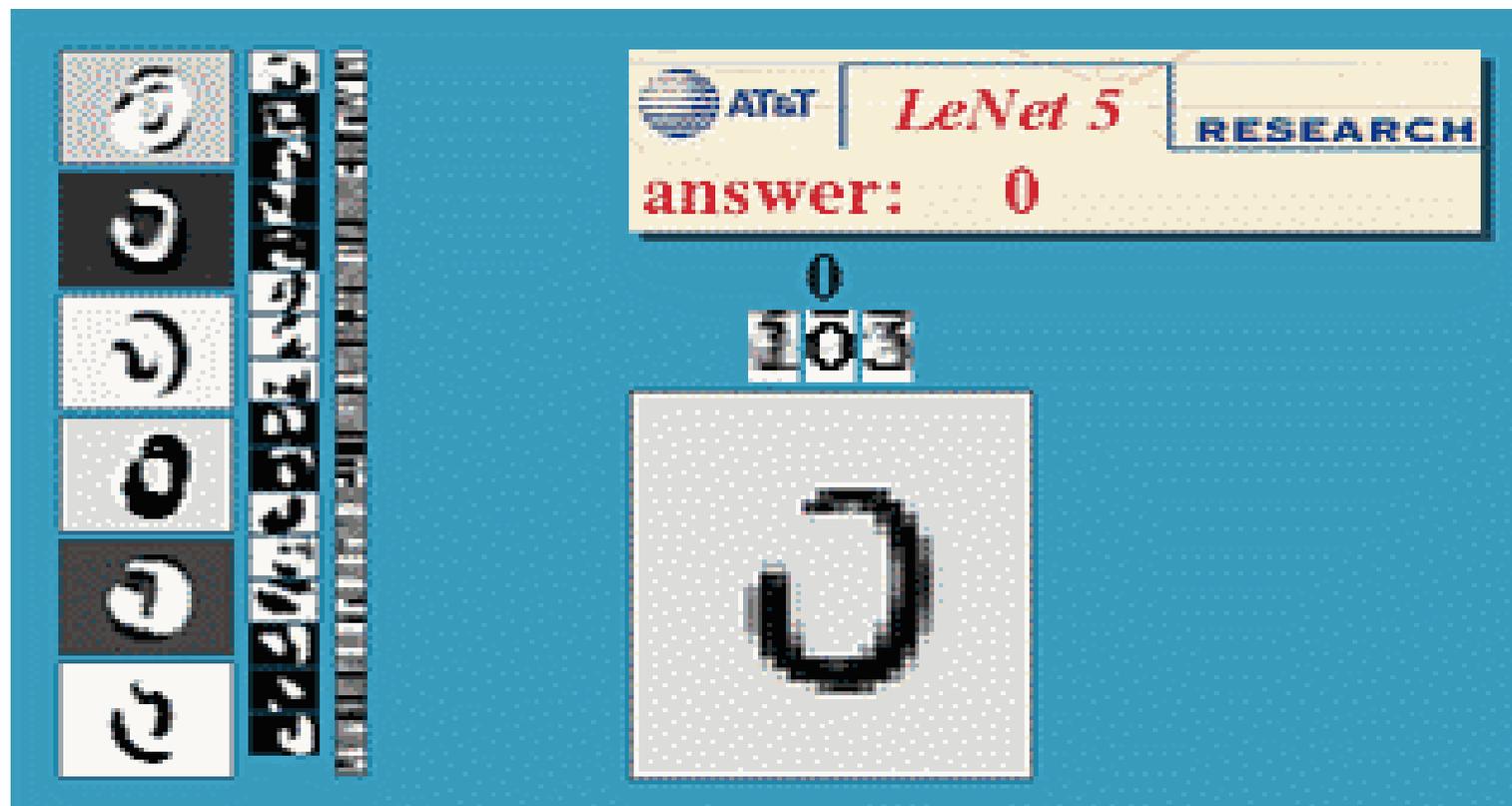
comment_text	toxic	severe_toxic	obsc
Explanation\nWhy the edits made under my usern...	0	0	0
D'aww! He matches this background colour I'm s...	0	0	0
Hey man, I'm really not trying to edit war. It...	0	0	0
"\nMore\nI can't make any real suggestions on ...	0	0	0
You, sir, are my hero. Any chance you remember...	0	0	0

<https://www.kaggle.com/c/jigsaw-toxic-comment-classification-challenge>

常见的学习任务

- 数字识别是_____问题?

- A. 二分类 B. 多分类 C. 回归 D. 排序



常见的学习任务

● 网络搜索是_____问题?

- A. 二分类
- B. 多分类
- C. 回归
- D. 排序

The image shows a Baidu image search interface for the term "机器学习" (Machine Learning). The search results are displayed in a grid of images. The top row includes a woman's face with circuitry, a "Machine Learning" title with data streams, a neural network visualization, and a cloud services diagram. The bottom row includes a flowchart of the machine learning process, a word cloud, a robot holding books, a man at a computer, and a brain with neural connections.

新闻 网页 贴吧 知道 音乐 图片 视频 地图 百科 文库

Baidu 图片 机器学习 百度一下 图片筛选

相关搜索: 人工智能 机器学习 机器学习 周志华 机器学习 素材 机器学习模型 机器学习应用 机器学习流程 模式识别与机器学习 机器学习算法 机器学习 训练 图解机器学习 机器学习应用领域

unknown target function $f: X \rightarrow Y$
ideal credit approval formula

training examples $D: (x_1, y_1), \dots, (x_n, y_n)$
(historical records in bank)

learning algorithm A

hypothesis set H
(set of candidate formula)

final hypothesis $g \approx f$
(learned formula to be used)

questions algorithm theoretical
information network behavior
complex objects machine
structure web algorithms results
different give people interactions many
patterns social extract analysis
deal use amounts social extract analysis
want often machine learning guarantees
describe classification able machine
graphs clustering important

Machine Learning

常见的学习任务

- 图像生成与翻译是_____问题?

- A. 二分类 B. 多分类 C. 回归 D. 排序



机器学习方法分类（怎样学？）



机器学习定义

- 例：房价预测

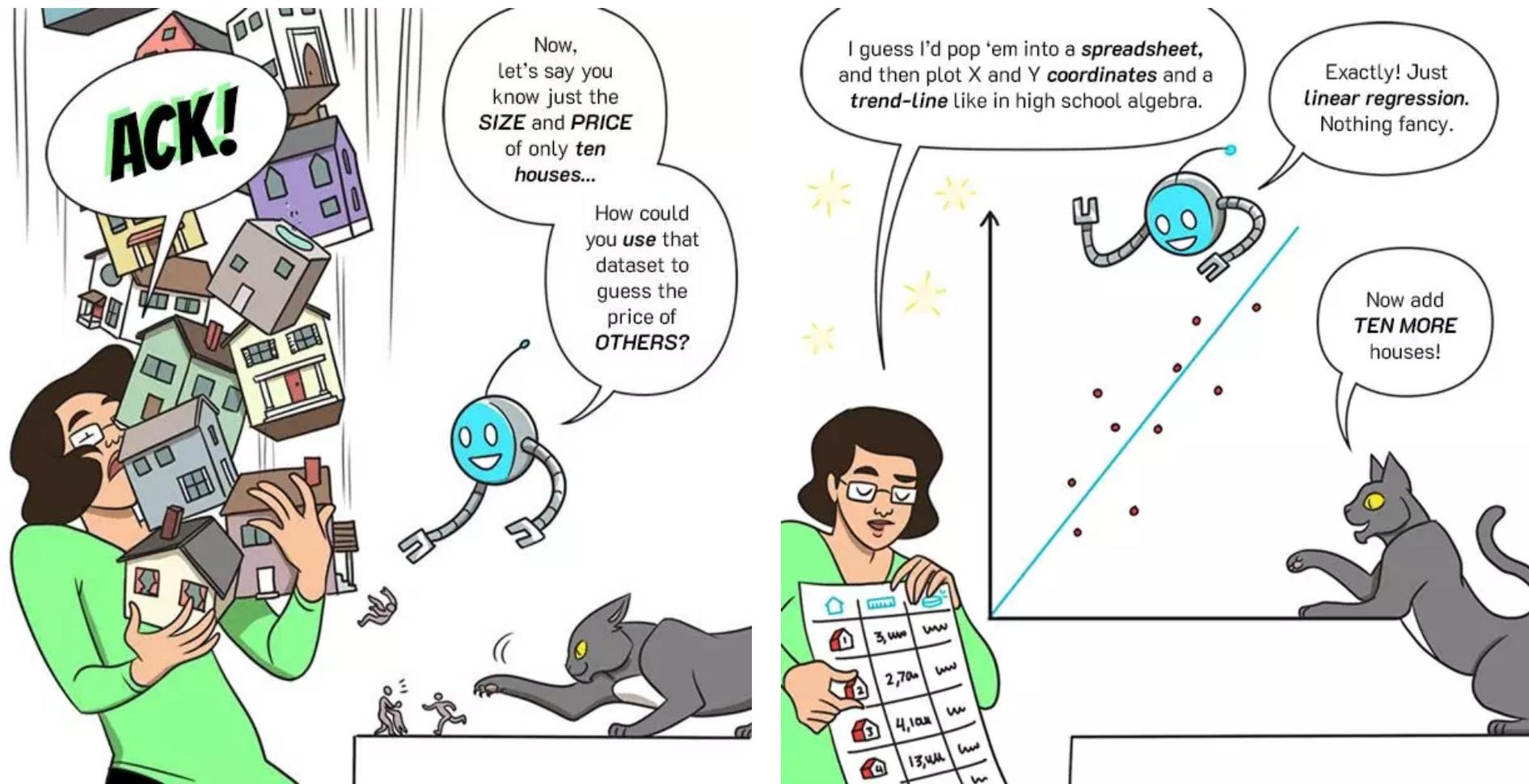
- 输入x：特征

- 输出y：房价

- 映射函数 f

- $y = f(x)$

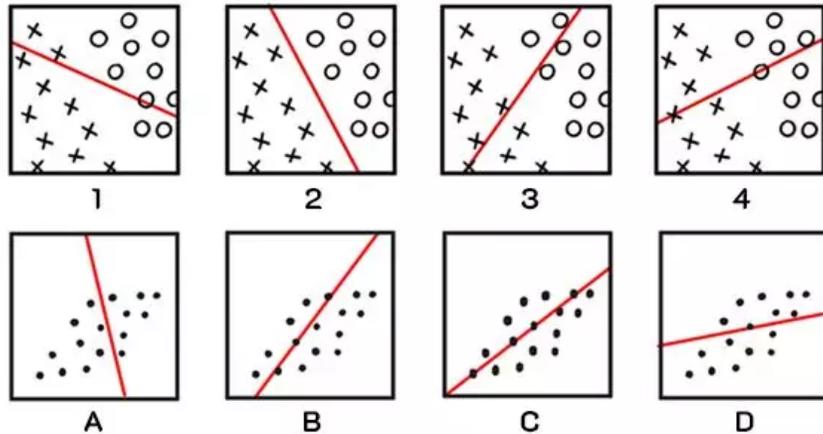
- $f: x \rightarrow y$



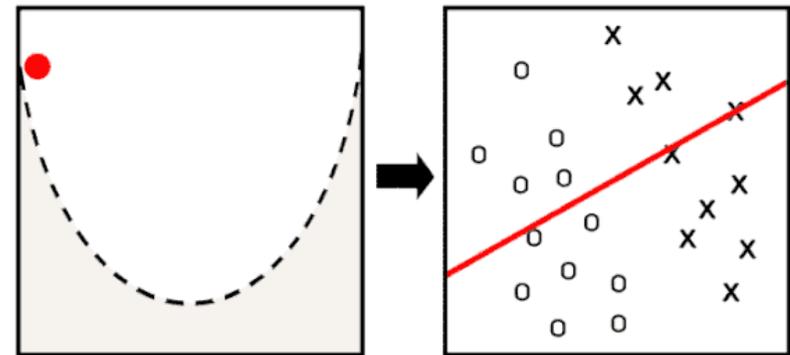
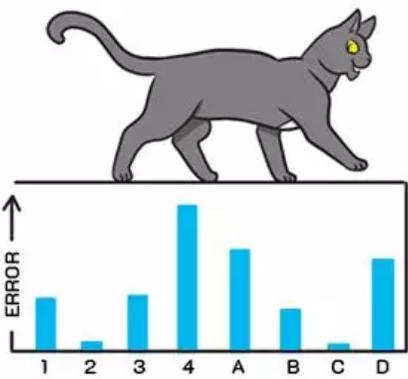
重要概念

- 优化 Optimize
- 极值
- 局部极值

- 假设 Hypothesis
- 误差 Error



Whether for classification or regression problems, ML typically starts with a **HYPOTHESIS** and, of course, some hypotheses will prove **MORE WRONG** than others.

