



The current states and progress of Artificial Intelligence (AI) and Machine Learning (ML)

SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY, CHINESE ACADEMY OF SCIENCES.

PRESENTER: QINGTIAN WU

DIRECTOR: YIMIN ZHOU

2018/9

CONTENTS

➤ 1TH PART : the current states and progress of AI

- 1) A **brief history** of AI (人工智能**简史**) and three school of thoughts (**三大学派**): symbolism, connectionism and behaviorism. (符号主义、连接主义和行为主义)
- 2) What is Machine Learning?
- 3) How to regard(看待) **Data**?
- 4) Debates in ML. (机器学习百花齐放、百家争鸣的**学派之争**)
- 5) Research on the application of AI in several famous IT companies

➤ 2TH PART : The future work about AI

- 1) Creating powerful data analysis capabilities. (打造有影响力的大**数据分析能力**)
- 2) **AI platform and Cloudy AI**: Intelligent operation and maintenance. Transforming operational data into intelligent decision making.
(**人工智能平台及云化**)：快速有效地解决具体业务问题；实现智慧运维，把运营数据转化为智能决策
- 3) **AI with IoT, Smart City**: an insight into big data. (AI与**物联网、智慧城市**：对大数据的洞察分析)
- 4) CONCLUSION

The current states and progress
of Artificial Intelligence.

AI : Being named at the Dartmouth Conference in 1956

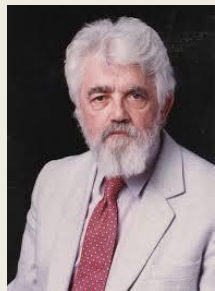


- ❑ 人工智能的研究可以追溯到亚里士多德的三段论 (syllogisms)、莱布尼兹、布尔、弗雷格、罗素、哥德尔等人在数理逻辑方面的工作
- ❑ Dartmouth 会议的 organizer 是 Marvin Minsky, John McCarthy, Claude Shannon, 与会者包括 Arthur Samuel, Newell 和 Simon 等大师
- ❑ 会议提出，**学习或者智能的任何特性都能够被精确地加以描述，使得机器可以对其进行模拟。**
- ❑ 人工智能早期的‘豪言壮语’：
 - 1958年, Simon, Newell: 十年之内，数字计算机将成为国际象棋世界冠军。十年之内，数字计算机将发现并证明一个重要的数学的定理
 - 1965年, Simon：二十年内，机器将能完成人能做到的工作。
 - 1967年, Marvin Minsky：一代之内.....创造人工智能的问题将获得实质上的解决。
 - 1970年, Marvin Minsky: 在三年到八年的时间里我们将得到一台具有人类平均智能的机器

The controversy(争辩) of artificial intelligence: Symbolism (符号) vs Connectionism (连接)

符号主义(symbolicism), 又称为逻辑主义(logicism)、心理学派(psychologism)或计算机学派(computerism), 其原理主要为物理符号系统(即符号操作系统)假设和有限合理性原理。

连接主义(connectionism), 又称为仿生学派(bionicsism)或生理学派(physiologism), 其主要原理为神经网络及神经网络间的连接机制与学习算法。



John McCarthy
(1927-2011)



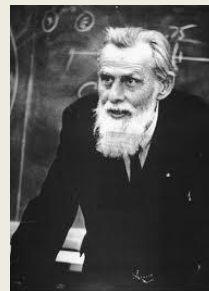
Allen Newell
(1927-1992)



Herbert Simon
(1916-2001)



Edward Feigenbaum
(1936-)



Warren S. McCulloch
(1898-1969)



Walter H. Pitts
(1923-1969)



Marvin Minsky
(1927-2016)

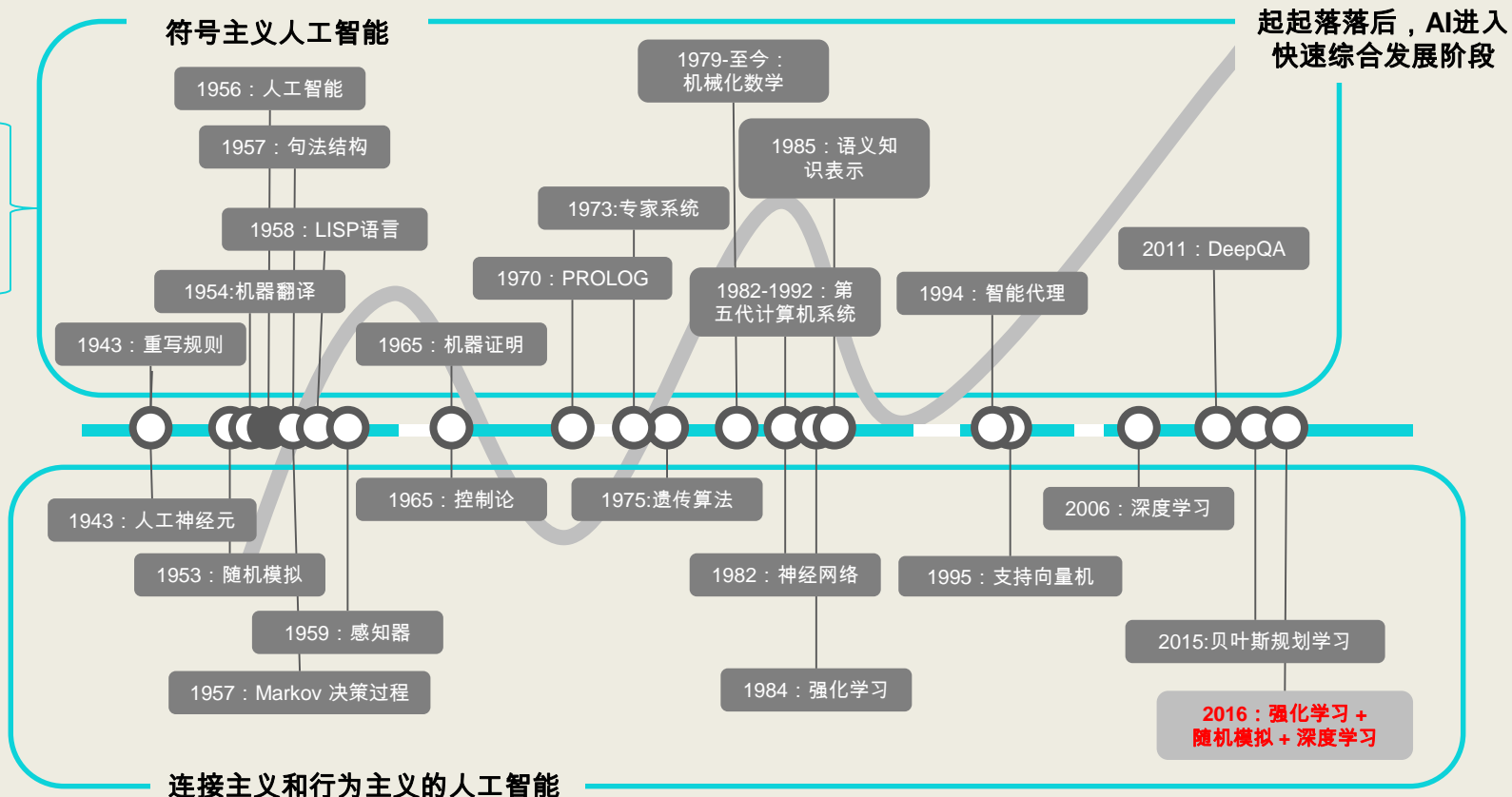
Brief History of AI (人工智能简史)

二十世纪数学哲学

□ **形式主义**：所有的数学分支都可以公理化

□ **逻辑主义**：一切数学都建立在数理逻辑的基础之上

□ **构造主义**：存在就是被构造



Artificial intelligence scholars and major achievements in the Turing Award

(图灵奖中的人工智能学者及主要成就)

Name	year	Institute	Main Contributions
Marvin Minsky	1969	MIT	神经网络和连接主义的奠基人
John McCarthy	1971	Stanford University	人工智能、LISP的开创者、机器证明
Allen Newell	1975	CMU	通用问题求解
Herbert Simon	1975	CMU	通用问题求解
Edward A. Feigenbaum	1994	Stanford University	专家系统之父
Dabbala R. Reddy	1994	CMU	语音识别
Leslie G. Valiant	2010	Harvard University	机器学习理论 (PAC learning)
Judea Pearl	2011	UCLA	贝叶斯网络、因果推断