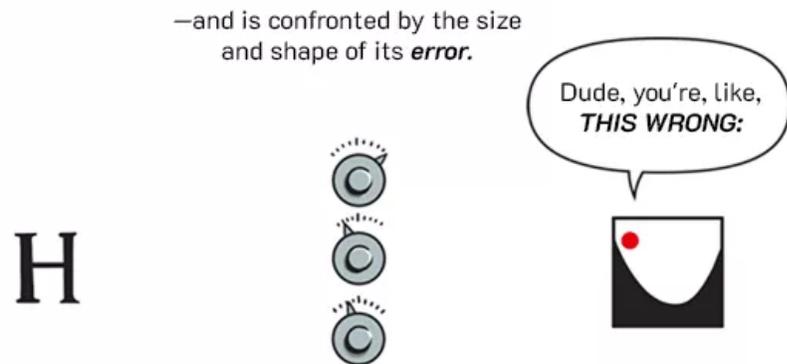
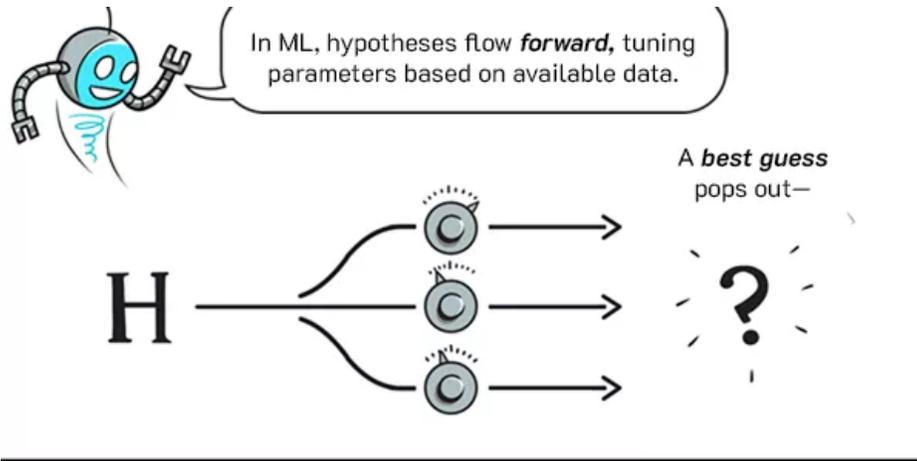


Think **descending = solving**, plain and simple. Gradient descent is the learning process, enacted in **space**.

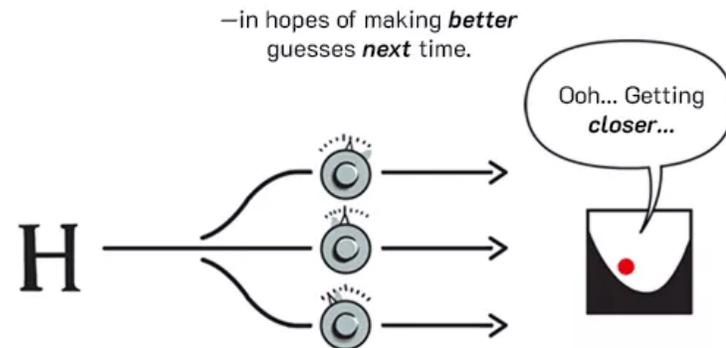
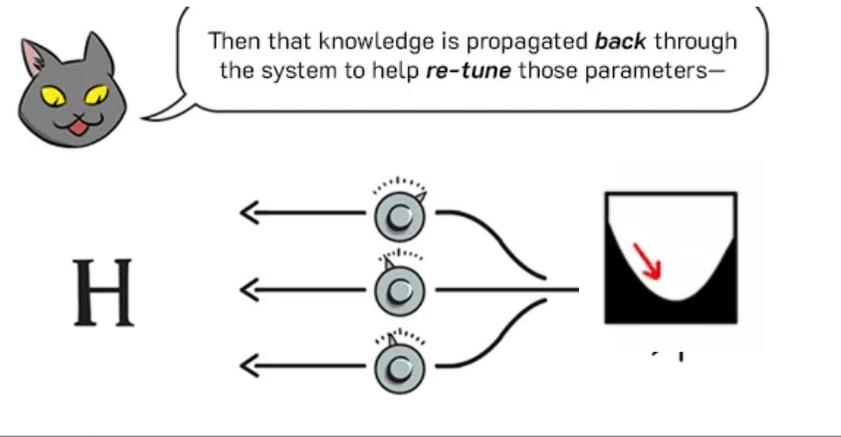


重要概念

- 反馈 Backward
- 优化 Optimization
- 调节 Tune



- 前馈 Forward
- 推理 Inference
- 预测 Prediction



学习算法的类型

● 有监督学习 (Supervised Learning)

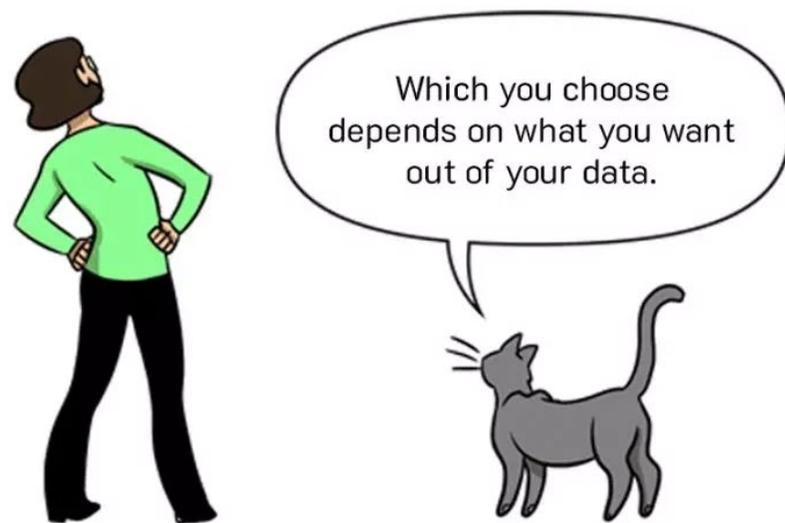
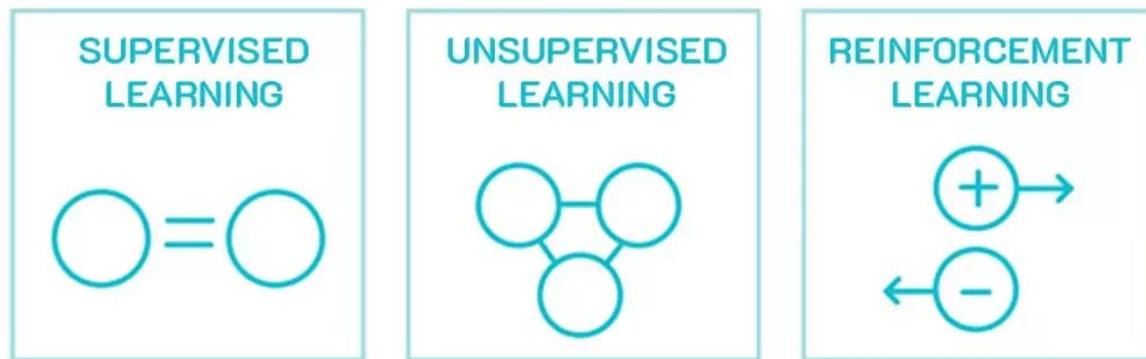
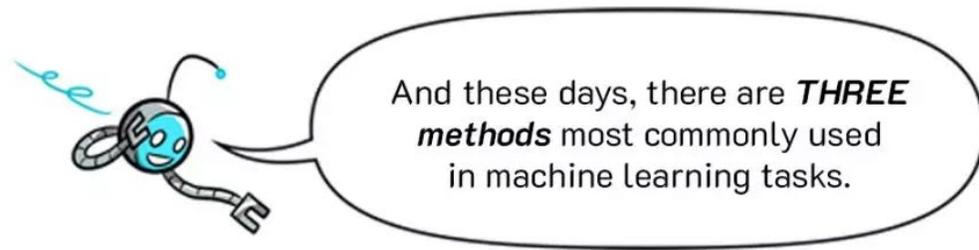
- 要求大量的有标记数据
- 如：图像分类、人脸识别

● 无监督学习

- 无标注，学习数据本身的规律
- 如：电影推荐、音乐推荐

● 强化学习

- 在与环境的交互中学习
- 如：游戏AI、机器人



"好" 数据!

Supervised learning is used in image classification—

—speech and text recognition—

—medical diagnostics—

—and fraud detection, to name a few...