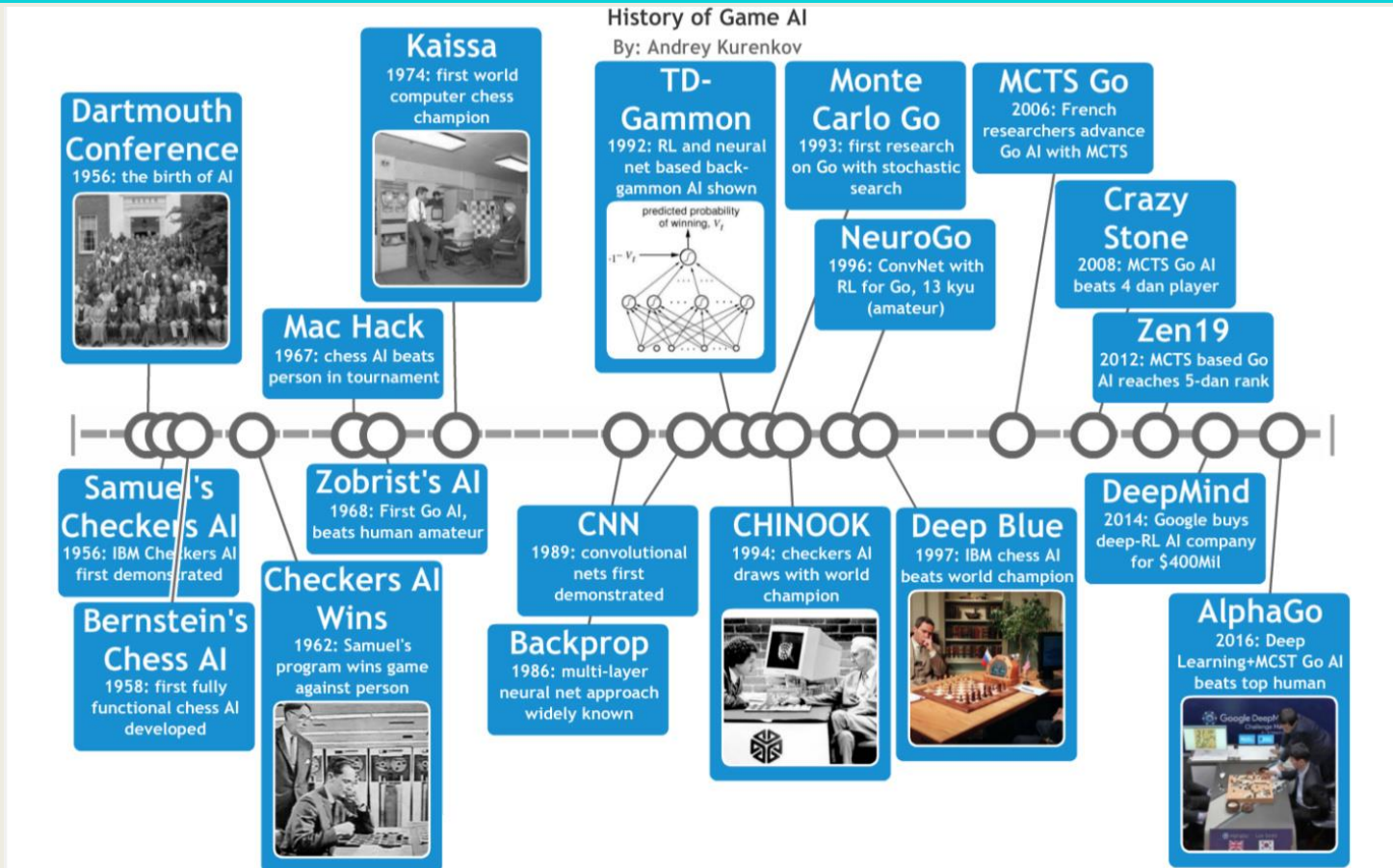
Analysis of the strengths and weaknesses of the three AI schools

(三大人工智能学派的优势和劣势分析)

人工智能三大学派	知识表达	黑箱	特征学习	可解释性	是否需要大样本	计算复杂性	组合爆炸	环境互动	过拟合问题
符号主义 (Symbolism)	强	否	无	强	否	高	多	否	无
连接主义 (Connectionism)	弱	是	有	弱	是	高	少	否	有
行为主义 (Behaviorism)	强	否	无	强	否	一般	一般	是	无

Brief history of AI in Chess Games: three principles toward integration

(人工智能的棋类游戏简史：三大主义走向融合)

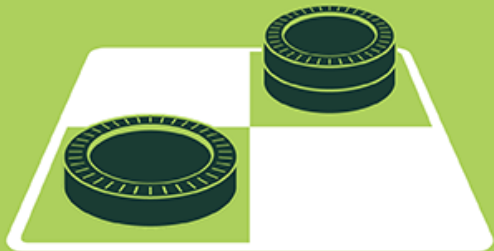


What's Machine Learning ?

The relationship in AI, ML, DL

ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



4 essentials: Algorithm, Computing, Data, Scene.
四要素：算法、算力、数据、场景

MACHINE LEARNING

Machine learning begins to flourish.



DEEP LEARNING

Deep learning breakthroughs drive AI boom.



What's Lightweight ML? (什么是轻量机器学习?)

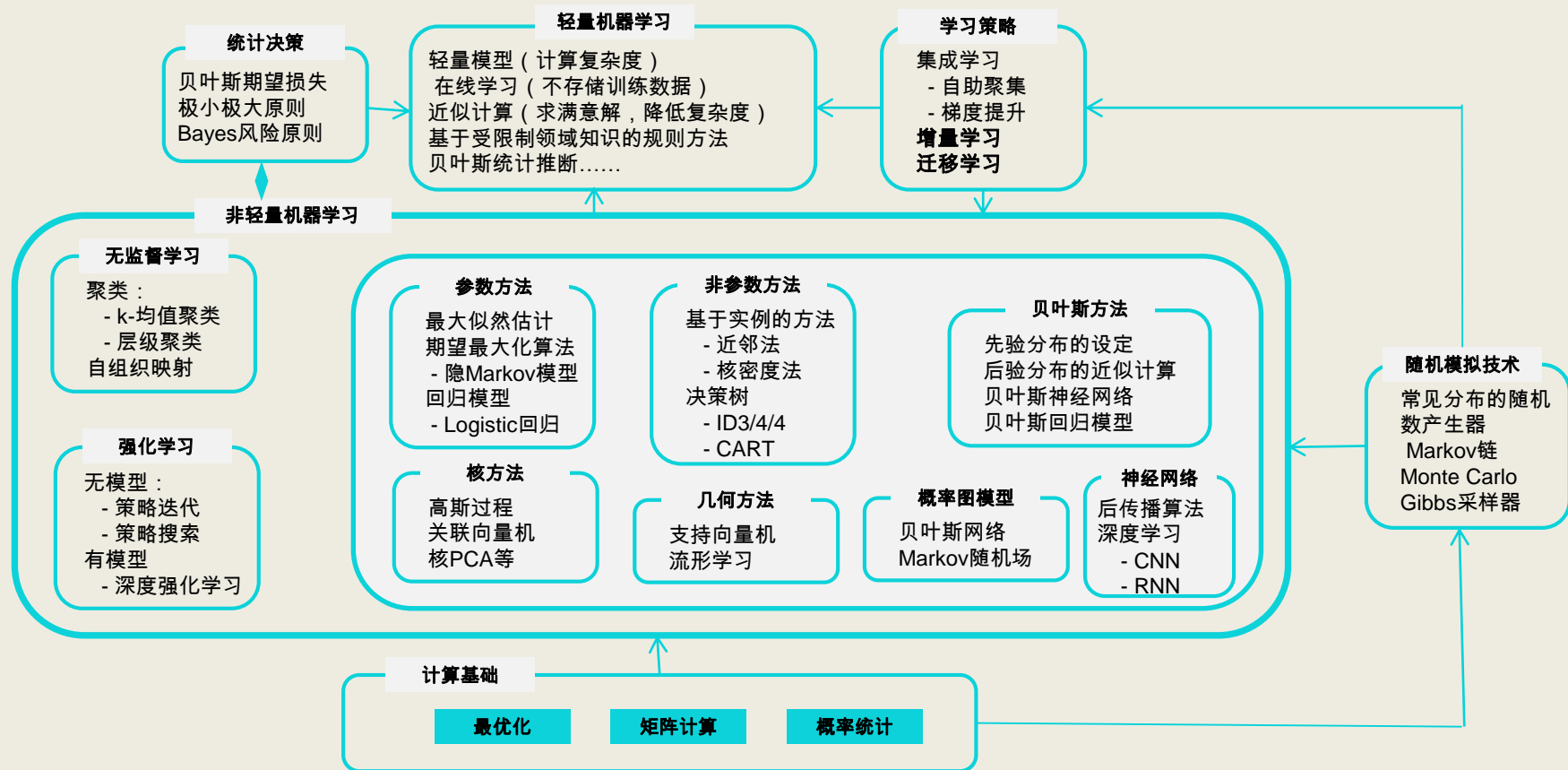
Many application scenarios have limited computing resources and it is impossible to deploy(部署) large-scale machine learning (mainly training).

Lightweight ML is a special ML technology. It has the following characteristics:

1. **Low computational complexity** : Under the premise of normal problem solving, the training of lightweight model is better than that of conventional model with less memory consumption and shorter training cycle.
2. **Have a certain online learning ability** : do not need to store a large number of training samples, to learn one by one, the model can be updated at all times.
3. **High prior knowledge** : in the case of small amount of data, model updating and inference can be still carried out. For example, Bias's statistical inference, pattern theory, and domain knowledge based on rule methods.
4. To use the **approximate calculation** by approximating the optimal solution to meet the demand of satisfactory solution the computation complexity can be reduced greatly.

Lightweight ML is a new concept. Its technology is still in a fragmented state.
轻量机器学习是一个新概念，其技术还处于零散状态，需要边用边储备。

Key technologies in ML/PR (机器学习/模式识别关键技术)



The mathematical foundation of AI (人工智能的数学基础)

基础部分

离散数学	数学分析
集合论	微分
组合数学	积分
图论	级数
数理逻辑	计算数学
代数学	数值分析
线性代数	误差分析
泛代数	随机模拟
抽象代数	运筹学
几何学	最优化
解析几何	排队论
仿射几何	博弈论
向量与张量分析	控制论
微分几何	可靠性理论
计算几何	

函数

函数论
实变函数
复变函数
逼近论
调和分析
特殊函数
泛函分析
线性算子
空间理论
非线性泛函
测度与积分
广义函数论
变分法
积分变换

方程

常微分方程
定性理论
稳定性理论
解析理论
偏微分方程
非线性椭圆、双曲和抛物方程
非线性波
非线性发展方程
积分方程
Fredholm方程
动力系统
微分动力系统
拓扑动力系统

概率统计

概率论	数理统计
概率分布	抽样理论
极限定理	假设检验
随机过程	参数估计
鞅论	方差分析
随机场	多元统计
随机分析	时间序列
随机微分方程	贝叶斯推断
	决策统计
	非参数统计

交叉分支

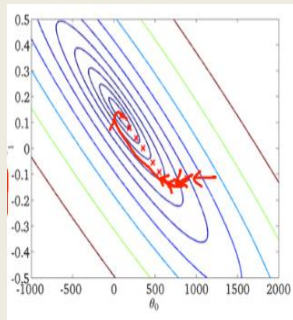
计算机科学	信息科学
可解性与可计算性	分形论
机器证明	生物数学
计算复杂性	经济数学
并行计算	信息论

AI has used most branches of **mathematics**. For example, the mathematical foundation of **reinforcement learning**, the **Markov** decision process comes from the optimal control theory (1950s).

Why do you need a variational method?

The general linear regression function is : $h_{\theta} = \sum_{j=0}^n \theta_j x_j$

The corresponding loss function is : $J(\theta) = 1/(2m) \sum_{i=1}^m (h_{\theta}(x^i) - y^i)^2$

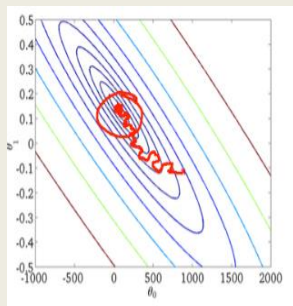


Gradient Descent

$$\frac{\partial J(\theta)}{\partial \theta_j} = \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^i) - y^i) x_j^i$$

Minimize the
loss function

$$\theta_j = \theta_j - \lambda \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^i) - y^i) x_j^i$$



Stochastic gradient descent

$$J(\theta) = 1/(2m) \sum_{i=1}^m (h_{\theta}(x^i) - y^i)^2 = \frac{1}{m} \sum_{i=1}^m \text{cost}(\theta, (x^i, y^i))$$

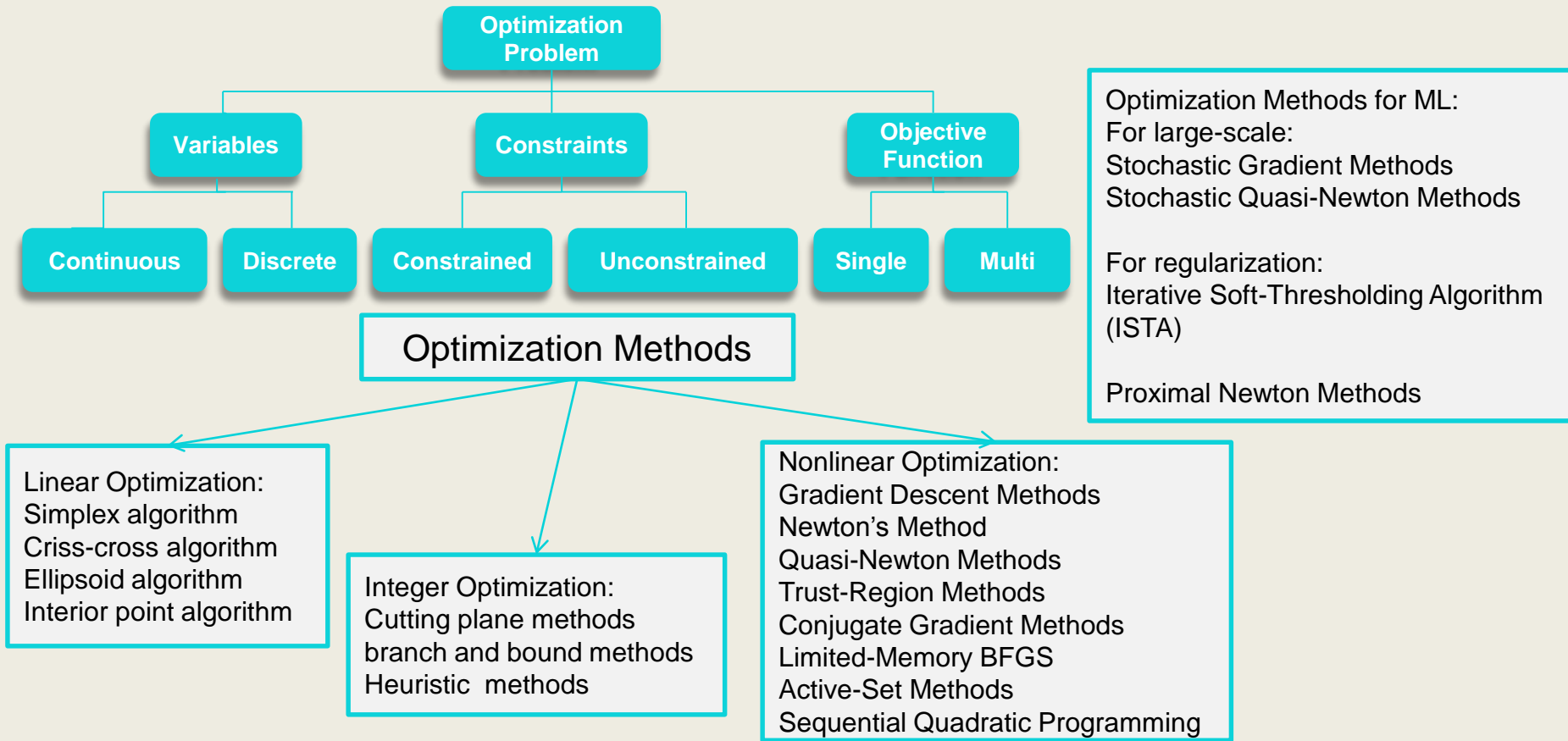
$$\text{cost}(\theta, (x^i, y^i)) = \frac{1}{2} (y^i - h_{\theta}(x^i))^2$$

Using the loss function
of each sample to obtain
a partial derivative of θ

$$\theta_j = \theta_j - \lambda (h_{\theta}(x^i) - y^i) x_j^i$$

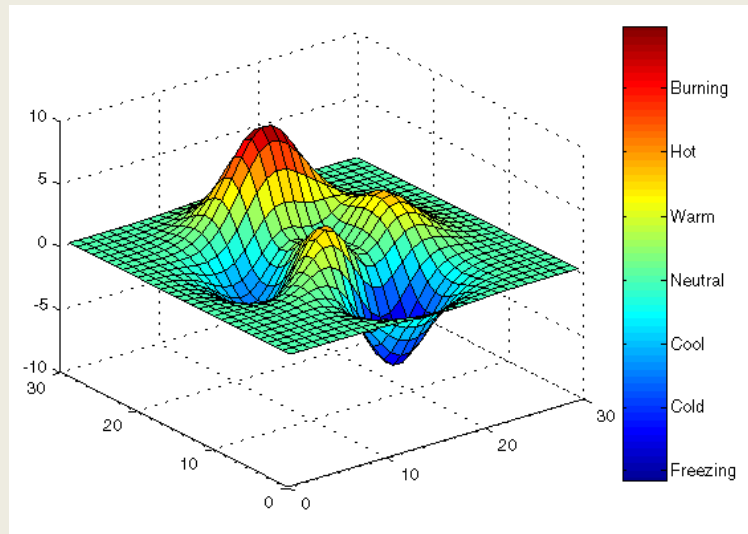
Improve training speed
Probably for non-convex
functions, eventually converge
to a better local extremum point,
even a global extremum point