!	\$ -1420.75
	¥ + 31.45
	£ - 2.74
	\$ -565.21
	€ +103.89

FORK

CAT

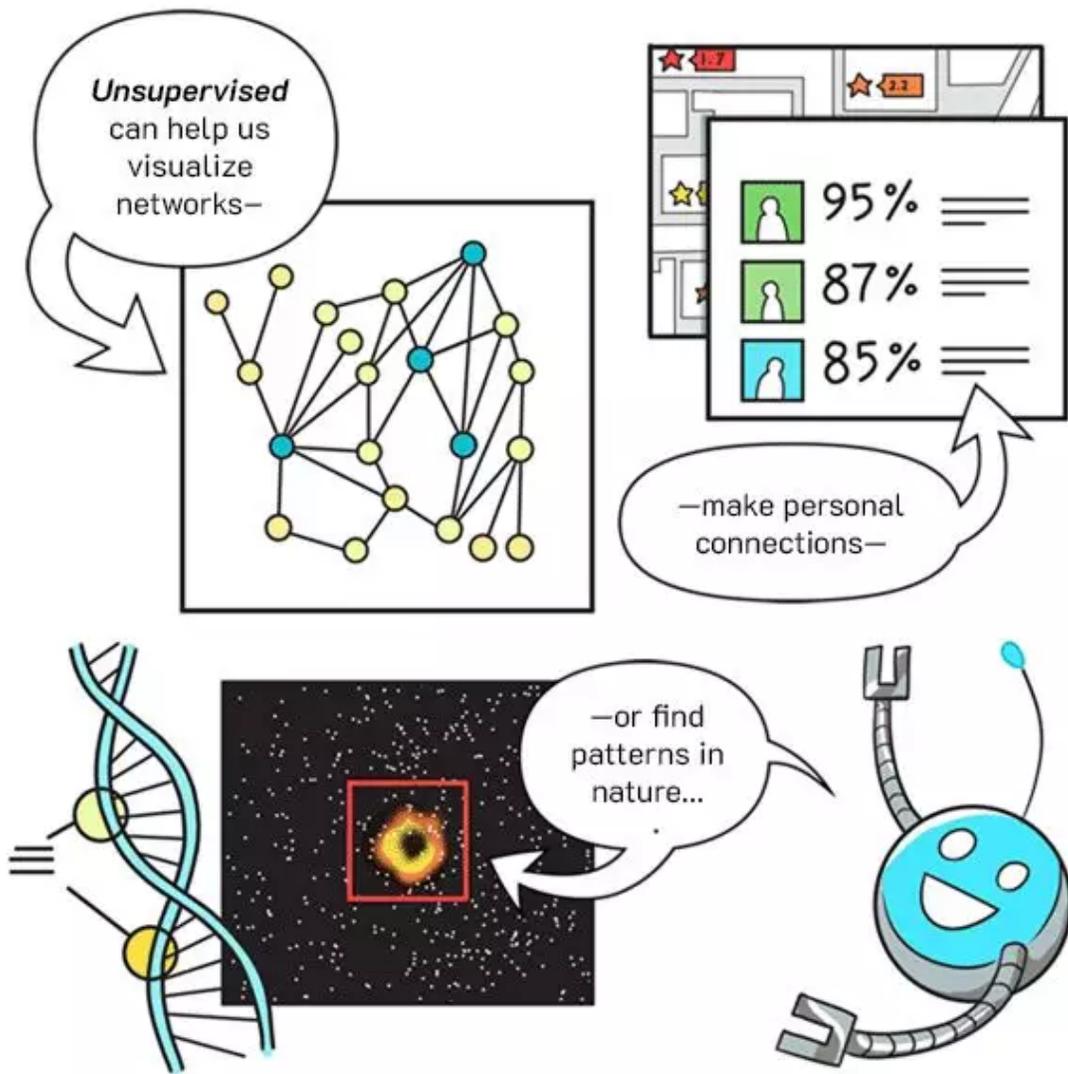
GRAPEFRUIT

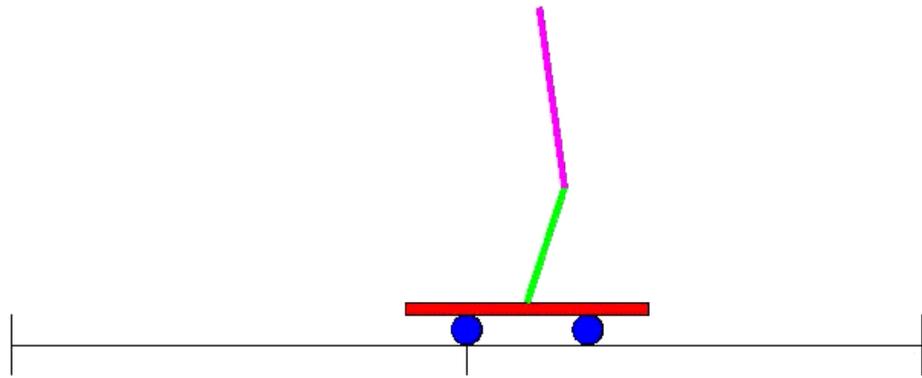
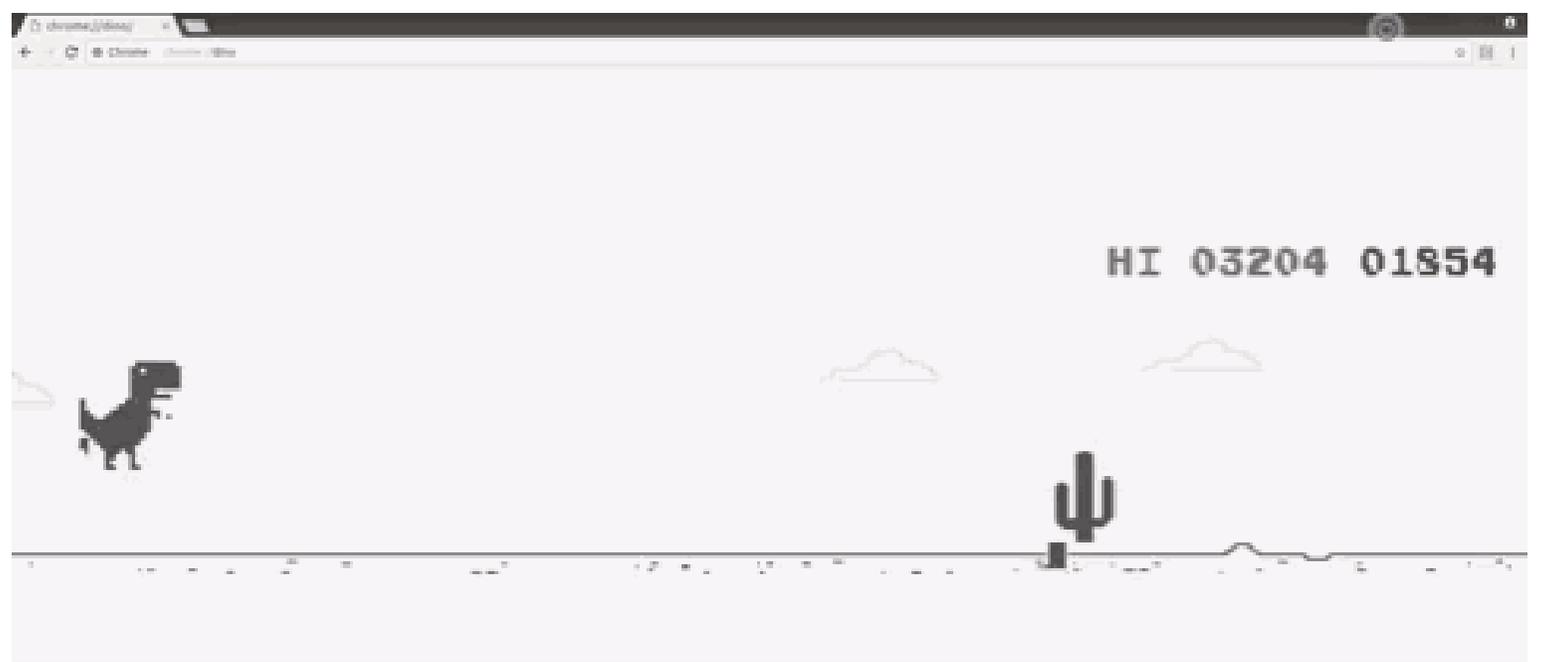
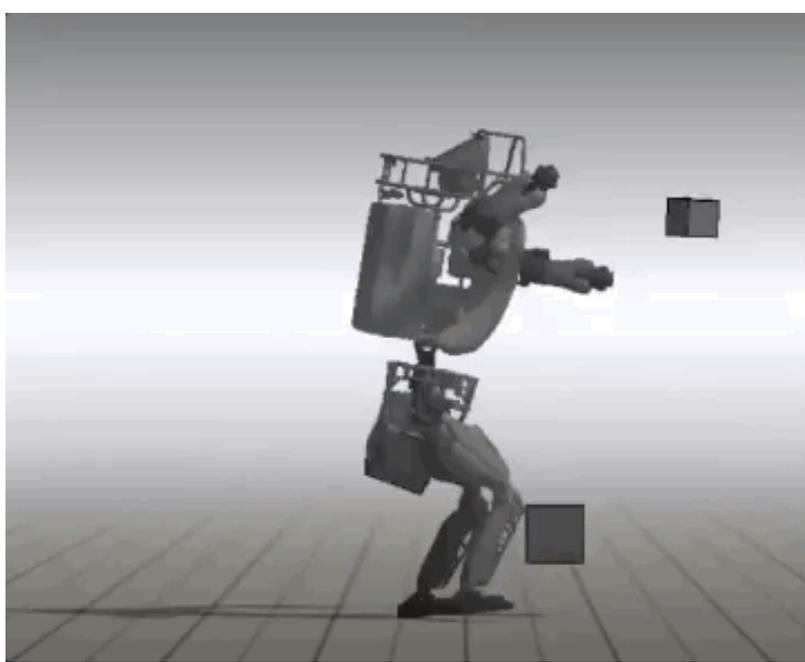
PIPE

Ceci n'est pas une pipe.

If your data is too sloppy, homogenous, or scarce, your results could fall apart in the real world...







- 请用自己的话解释什么是人工智能，什么是机器学习。
- 什么是图灵测试 (The Turing Test) ?
- 辨析下面的问题是回归问题还是分类问题：
 - 工程师正在训练模型，用于自动驾驶汽车的制动系统。模型的输入是一组来自前视摄像头的实时画面。模型需要计算一个关键指标：当前车速下，本车与前方障碍物之间的“碰撞时间”。一旦系统计算出“碰撞时间”低于 1.5秒，则触发红灯警报；如果低于 0.8秒，则直接触发刹车。训练一个模型来直接输出这个“碰撞时间”数值，属于什么类型的问题？
 - 银行开发了一套实时交易监测系统。当一笔交易发生时，模型需要输出风险评级，以估计交易涉嫌欺诈的可能性。这属于哪种类型的问题？

● 有监督学习、无监督学习和强化学习辨析：

- DeepMind 开发的 AlphaGo 在战胜人类围棋冠军的过程中，使用了多种技术组合。考虑以下两种训练阶段：

阶段一： 利用人类专业棋手对弈的历史棋谱数据，让模型学习在给定当前棋盘布局的情况下，人类最有可能在哪个位置落子。这相当于让模型模仿人类高手。

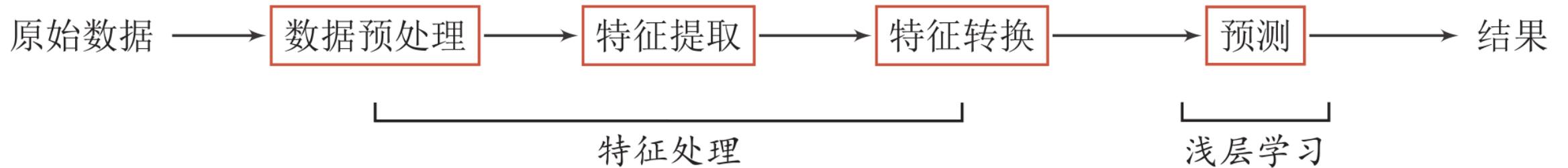
阶段二： 让AlphaGo和自己（或者之前的版本）对弈。对弈过程中，没有任何人类棋谱作为指导，只有最终的胜负结果（赢或输）作为反馈信号。模型不断调整策略，试图找到能够最大化获胜概率的下法。

从机器学习的范式来看，阶段一和阶段二分别属于什么类型？

如何开发一个机器学习系统？



- 当我们用机器学习来解决一些模式识别任务时，一般的流程包含以下几个步骤：

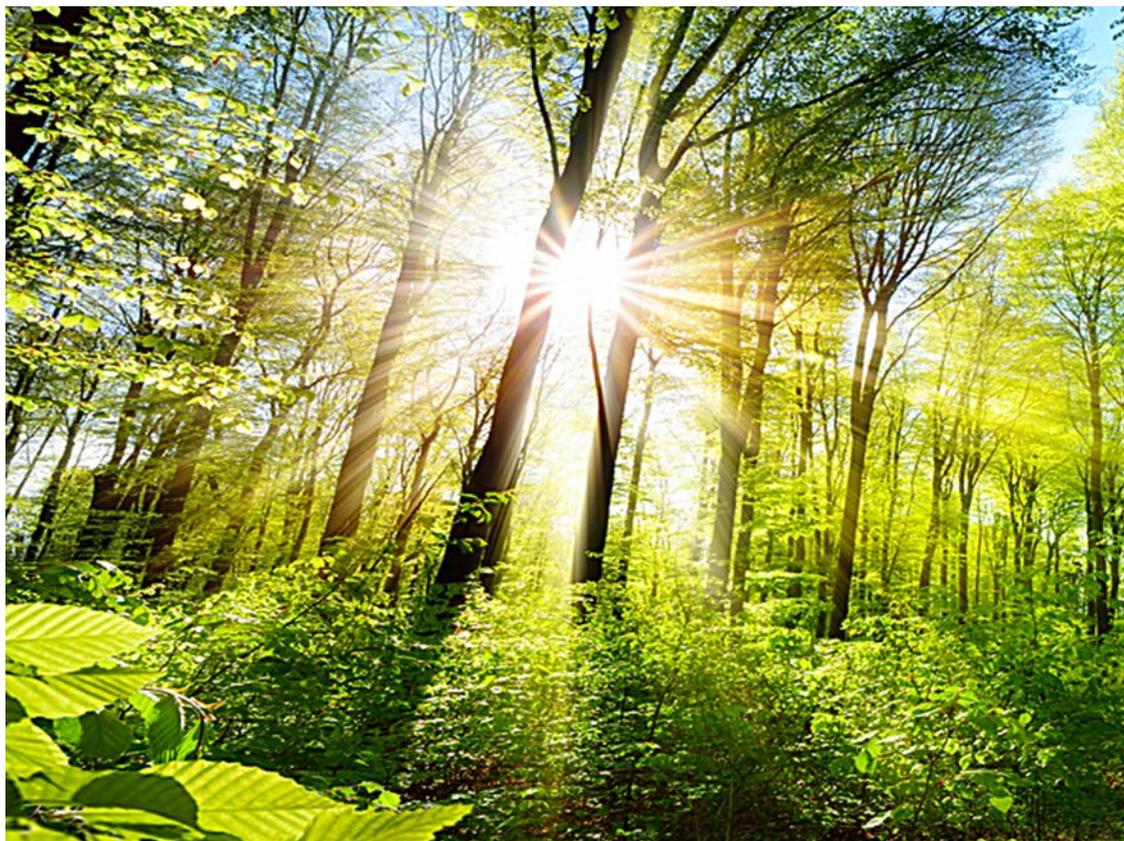


特征工程 (Feature Engineering)

- 浅层学习 (Shallow Learning)：不涉及特征学习，其特征主要靠人工经验或特征转换方法来抽取。

- 例如：图像纹理、颜色等特征

- 季节预测：夏天 vs. 秋天



语义鸿沟：人工智能的挑战之一

● 底层特征 VS 高层语义

- 人们对文本、图像的理解无法从字符串或者图像的底层特征直接获得



语义表示

符号表示

知识库规则

分布式表示

嵌入：压缩、低维、稠密向量

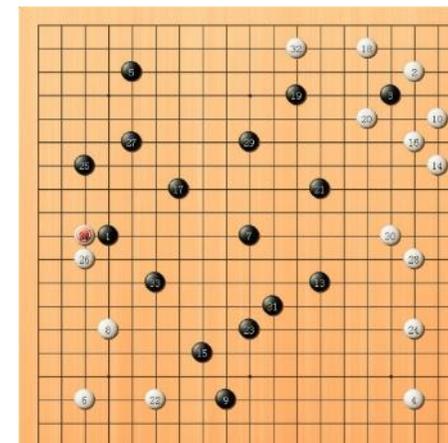
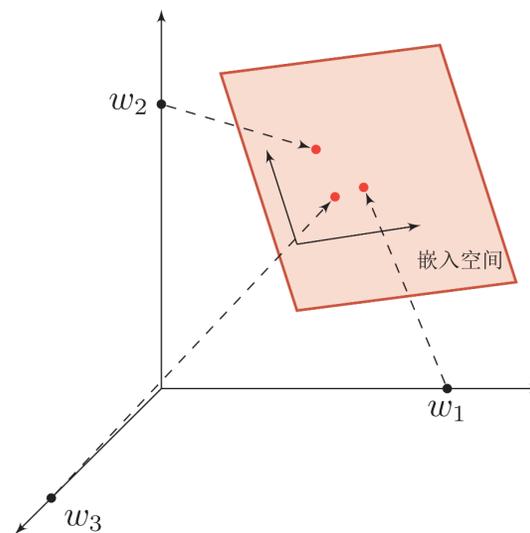
● 符号表示

- 离散表示、局部表示
- One-Hot向量

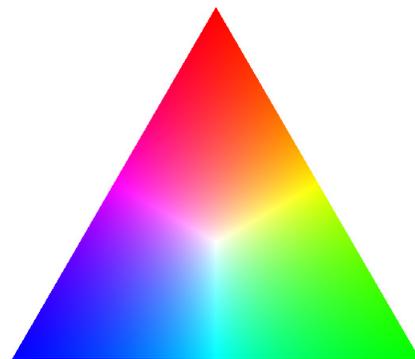
● 分布式(distributed)表示

- 压缩、低维、稠密向量
- 用 $O(N)$ 个参数表示 $O(2k)$ 区间
 - k 为非0参数, $k < N$

	离散表示	连续表示
A	[1 0 0 0]	[0.25 0.5]
B	[0 1 0 0]	[0.2 0.9]
C	[0 0 1 0]	[0.8 0.2]
D	[0 0 0 1]	[0.9 0.1]