The core technology of optimization (最优化的核心技术)

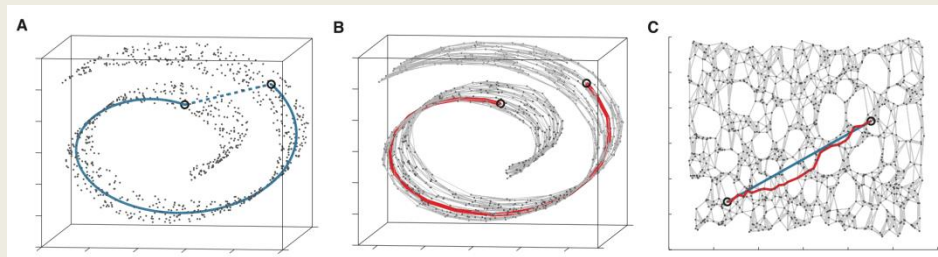


How to regard data ?

Manifold(流形) learning: reconsidering the data space



The distribution of data in space and their manifolds are generally unknown. Manifold can be very complicated. How does its geometric property be "excavated" through data?



Locally manifolds can be considered as Euclidean spaces.
流形局部可看作欧式空间

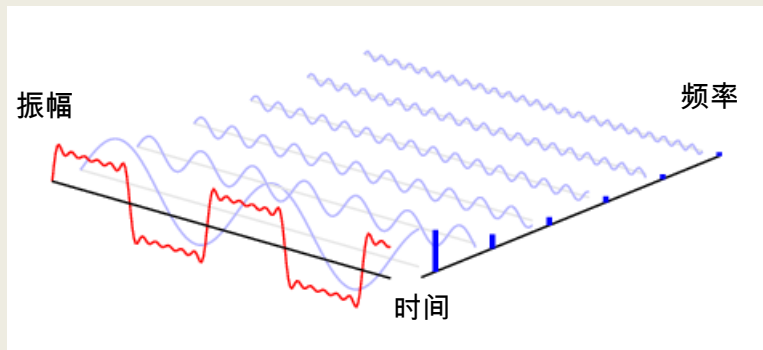
Data on Swiss rolls
瑞士卷上的数据

Artificial intelligence has an inconformity to mathematical is nonsense.

--John McCarthy

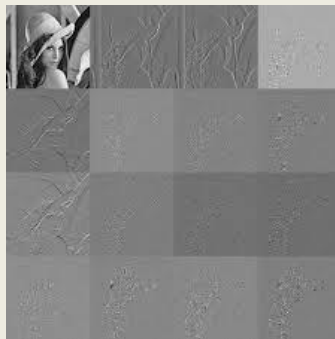
人们常问什么是数据的本质？我们可以从多个角度看，例如，统计分布的角度、数据压缩与重构的角度等。事实上，对数据的本质可以有多种理解，但每个理解都要落实到具体的数学方法上，根据问题和效果来决定是否使用它们。拿数据压缩与重构为例，常见方法有矩阵（或张量）分解、Fourier变换、小波变换等

矩阵与张量分解：数据压缩、特征抽取、推荐系统等等



Fourier Transform: transform signals in time and frequency domain.

傅里叶变换：将信号在时域和频域之间变换

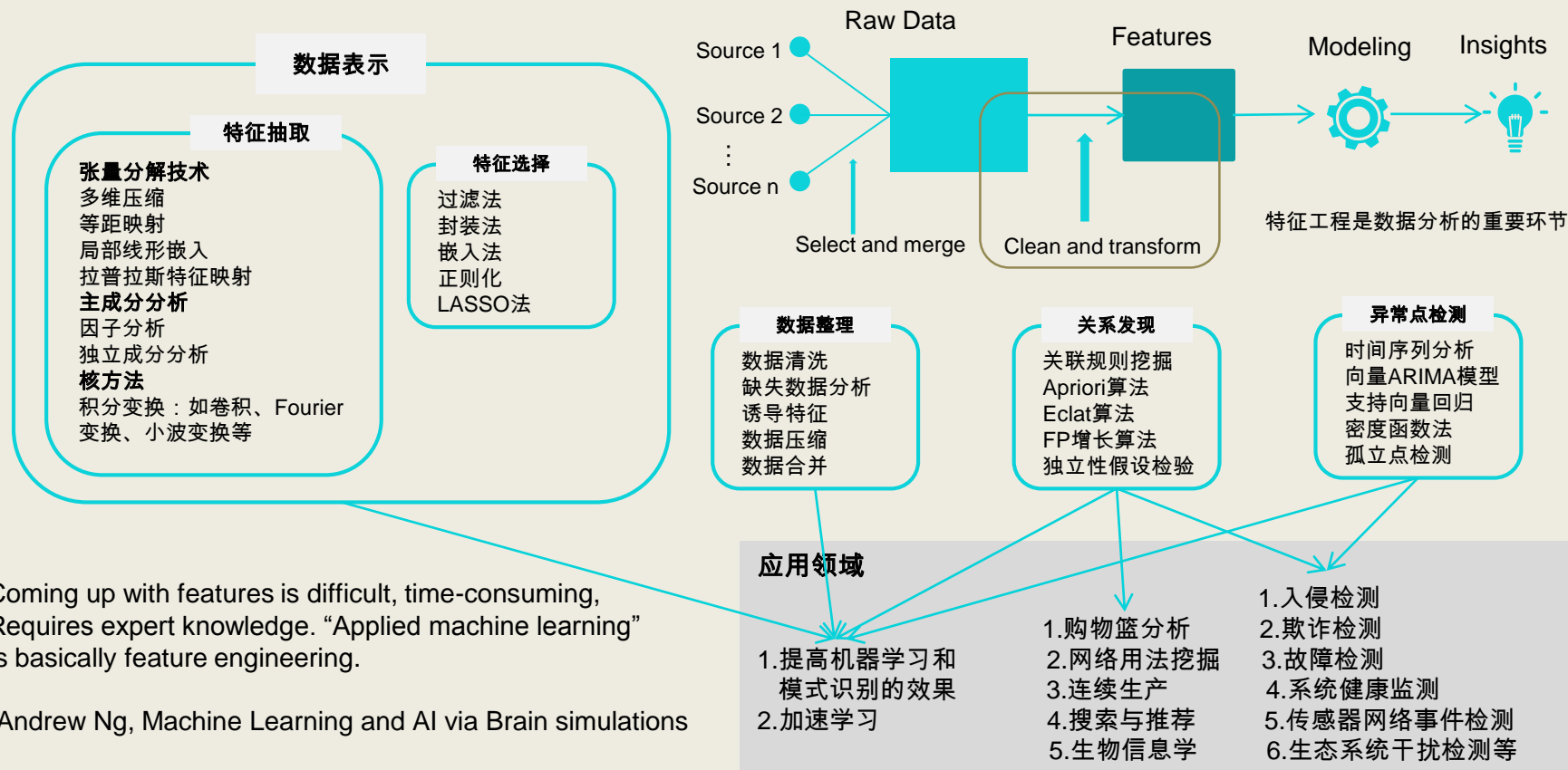


Wavelet transform: viewing data at different levels

小波变换：在不同层面看数据

如何看待数据：特征工程关键技术洞察

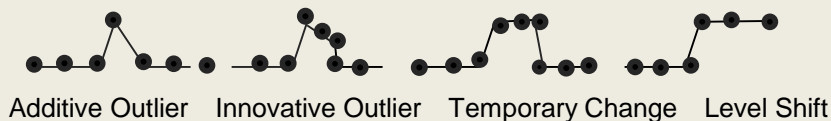
How to view data: key technology insights into feature Engineering