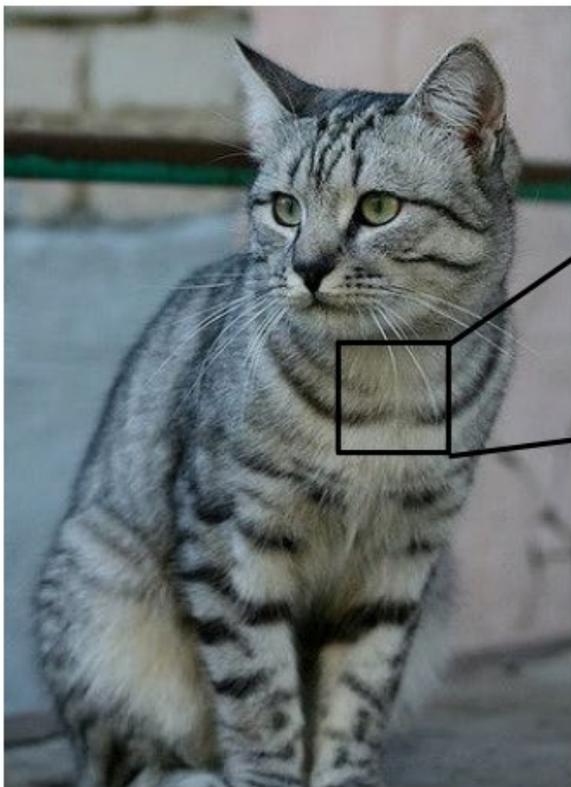


语义表示



颜色	局部表示	分布式表示
琥珀色	$[1, 0, 0, 0]^T$	$[1.00, 0.75, 0.00]^T$
天蓝色	$[0, 1, 0, 0]^T$	$[0.00, 0.5, 1.00]^T$
中国红	$[0, 0, 1, 0]^T$	$[0.67, 0.22, 0.12]^T$
咖啡色	$[0, 0, 0, 1]^T$	$[0.44, 0.31, 0.22]^T$

The Problem: Semantic Gap



This image by Nikita is licensed under [CC-BY 2.0](https://creativecommons.org/licenses/by/2.0/)

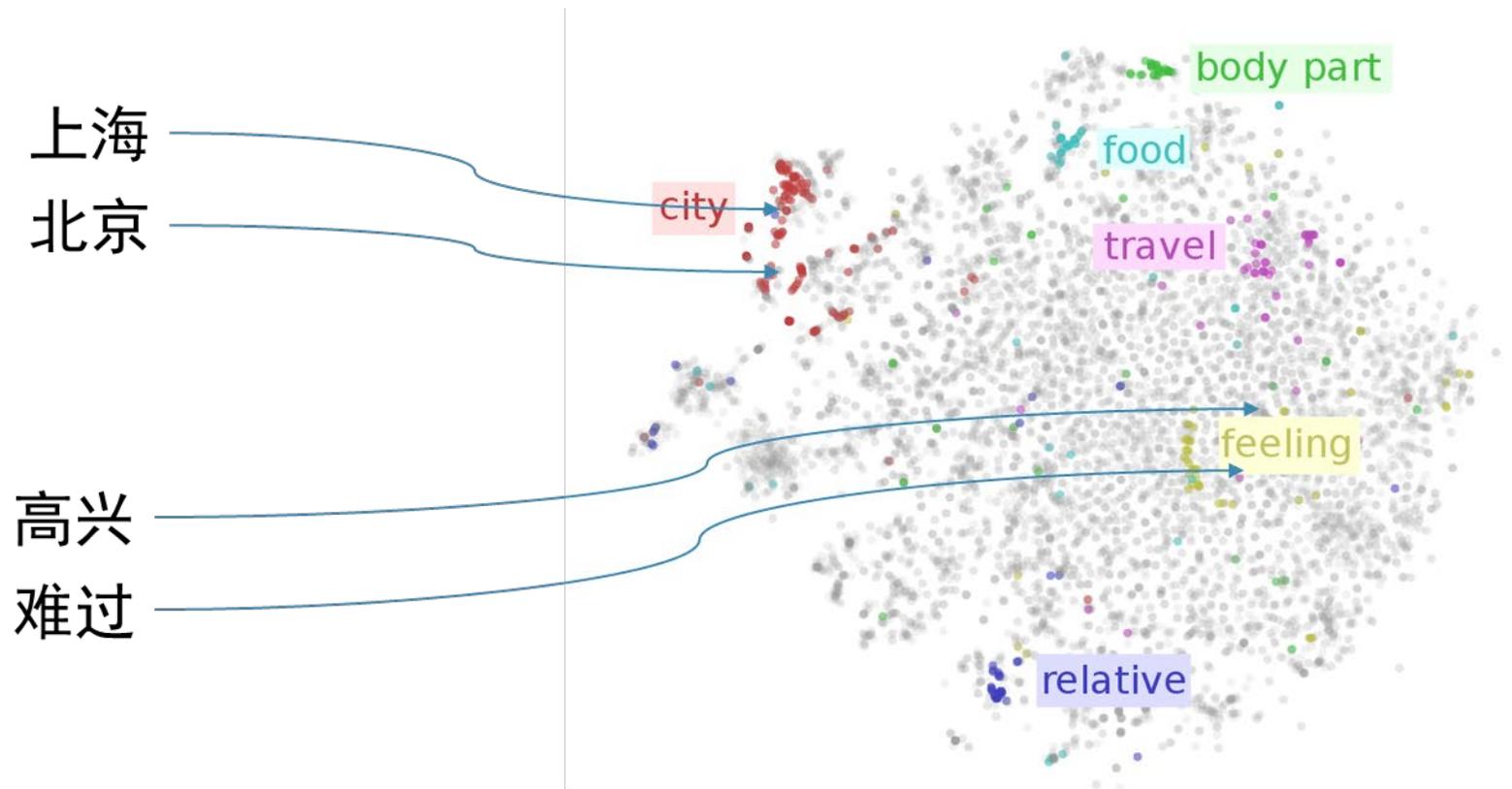
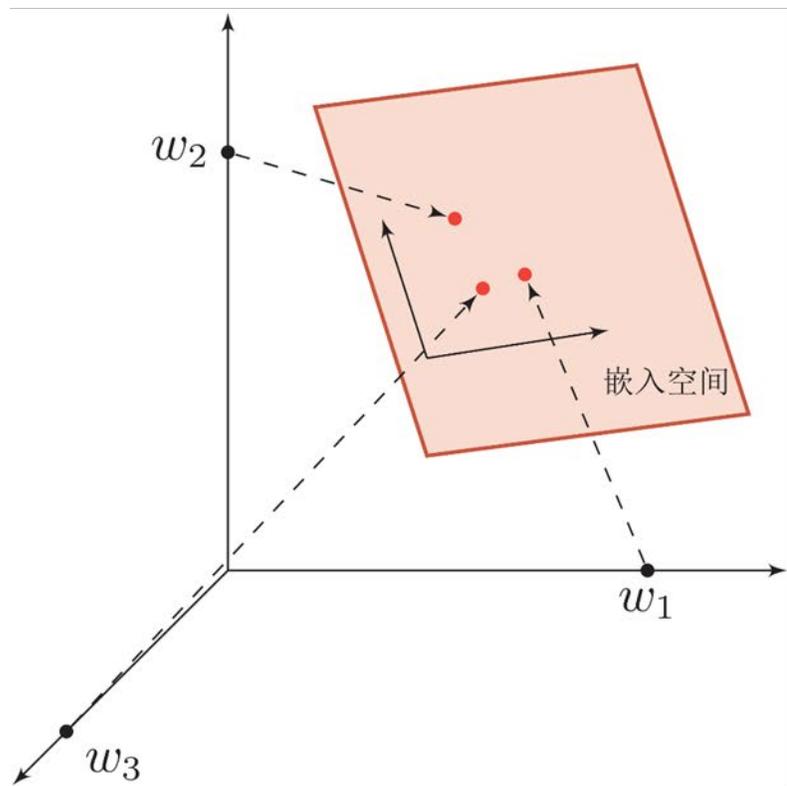
```
[[105 112 108 111 104 99 106 99 96 103 112 119 104 97 93 87]  
[ 91 98 102 106 104 79 98 103 99 105 123 136 110 105 94 85]  
[ 76 85 90 105 128 105 87 96 95 99 115 112 106 103 99 85]  
[ 99 81 81 93 120 131 127 100 95 98 102 99 96 93 101 94]  
[106 91 61 64 69 91 88 85 101 107 109 98 75 84 96 95]  
[114 108 85 55 55 69 64 54 64 87 112 129 98 74 84 91]  
[133 137 147 103 65 81 80 65 52 54 74 84 102 93 85 82]  
[128 137 144 140 109 95 86 70 62 65 63 63 60 73 86 101]  
[125 133 148 137 119 121 117 94 65 79 80 65 54 64 72 98]  
[127 125 131 147 133 127 126 131 111 96 89 75 61 64 72 84]  
[115 114 109 123 150 148 131 118 113 109 100 92 74 65 72 78]  
[ 89 93 90 97 108 147 131 118 113 114 113 109 106 95 77 80]  
[ 63 77 86 81 77 79 102 123 117 115 117 125 125 130 115 87]  
[ 62 65 82 89 78 71 80 101 124 126 119 101 107 114 131 119]  
[ 63 65 75 88 89 71 62 81 120 138 135 105 81 98 110 118]  
[ 87 65 71 87 106 95 69 45 76 130 126 107 92 94 105 112]  
[118 97 82 86 117 123 116 66 41 51 95 93 89 95 102 107]  
[164 146 112 80 82 120 124 104 76 48 45 66 88 101 102 109]  
[157 170 157 120 93 86 114 132 112 97 69 55 70 82 99 94]  
[130 128 134 161 139 100 109 118 121 134 114 87 65 53 69 86]  
[128 112 96 117 150 144 120 115 104 107 102 93 87 81 72 79]  
[123 107 96 86 83 112 153 149 122 109 104 75 80 107 112 99]  
[122 121 102 80 82 86 94 117 145 148 153 102 58 78 92 107]  
[122 164 148 103 71 56 78 83 93 103 119 139 102 61 69 84]]
```

What the computer sees

An image is just a big grid of numbers between [0, 255]:

e.g. 800 x 600 x 3
(3 channels RGB)

词嵌入 (Word Embeddings)



<https://indico.io/blog/visualizing-with-t-sne/>

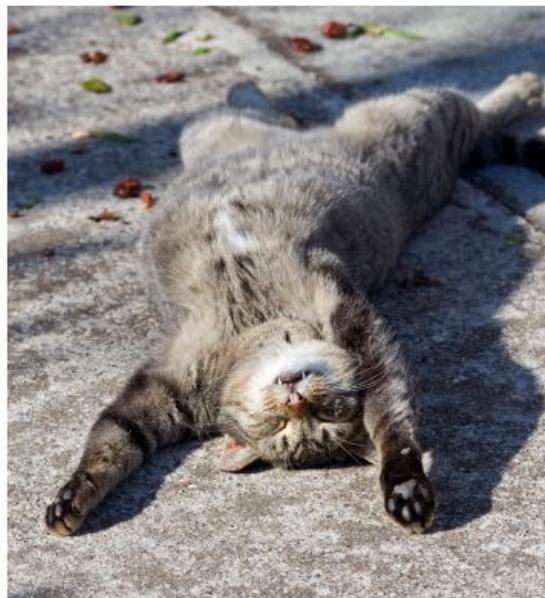
表示学习与深度学习

- 一个好的表示学习策略必须具备一定的深度

- 特征重用：指数级的表示能力
- 抽象表示与不变性：抽象表示需要多步的构造



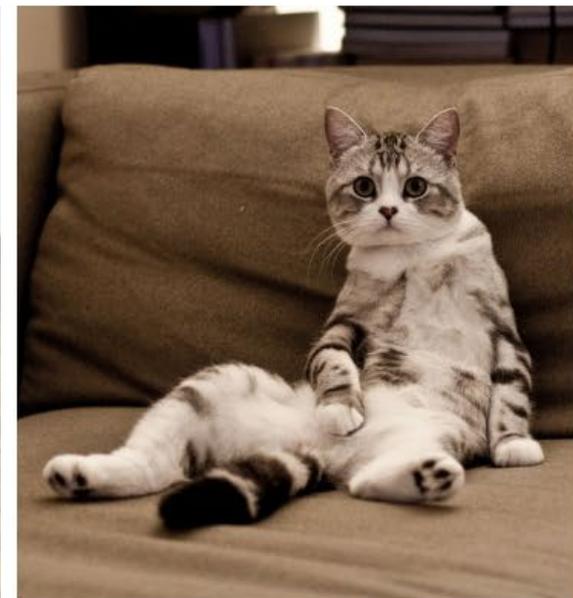
This image by [Umberto Salvagnin](#) is licensed under [CC-BY 2.0](#)



This image by [Umberto Salvagnin](#) is licensed under [CC-BY 2.0](#)



This image by [sare bear](#) is licensed under [CC-BY 2.0](#)