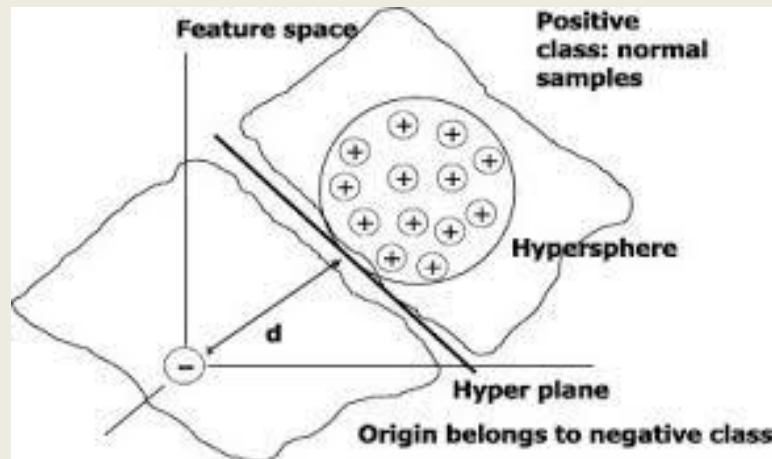
异常点检测：时间序列方法和机器学习方法

Outlier detection: time series method and machine learning method

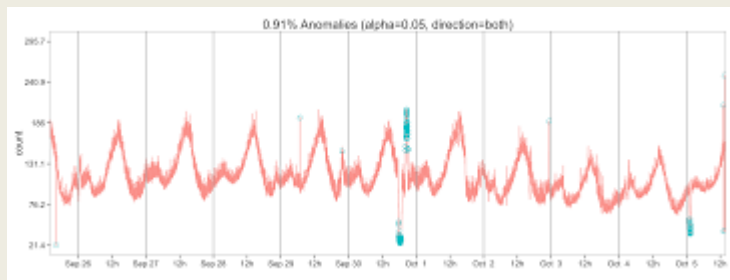
异常点检测



基于机器学习的异常点检测

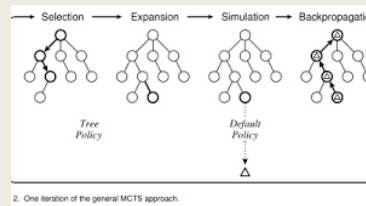
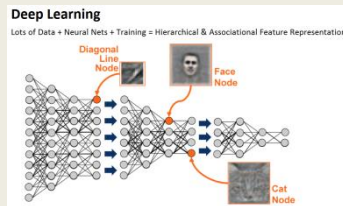
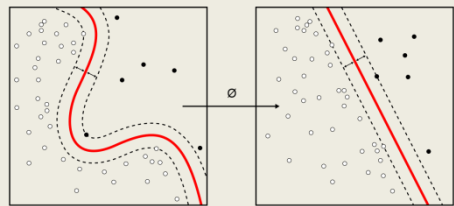


基于时间序列的异常点检测

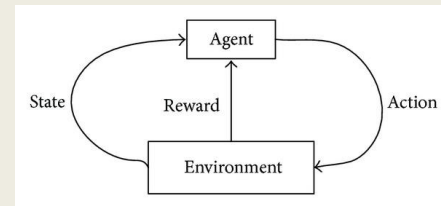


机器学习的现状：百花齐放、百家争鸣

The present situation of machine learning: a bloom



$$\text{posterior} \propto \text{prior} \times \text{likelihood}$$



经典方法 (2006年前)

深度学习 (2006-?)

Monte Carlo计算方法

统计决策学习

基于单一数据表示的最优化

特征的学习和数据表示的层级化

用随机模拟方法求近似解

基于知识的小样本学习和决策规划学习

方法	优点	缺点
经典方法	<ol style="list-style-type: none">1. 数学理论清楚，一般都经过严格论证；2. 很多问题归结为最优化。	<ol style="list-style-type: none">1. 数据质量和特征选择直接影响结果；2. 过拟合现象比较普遍，稳健性相对于贝叶斯方法差一些。
深度学习	<ol style="list-style-type: none">1. 在数据表示上具有天生优势；2. 在图像分析、语音识别等具体应用上效果明显	<ol style="list-style-type: none">1. 完全数据驱动，在知识表示上欠缺；2. 缺乏坚实的理论基础，难于推广；3. 调参数困难，学习周期太长。

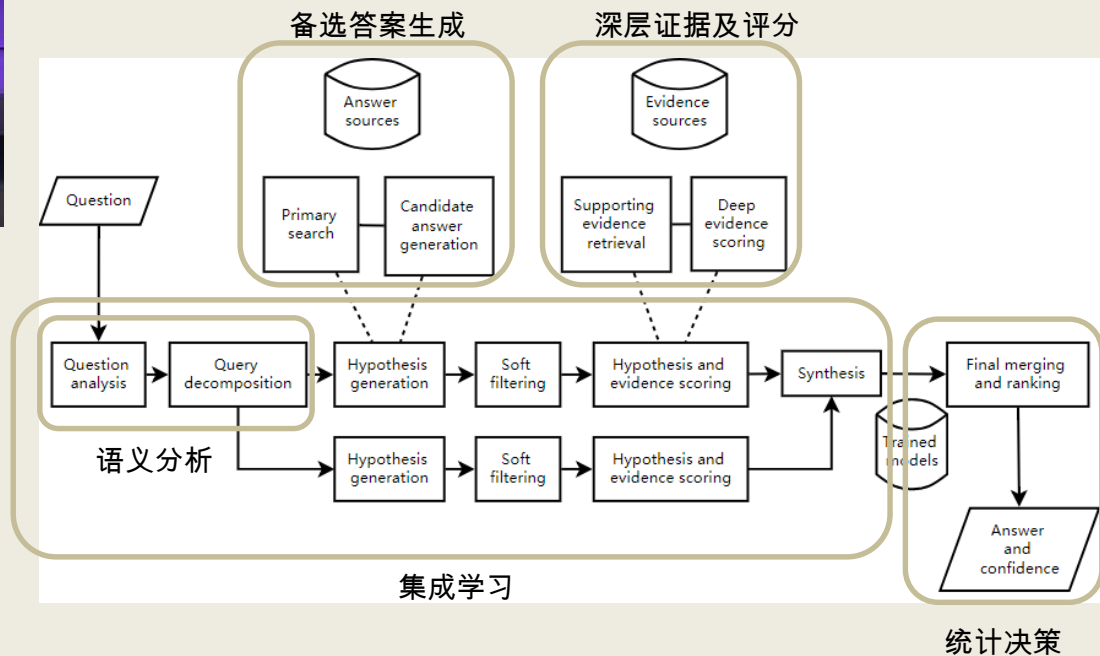
随机模拟方法的大量应用：例如，AlphaGo的核心技术Monte Carlo树搜索

更接近人类的学习方式

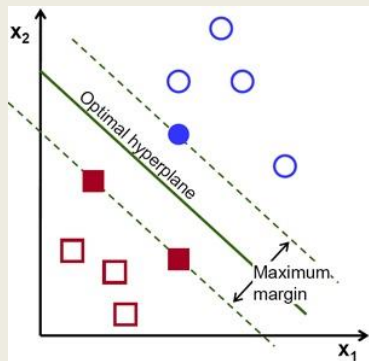
The inspiration from IBM Watson intelligent question answering system : traditional methods are also good.



Watson智能问答系统给我们的启发是：传统机器学习的方法，外加知识推理、语义分析、学习策略（如集成学习）、智能策略，经过精心的设计依然有强大的爆发力和令人瞩目的效果



Deep learning: the rise of connectionist AI (2006-?)



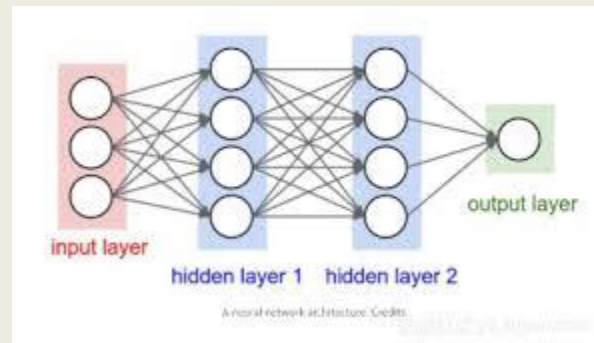
SVM



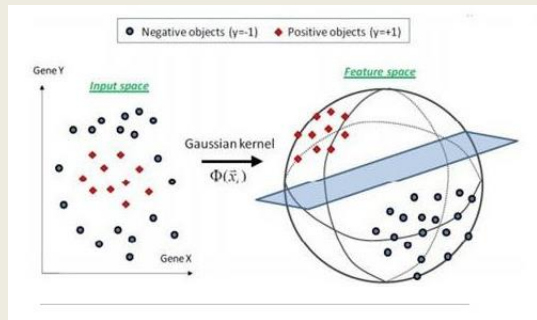
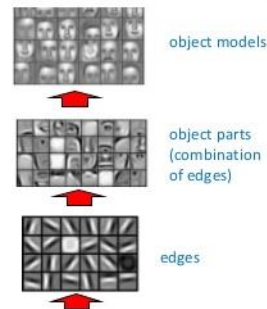
如何形式化地描述人脸表情的各种模式？