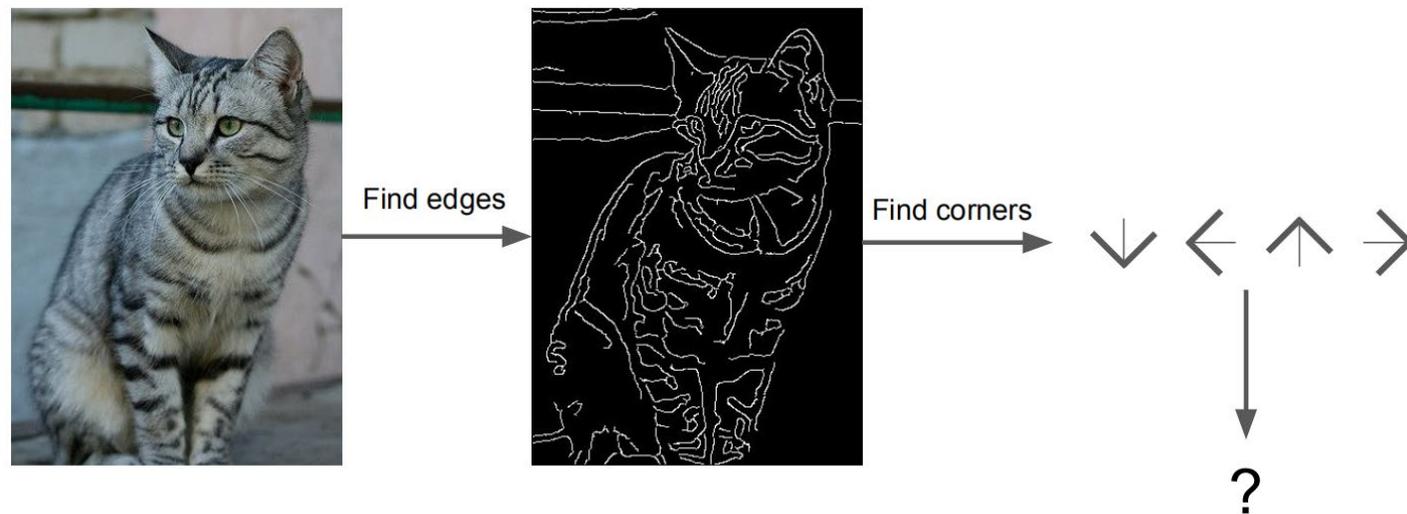


This image by [Tom Thai](#) is licensed under [CC-BY 2.0](#)

传统的特征提取

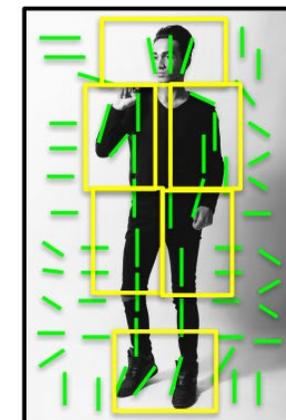
● 特征提取

- 线性投影 (子空间)
 - PCA、LDA
- 非线性嵌入
 - LLE、Isomap、谱方法
- 自编码器



● 特征提取VS表示学习

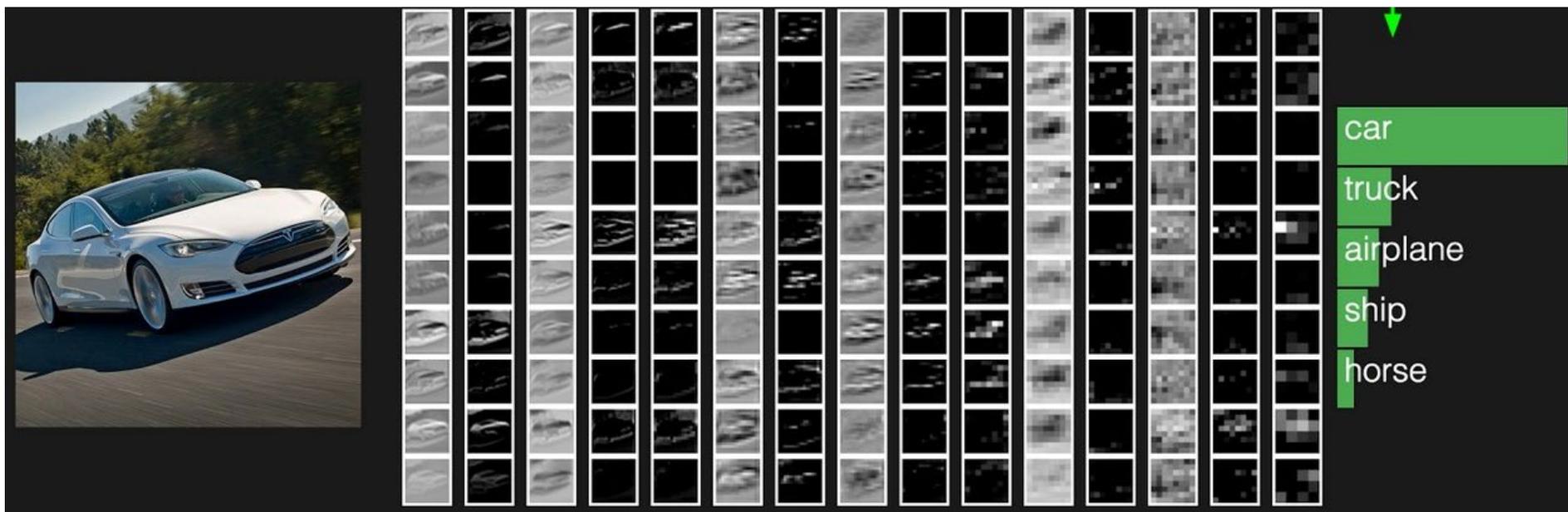
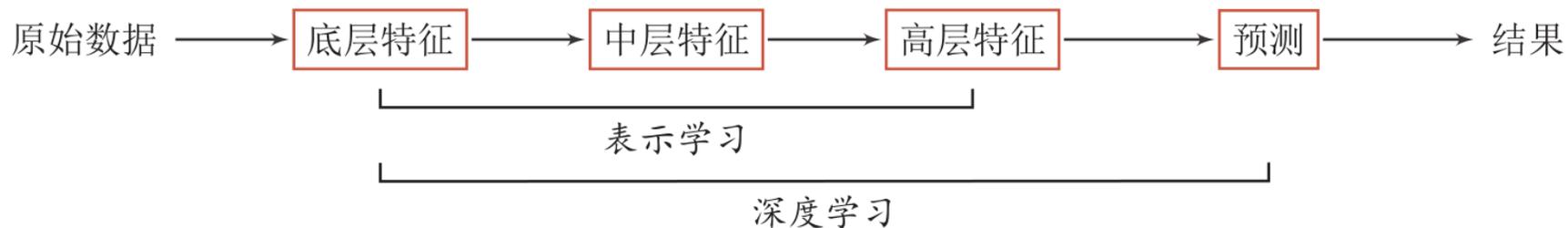
- 特征提取：基于任务或先验对去除无用特征
- 表示学习：通过深度模型学习高层语义特征



● “深度”模型

- 可以让模型来自动学习好的特征表示（从底层特征，到中层特征，再到高层特征），从而最终提升预测或识别的准确性。

关键问题：贡献度分配



$$y = f(x)$$

浅层学习



$$y = f(f(x))$$



⋮



$$y = f(f(f(f\dots(f(x))))))$$

深度学习

神经网络的发展



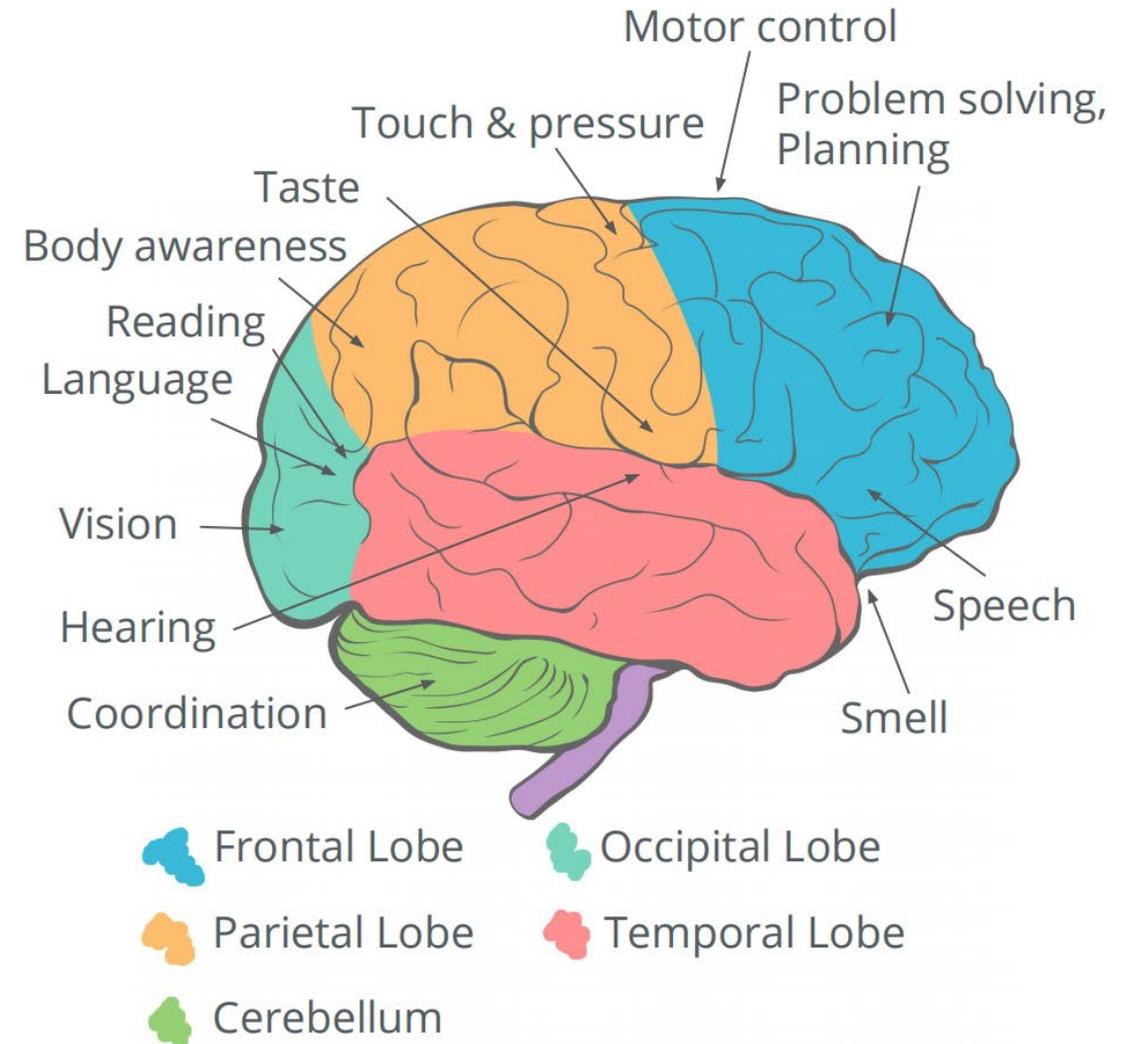
Learning in Nature how does the brain work?



Architecture of the brain

Right side/left side (hemispheres)

- Frontal lobe
 - executive functions, memory and planning
- Parietal lobe
 - sensation and spatial awareness
- Temporal lobe (banana shape)
 - hearing and language
- Occipital lobe at back
 - Vision from front along optic nerves



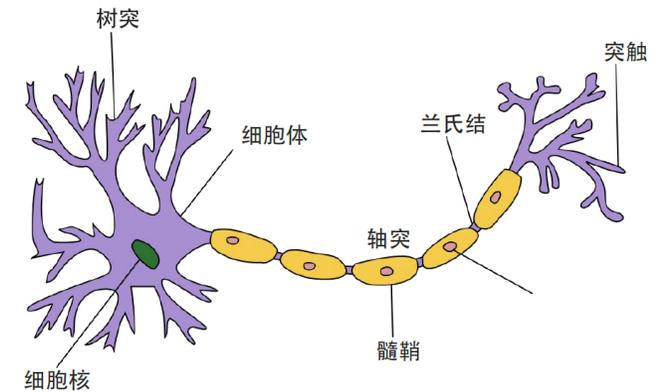
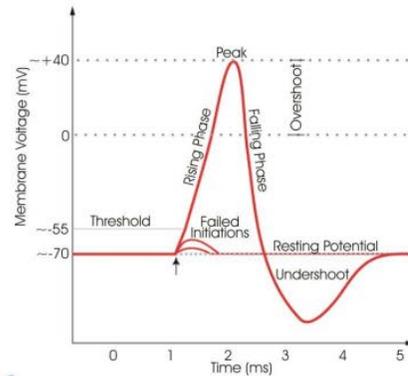
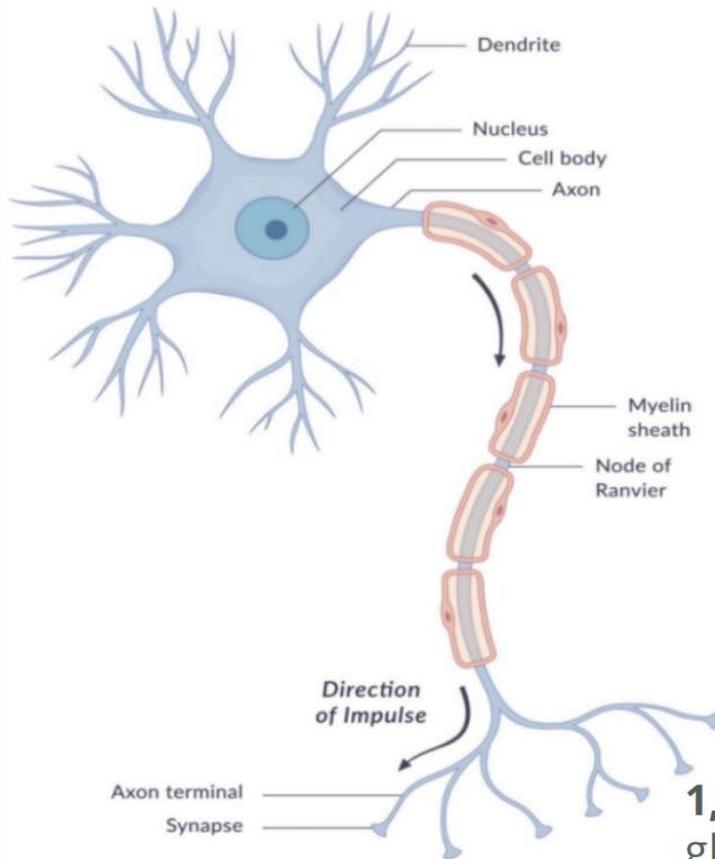


Learning in Nature how does the brain work?

1,000's of inputs (other neurons, sen

- 细胞体 (Soma) 中的神经细胞膜上有各种受体和离子通道，胞膜的受体可与相应的化学物质神经递质结合，引起离子通透性及膜内外电位差发生改变，产生相应的生理活动：兴奋或抑制。
- 细胞突起是由细胞体延伸出来的细长部分，又可分为树突和轴突。

- 树突 (Dendrite) 可以接受刺激并将兴奋传入细胞体。每个神经元可以有一或多个树突。
- 轴突 (Axons) 可以把自身的兴奋状态从胞体传送到另一个神经元或其他组织。每个神经元只有一个轴突。

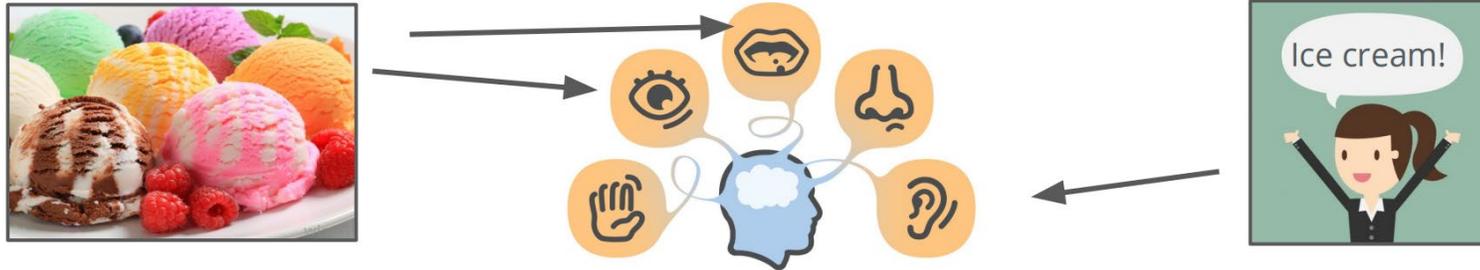


1,000's of output targets (e.g. other neurons, muscle cells, gland cells, blood vessels to release hormones...)

神经网络如何学习?

● 赫布法则 Hebb's Rule

- 如果两个神经元总是相关联地受到刺激，它们之间的突触强度增加。这样的学习方法被称为赫布型学习 (Hebbian learning)

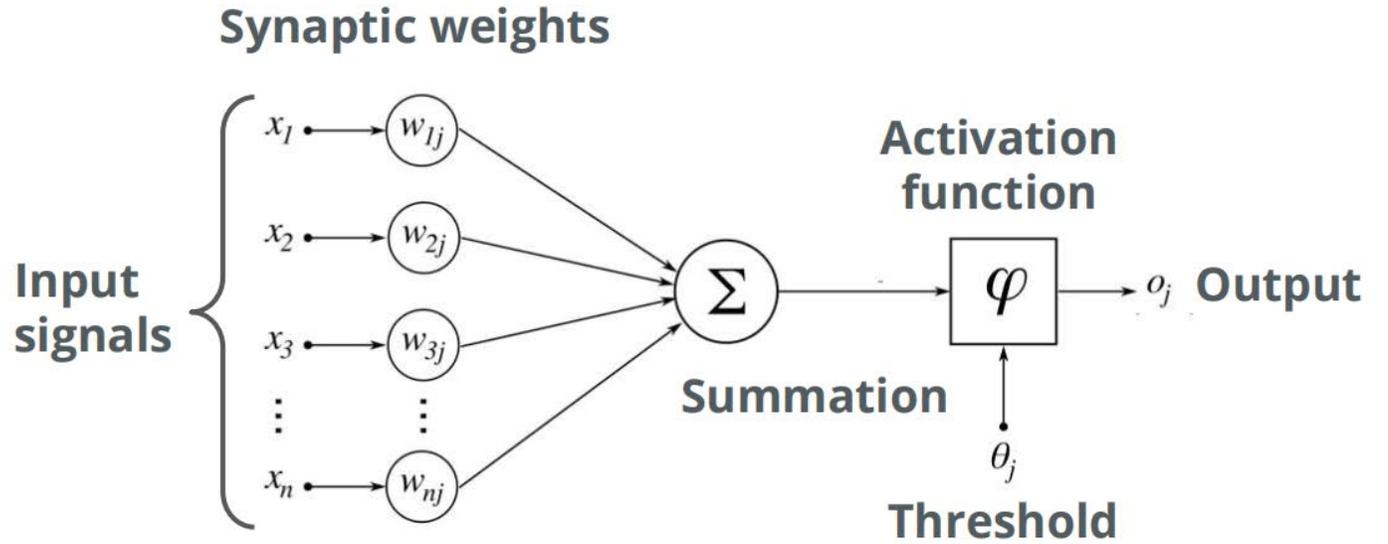
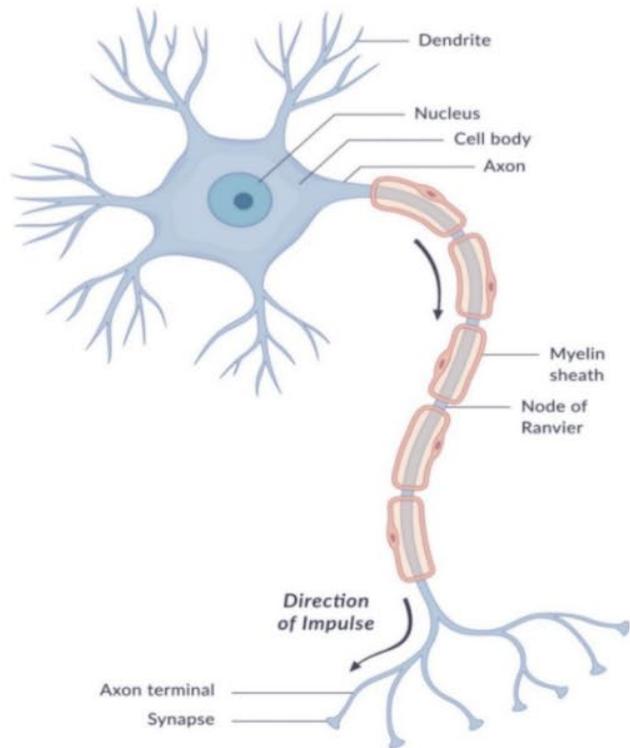


- Hebbian theory
 - If two neurons fire at the same time, the connections between them are strengthened, and thus are more likely to fire again together in the future
 - If two neurons fire in an uncoordinated manner, their connections are weakened and they are more likely to act independently in the future
- Updated hebbian hypothesis based on recent findings
 - If the presynaptic neuron fires within a window of 20ms before the postsynaptic neuron, the synapse will be strengthened
 - However if within a window of 20ms after, the synapse will be weakened

人工神经元

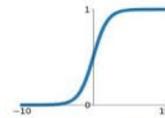
Pitts and McCulloch, 1943

The Artificial Neuron



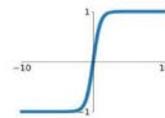
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



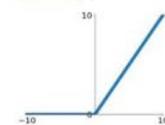
tanh

$$\tanh(x)$$



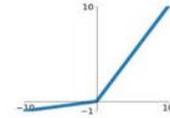
ReLU

$$\max(0, x)$$



Leaky ReLU

$$\max(0.1x, x)$$

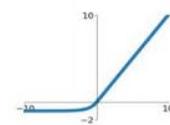


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

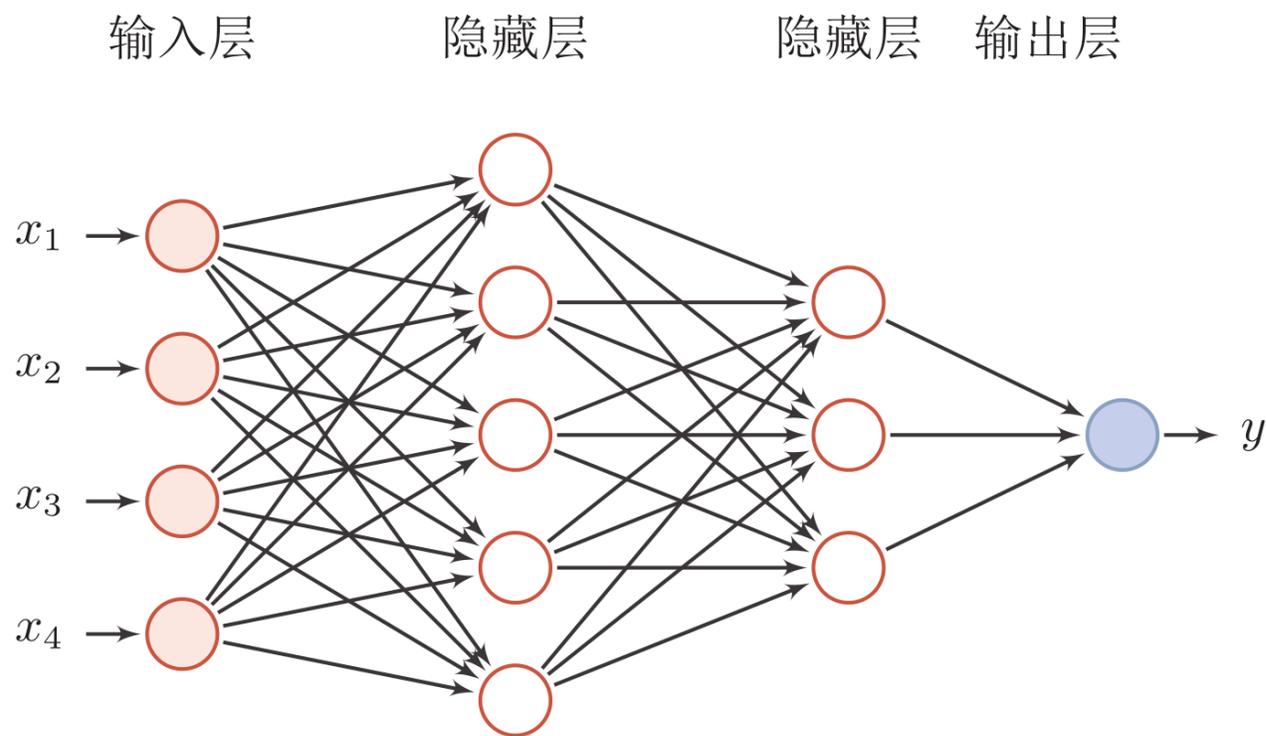
ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



$$y = (f^3(f^2(f^1(x))))$$

$$f^1(x) = \sigma(W^1 f^{1-1}(\cdot))$$



神经网络地这种结构及反向传播算法较好地解决了“贡献度分配”问题!

人工神经网络

- 人工神经网络主要由大量的神经元以及它们之间的有向连接构成。因此考虑三方面：

- 神经元的激活规则

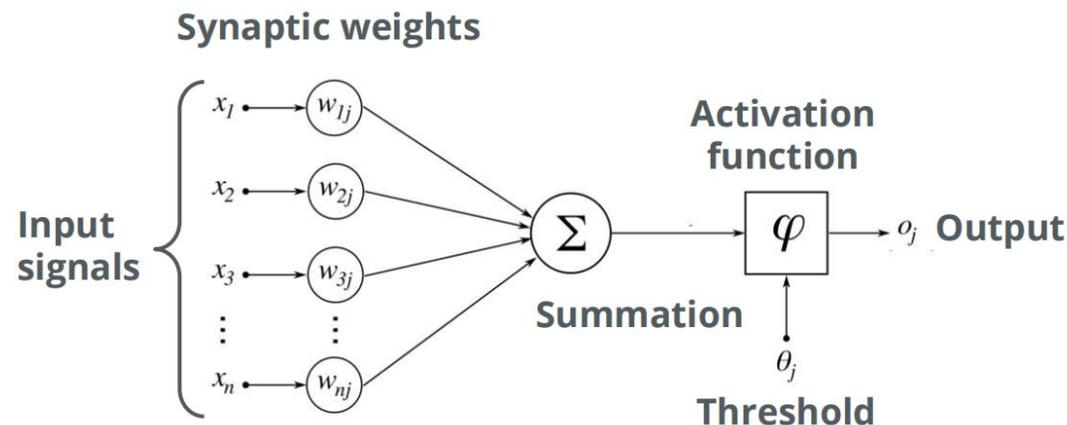
- 主要是指神经元输入到输出之间的映射关系，一般为非线性函数。

- 网络的拓扑结构

- 不同神经元之间的连接关系。

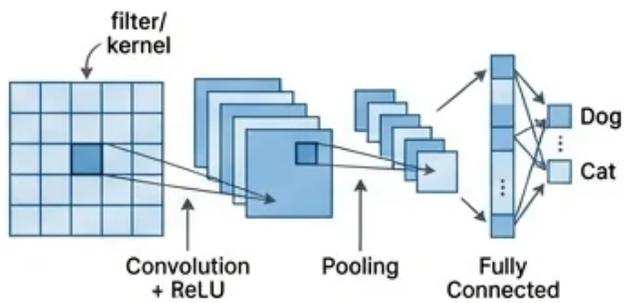
- 学习算法

- 通过训练数据来学习神经网络的参数。



人工神经网络

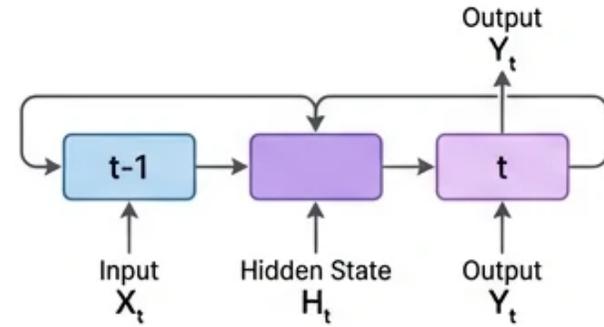
CNN (Convolutional Neural Network)



Best for:
Images,
Computer Vision

Specialized for processing grid-like data. Uses convolutional layers to automatically learn spatial hierarchies of features from input images.