数据表示成为机器学习的难言之隐和巨大障碍

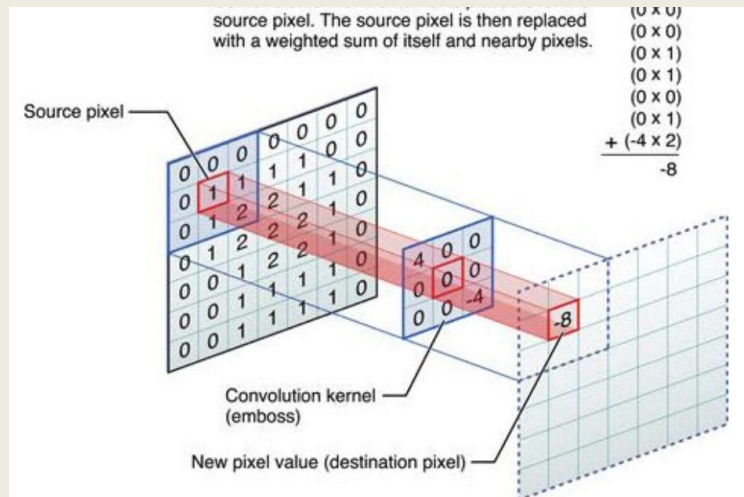


Deep Learning: Learns increasingly complex features



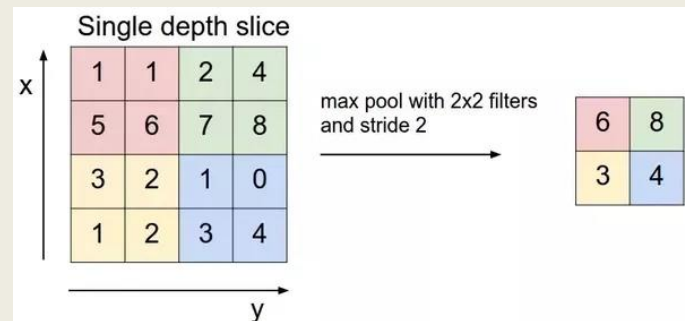
Kernel Method

Disadvantages of deep learning



卷积变换

- 无理论支撑，完全数据驱动，可解释性较弱
- 深度神经网络复杂，需要大规模数据训练
- 深度学习当前只适合做感知，不适合做认知和推理



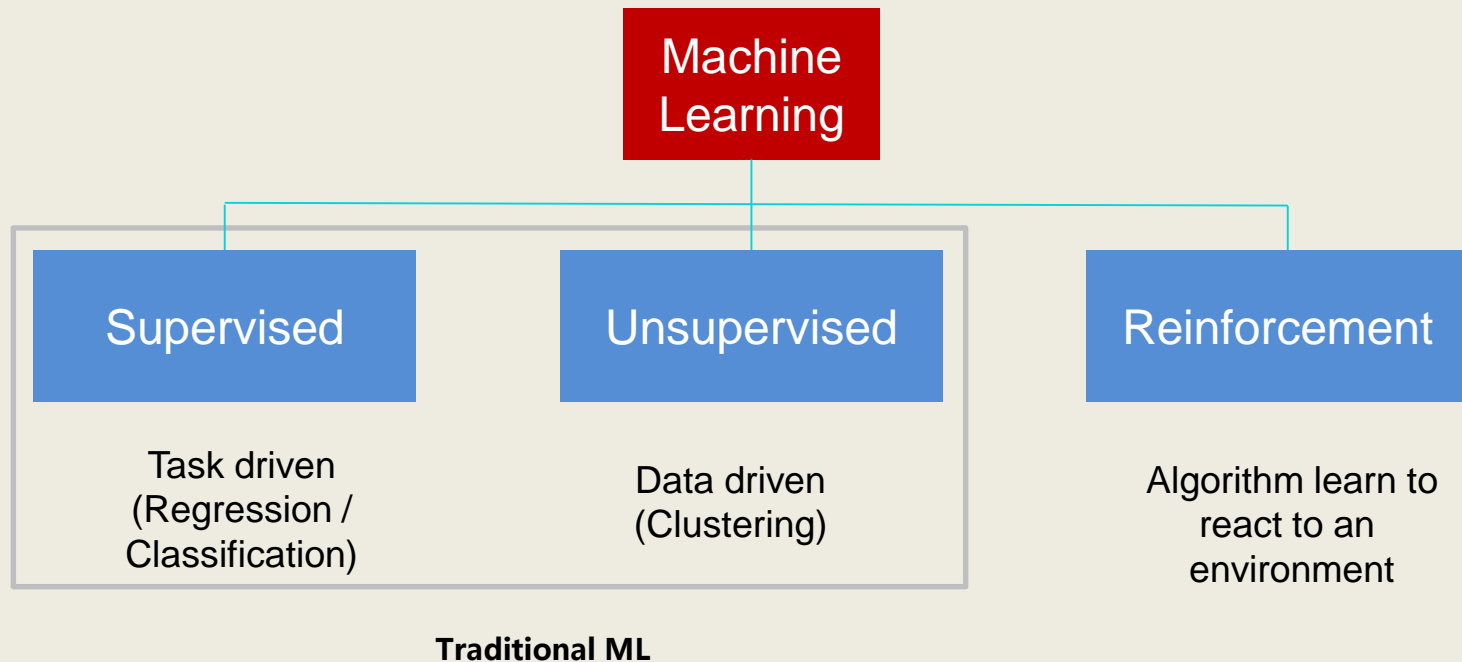
Max pooling



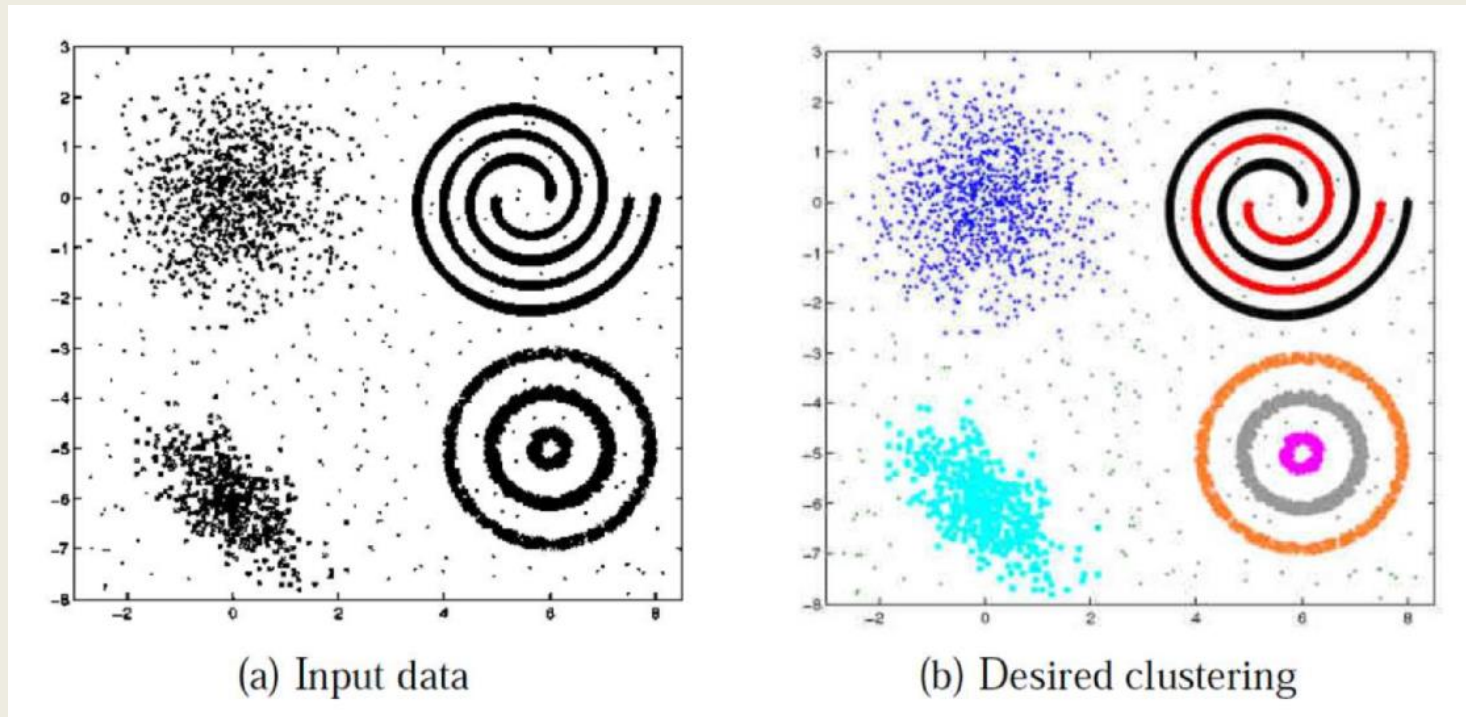
深度学习的民间行情

Three supervisor in ML (机器学习的三种 “老师”)

Types of Machine Learning

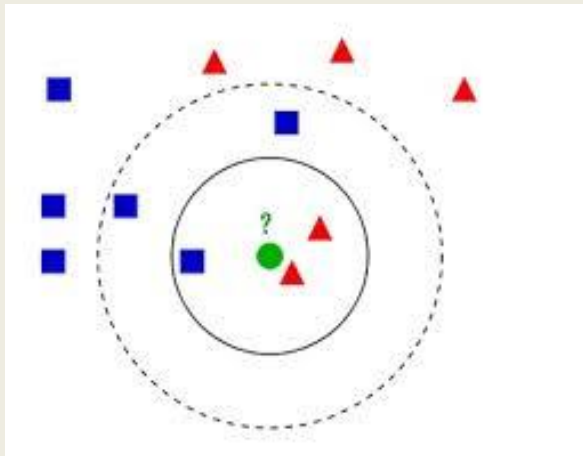


无监督学习的一个例子：K-means聚类

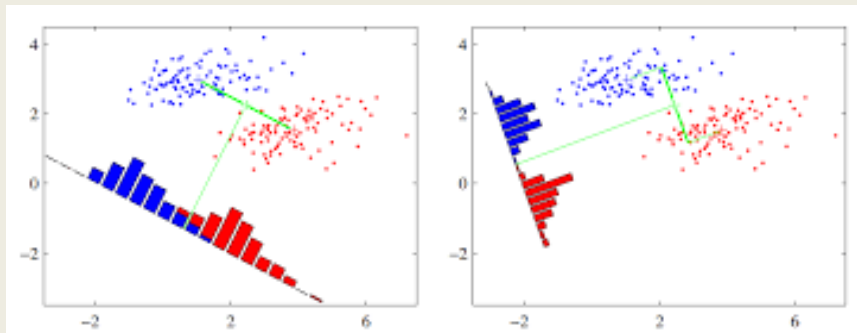


An example of unsupervised learning: K-means clustering

有监督学习的两个例子：k-近邻法和 Fisher 线性判别分类器



K-近邻法：近朱者赤，近墨者黑



Fisher线性判别法：使得Bayes error最小

Two examples of supervised learning: k-nearest neighbor and Fisher linear discriminant classifier.

Data-driven ML *vs* knowledge-based small-sample learning



Bayes Learning

Tenenbaum 教授 (MIT)



Reinforce Learning



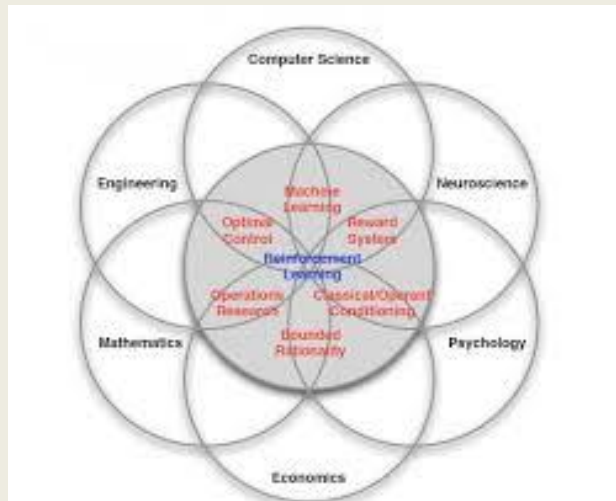
Jordan 教授(UC Berkeley)
Traditional ML



Sutton 教授 (Alberta大学)

Vapnik 教授 (Google)

- 机器学习教父Jordan为何不喜欢深度学习
- 机器学习大师Vapnik坚信下一次机器学习的革命是与知识紧密结合，以更贴近人类的方式进行。
 - 1) 强化学习可以不需要训练样本，直接从试验和错误中学习。
 - 2) 贝叶斯学习只需要极少量的样本。小样本时先验知识起关键作用；大样本的数据淹没先验，效果和经典统计学类似



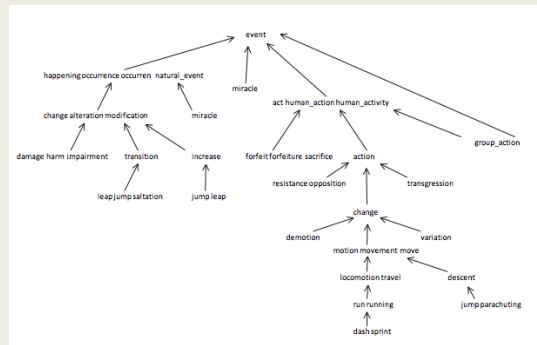
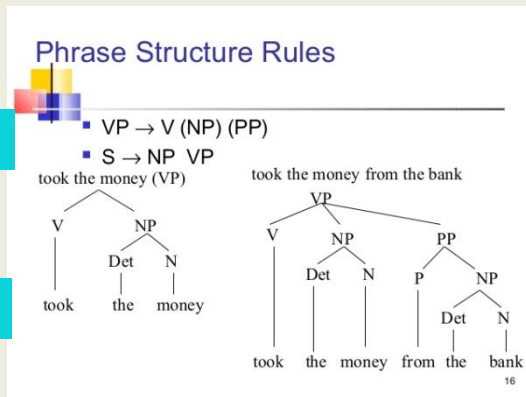
The relationship between reinforcement learning and other subjects
强化学习与其他学科的关系

什么是知识？如何表达？怎样使用不同来源的知识？

句法知识

句法分析

句法树



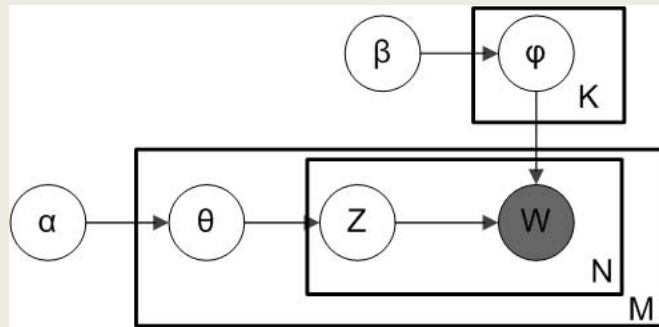
语义分析

浅层理解

语义知识

$\frac{1}{(3+a-b\,I)}$	
$\frac{1}{3+a-I\,b}$	
$\text{eval}(\%)$	$\frac{a+3}{(a+3)^2+b^2} + \frac{I\,b}{(a+3)^2+b^2}$
$\text{abs}(\%)$	
$\frac{1}{ 3+a-I\,b }$	
$\text{eval}(\%)$	$\frac{1}{\sqrt{(a+3)^2+b^2}}$
$\text{sqrt}(\%)$	$\frac{1}{((a+3)^2+b^2)^{1/4}}$
$\text{sqrt}(a+b\,I)$	$\sqrt{a+b\,I}$
$\text{eval}(\%)$	$\frac{1}{2}\sqrt{2\sqrt{a^2+b^2}+2a} - \frac{1}{2}I\,\text{csgn}(-b+I\,a)\sqrt{2\sqrt{a^2+b^2}-2a}$

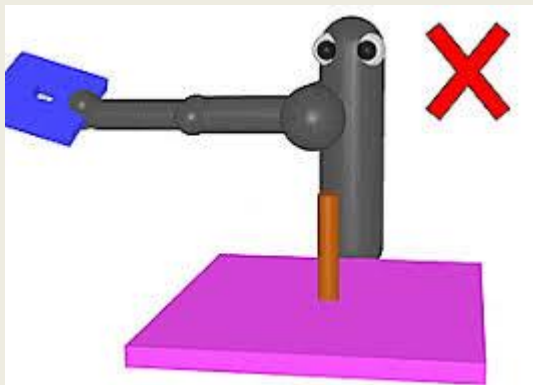
符号计算



LDA中的先验分布和变量约束关系

Reinforce Learning: the future star in ML

(强化学习：机器学习的未来之星)



教给机器一些基本动作



利用强化学习教会机器根据图片画水彩画

Research on the application of AI in several famous IT companies

几家著名IT公司的人工智能应用研究侧重点

Companies



The highlights of AI



R&D focus

自然语言处理
智能问答
认知计算
知识工程
云服务











通用机器学习平台
深度学习与强化学习
知识图谱
语义理解 (图像、自然语言等)
机器翻译
智能交通、自动驾驶



人机对话、智能助手
机器翻译
命名实体识别与搜索

Research on the application of AI in several famous IT companies

几家著名IT公司的人工智能应用研究侧重点

Companies	The highlights of AI	R&D focus
		Human-machine interaction, Intelligent Serve 智能助手
	  	通用机器学习平台 推荐系统 云服务 智慧语音助理（例如，用于智能音箱）
		人机对话、智能助手 计算机视觉