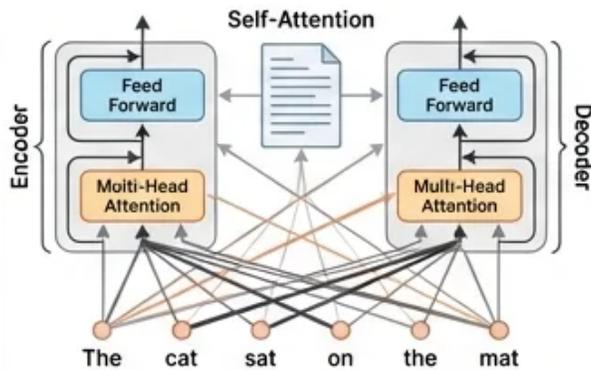
RNN/LSTM (Recurrent Neural Network)



Best for:
Sequential Data,
Time Series

Designed for sequential data. Remembers past information to influence current processing, making it ideal for tasks with time dependencies.

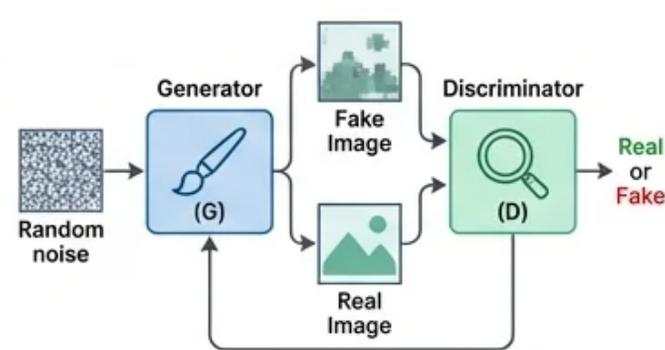
Transformer



Best for: NLP,
Language Models
(GPT, BERT)

Relies entirely on attention mechanisms to draw global dependencies between input and output. Enables parallel processing, revolutionizing natural language understanding.

GAN (Generative Adversarial Network)



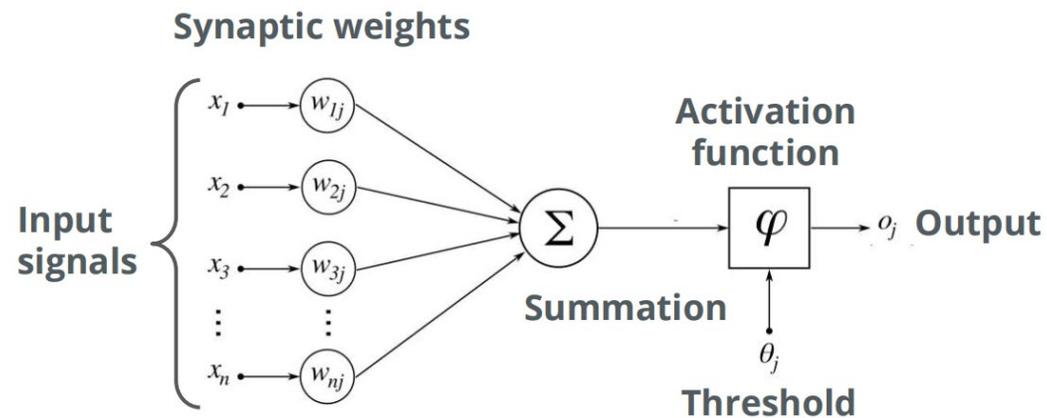
Best for:
Image Generation,
Deepfakes

Consists of two neural networks, a generator and a discriminator, contesting with each other in a game. The generator creates data, and the discriminator evaluates its authenticity.

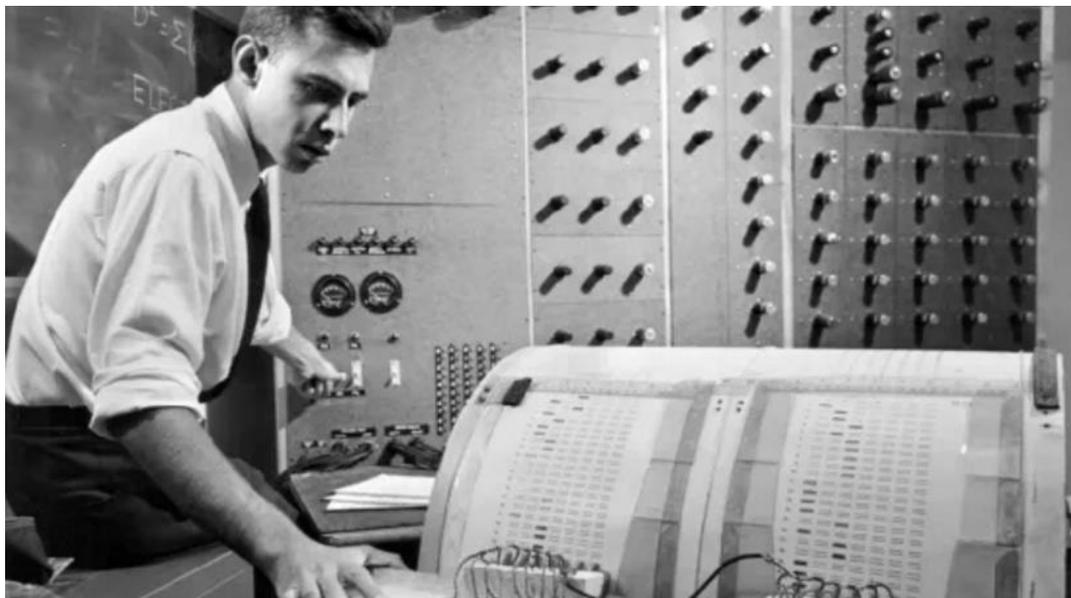
神经网络发展史

● 第一阶段：模型提出

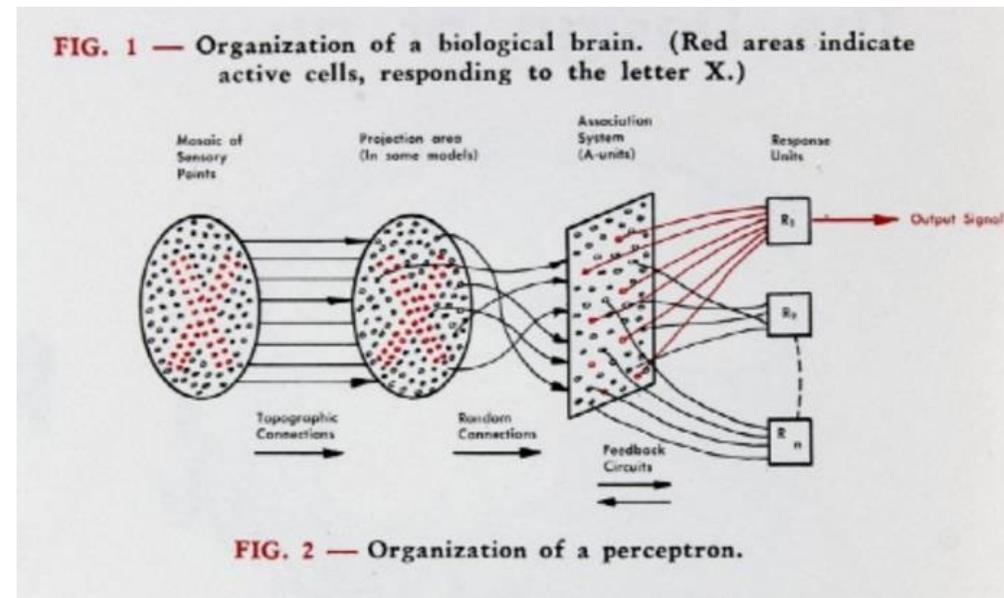
- 在1943年，心理学家Warren McCulloch和数学家Walter Pitts和最早描述了一种理想化的人工神经网络，称为MP模型。
- 阿兰·图灵在1948年的论文中描述了一种“B型图灵机”。(赫布型学习)
- 1951年，McCulloch和Pitts的学生Marvin Minsky建造了第一台神经网络机，称为SNARC。
- Rosenblatt [1958]提出**感知器 (Perceptron)**，并提出了一种接近于人类学习过程（迭代、试错）的学习算法。



神经网络发展史



Rosenblatt & IBM 704

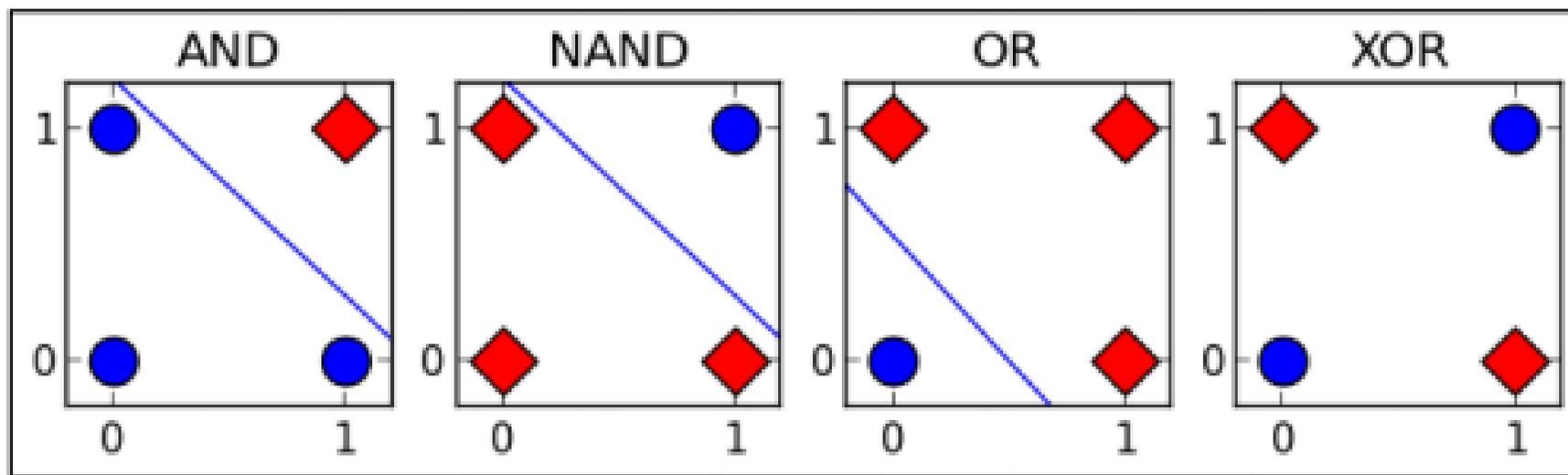


The Design of an Intelligent Automaton
一文中的感知机原理图

神经网络发展史

● 第二阶段：冰河期

- 1969年，Marvin Minsky 《感知机》一书，导致神经网络十多年的“冰河期”。他们发现了神经网络的两个关键问题。
 - 第一是**基本感知机无法处理异或回路**。
 - 第二个重要的问题是电脑没有足够的能力来处理大型神经网络所需要的很长的计算时间。



神经网络发展史

● 第二阶段：冰河期

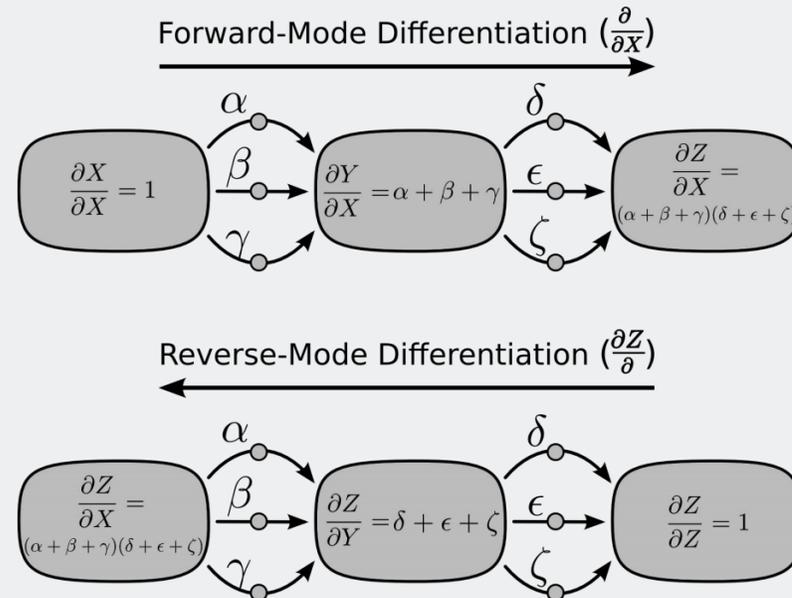
- 但是，这一时期的很多研究成果成为神经网络的后续完善提供了重要理论支撑，例如：
- 1974年，哈佛大学的Paul Werbos发明**反向传播算法**；
- 1980年，Kunihiko Fukushima（福岛邦彦）提出了一种带卷积和子采样操作的多层神经网络：新
知机（Neocognitron）

Backpropagation, 1970

The backpropagation algorithm (also called the reverse mode of automatic differentiation) was independently discovered by different researchers

Seppo Linnainmaa, 1970
Werbos, 1974 (with neural networks)

Reverse-mode differentiation



神经网络发展史

● 第三阶段：反向传播算法引起的复兴

- 1983年, John Hopfield引入能量函数, 并提出Hopfield网络;
- 1984年, Geoffrey Hinton提出玻尔兹曼机;
- 1986年, Hinton等人将反向传播引入到多层感知器;
- Yann Lecun等人在这一时期探索了卷积神经网络的应用。

History of CNNs

Convolutional neural networks were first introduced by Kuniyuki Fukushima in 1980

- **1989** - Yann Lecun et al., trained a CNN with “Backpropagation Applied to Handwritten Zip Code Recognition”
- **1998** - Yann Lecun et al., released LeNet5 “Gradient-based learning applied to document recognition”

Example: LeNet5

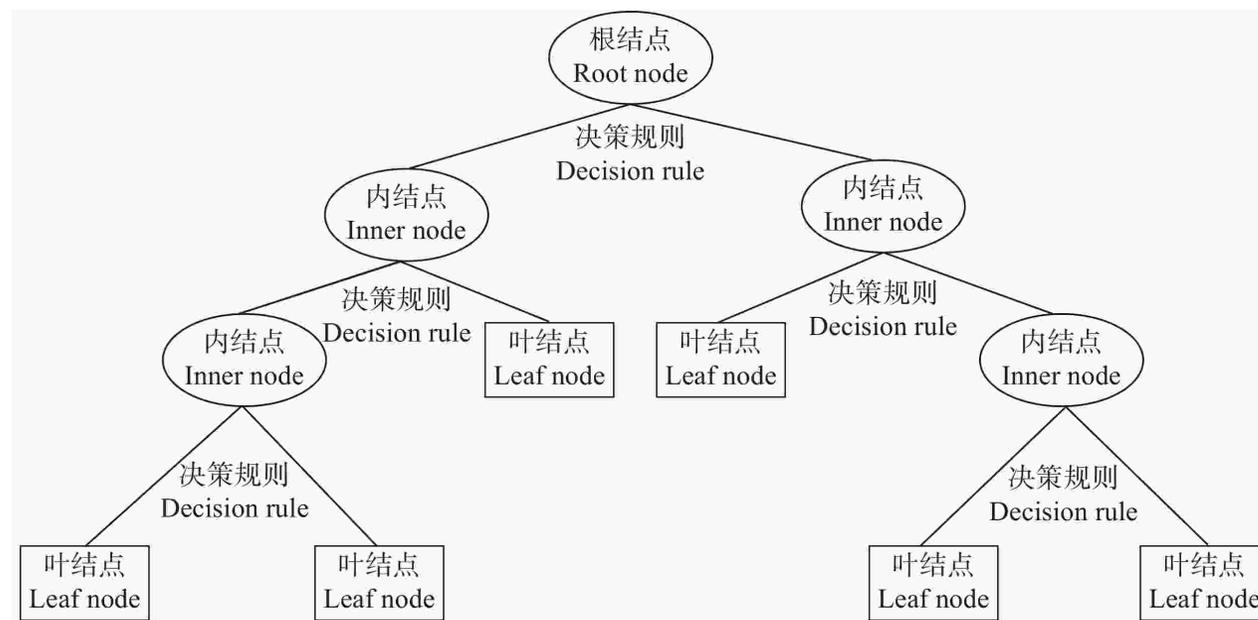
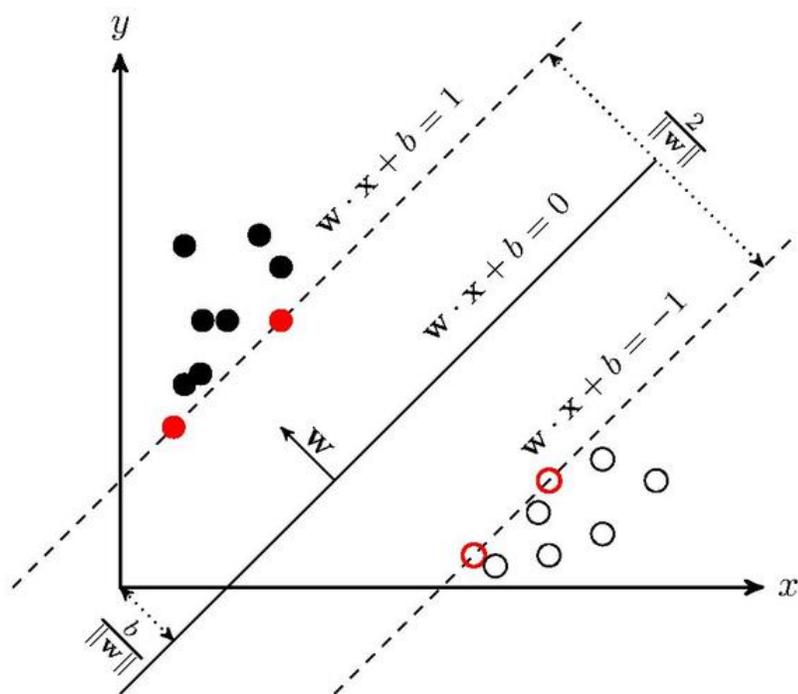
```
class LeNet(nn.Module):
    def __init__(self):
        super(LeNet, self).__init__()
        self.conv1 = nn.Conv2d(1, 6, 5, padding=2)
        self.conv2 = nn.Conv2d(6, 16, 5)
        self.fc1 = nn.Linear(16*5*5, 120)
        self.fc2 = nn.Linear(120, 84)
        self.fc3 = nn.Linear(84, 10)

    def forward(self, x):
        x = F.max_pool2d(F.relu(self.conv1(x)), (2, 2))
        x = F.max_pool2d(F.relu(self.conv2(x)), (2, 2))
        x = flatten(x)
        x = F.relu(self.fc1(x))
        x = F.relu(self.fc2(x))
        x = self.fc3(x)
        return x
```

神经网络发展史

● 第四阶段：流行度降低

- 在20世纪90年代中期，统计学习理论和以支持向量机为代表的机器学习模型开始兴起。
- 相比之下，神经网络的理论基础不清晰、优化困难、可解释性差等缺点更加凸显，神经网络的研究又一次陷入低潮。



● 第五阶段：深度学习的崛起

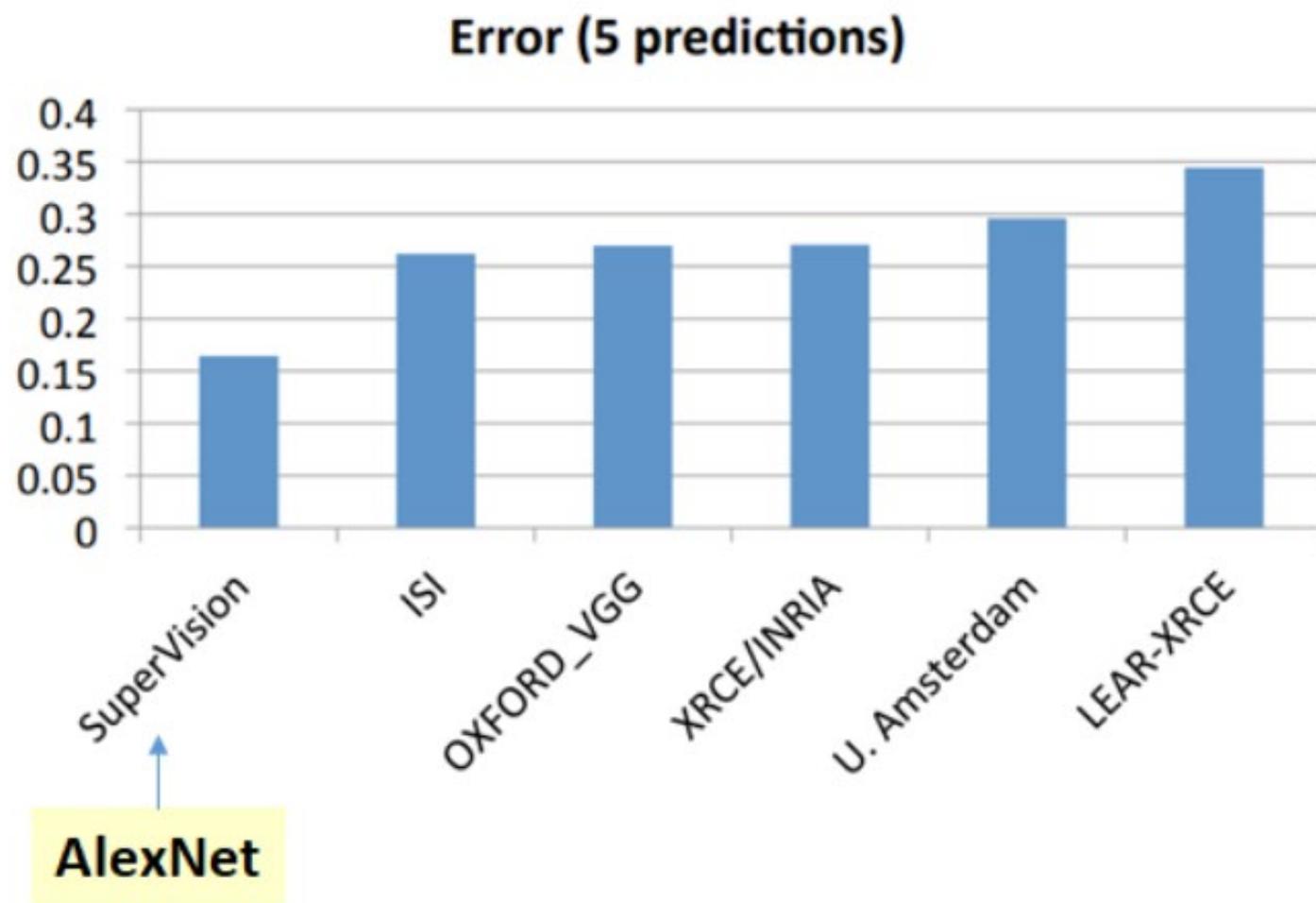
- 2006年，Hinton等人发现多层前馈神经网络可以先通过逐层预训练，再用反向传播算法进行精调的方式进行有效学习。
 - 深度神经网络在语音识别和图像分类等任务上的巨大成功。
- 2012年，AlexNet：第一个现代深度卷积网络模型，是深度学习技术在图像分类上取得真正突破的开端。
 - AlexNet不用预训练和逐层训练，首次使用了很多现代深度网络的技术
- 随着大规模并行计算以及GPU设备的普及，计算机的计算能力得以大幅提高。此外，可供机器学习的数据规模也越来越大。在计算能力和数据规模的支持下，计算机已经可以训练大规模的人工神经网络。

神经网络发展史



ImageNet是一个包含超过1500万张高分辨率图像的分类数据集，共划分为约22000个类别。从2010年开始，每年举办ImageNet大型视觉识别挑战赛(ILSVRC)，比赛中选手需报告top-1错误率和top-5错误率。

神经网络发展史



深度学习与AI革命

- 语音识别：词错误率大幅降低
- 计算机视觉：目标识别、图像分类等
- 自然语言处理：分布式表示、机器翻译、问题回答等
- 信息检索、社会化网络



AI“数智员工”：迈出人机协同治理重要一步

2025年02月21日08:58 | 来源：科技日报

Tf 小字号

【科技观察家】

◎罗云鹏

近期，广东省深圳市福田区上线政务大模型2.0，并推出了70名基于DeepSeek开发的AI“数智员工”，这一创新举措引发了广泛关注。作为政务服务智能化转型的前沿实践，福田区的探索不仅提升了政务工作效率，更在人机协同的治理新模式上迈出了重要一步。

首先，AI“数智员工”的出现，是对政务服务高度标准化、规范化特点的精准契合。AI技术的核心优势在于处理标准化、重复性任务的高效性和精准性，这与政务服务的内在需求高度一致。在福田区的实践中，AI“数智员工”被应用于公文处理、民生服务、应急管理等多个场景，显著提升了政务工作的效率和质量。例如，“执法文书生成助手”能够将笔录秒级转化为初稿，公文修正的准确率超过95%，审核时间缩短了90%；在民生服务领域，诉求分拨的准确率从70%跃升至95%，传统需要数日的流程被压缩至分钟级。这些数据背后，是AI技术对政务流程的深度优化，更是对政务工作“以人民为中心”理念的有力践行。

其次，AI“数智员工”的价值不仅在于提升效率，更在于推动政务服务从“事务处理”向“价值创造”的跃迁。福田区明确，将AI“数智员工”定位为辅助工具，强调人机协同的治理逻辑。AI

- **Toronto 大学**

- Hinton 75 年 Edinburgh 大学博士

- **NYU**

- Lecun (Now Facebook) 87 年 Hinton 博士后

- **Montreal大学**

- Bengio 91 年 M. Jordan 博士后

- **IDSIA (Dalle Molle人工智能研究所)**

- Jürgen Schmidhuber



2018图灵奖获得者



2024诺贝尔物理学奖获得者

深度学习框架

- 简易和快速的原型设计
- 自动梯度计算
- 无缝CPU和GPU切换



Q&A