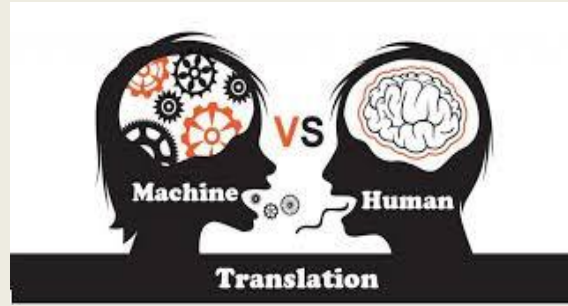
In 2016, Facebook, Amazon, Alphabet, IBM and Microsoft announced the formation of the Partnership for Artificial Intelligence, began to compete for standards and voice in artificial intelligence

The future work about AI

Application scenarios of machine learning in artificial intelligence and big data analysis (机器学习在人工智能和大数据分析中的应用场景)



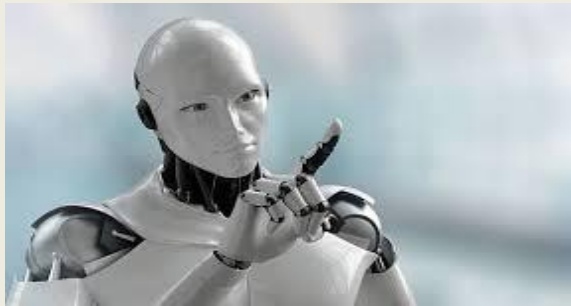
自动驾驶技术
Automatic driving technology



人机对话、机器翻译等
human-machine dialoge



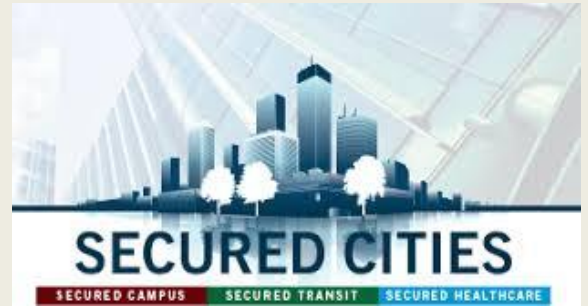
虚拟现实
Virtual Reality



机器人



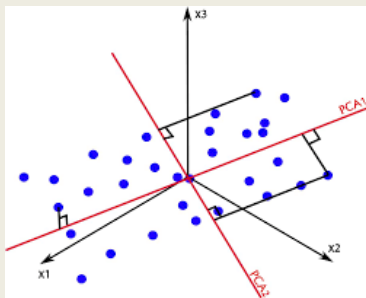
从大数据到小数据：数据清洗、特征工程、在线学习



Safe city, health monitoring
intelligent transportation

利用机器学习打造有影响力的大数据分析能力

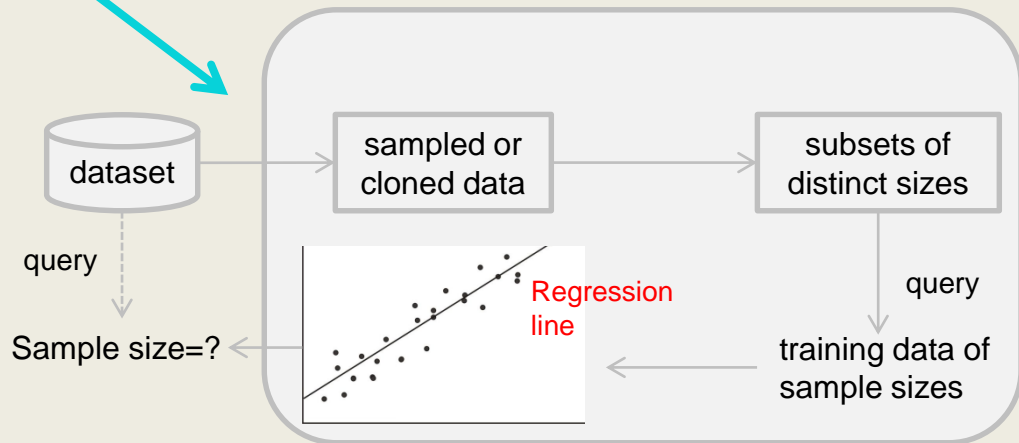
特征提取技术：
概率主成分分析



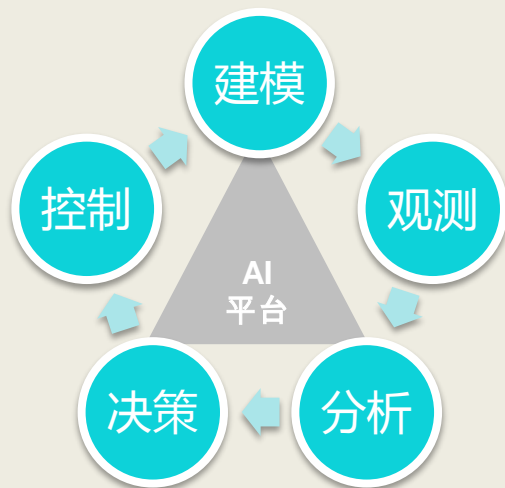
数据克隆技术



查询优化技术



采用有效的人工智能方法解决实际问题

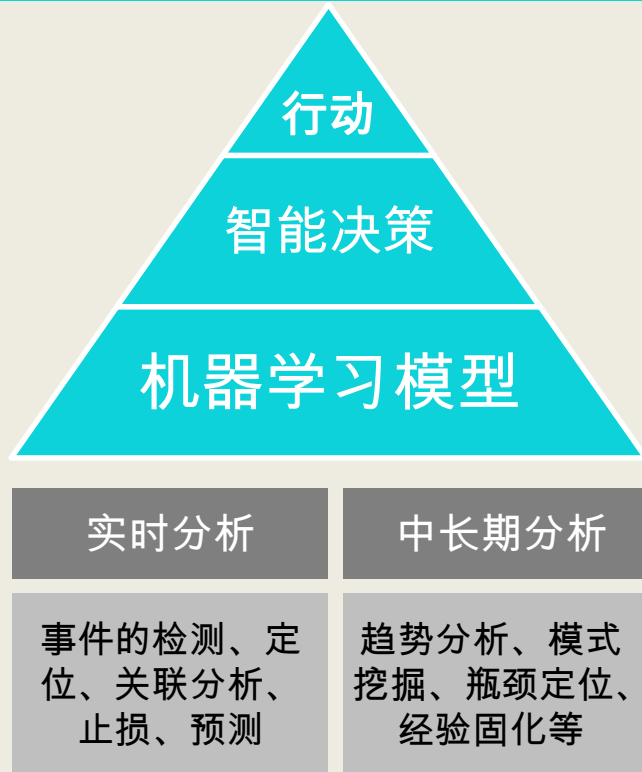


利用人工智能平台，基于领域知识迅速搭建求解策略，并利用平台能力解决实际问题。

人工智能和机器学习方法

- 1) 领域知识图谱与专家系统
- 2) 时间序列分析
- 3) 强化学习
- 4) 深度学习
- 5) 贝叶斯学习
- 6) 传统机器学习方法
- 7) 特征选择和特征提取
- 8) 图像处理和分析
- 9) 自然语言处理
- 10) 统计决策
- 11) 启发式搜索
- 12) 数据压缩与重构
- 13) 图计算
- 14)

Intelligent operation and maintenance: cloud computing transforms operational data into intelligent decision making and action through algorithm cloud.
(智能运维：通过算法云把运营数据转化为智能决策和行动)



AI with Smart City (人工智能与智慧城市)



AI with Smart City

Exponential growth



Conclusion

◆ Rationally(理性) looking at the status of Artificial Intelligence :

(1) DL has achieved brilliant achievements in the field of images and sounds perception (or recognition), even surpass human beings in some areas, but they are still weak in reasoning and cognition.

深度学习在图像、语音的感知（或识别）领域取得了辉煌的成就，甚至超越了人类，但在推理、认知等方面仍十分欠缺。

(2) if there is no simulation environment or without a priori knowledge, reinforcement learning will be difficult to move a single step.

强化学习如果没有仿真环境、没有先验知识也将寸步难行。

◆ The development trend of artificial intelligence is towards integration:

traditional machine learning + Deep Learning + reinforcement learning + knowledge reasoning + intelligent decision making.

人工智能的发展趋势是走向融合：**传统机器学习+深度学习+强化学习+知识推理+智能决策**

◆ The future of artificial intelligence lies in its application:

Integrating AI into core business and demonstrating AI technology with actual results

人工智能的前途在于应用：将AI融入到核心业务中去，用实际效果来证明AI技术