



医学信息学 第四章

智慧医疗

医学信息学是以信息学、信息管理和信息技术为依托，研究医学领域中的信息现象和信息规律，用于医学决策和管理的一门交叉学科。学科以生命信息现象和人类社会活动中的医学信息为研究对象。

谭积斌

讨论

课堂讨论

谈一谈你平时遇上的人工智能



智能初探

A preliminary study of Intelligence



近几年来我国人工智能领域在政策层面上提速非常明显

有来自国务院、工信部、发改委、财政部、科技部、网信办等部门的若干重要政策文件。

时间	部门	文件
2015-5-20	国务院	《中国制造2025》：智能制造被定位为中国制造的主攻方向；
2015-7-5	国务院	《“互联网+”行动指导意见》
2016-4	工信部、发改委、财政部	《机器人产业发展规划(2016-2020年)》
2016-5-23	发改委、科技部、工信部、网信办	《“互联网+”人工智能三年行动实施方案》
2016-7-28	国务院	《“十三五”国家科技创新规划》
2016-9-01	发改委	《国家发展改革委办公厅关于请组织申报“互联网+”领域创新能力建设专项的通知》
2016-9-12	国务院	《费品标准和质量提升规划(2016—2020年)》
2016-12-19	国务院	《“十三五”国家战略性新兴产业发展规划的通知》：
2017-1-16	中共中央办公厅、国务院	《关于促进移动互联网健康有序发展的意见》
2017-07-20	国务院	《新一代人工智能发展规划》

人工智能概念

Artificial intelligence concept

“

人工智能是研究人类智能活动规律，构造具有一定智能的人工系统，研究如何让计算机去完成以往需要人的智能才能胜任的工作的基本理论、方法和技术。

Artificial intelligence is a basic theory, method and technology that studies the law of human intelligence activities, constructs an artificial system with certain intelligence, and studies how to make computers complete the work that requires human intelligence in the past.

— 人工智能 —

本质是？目的是？方法和过程是？



计算机科学的一个分支，用一定的技术和应用系统生产出新的机器，使之实现目前必须借助人类智慧才能实现的任务。

数 学	逻辑学	归纳学	统计学	系统学	控制学	工程学	计算机科学
哲学	心理学	生物学	神经科学	认知科学	仿生学	经济学	语言学



让机器像人类一样思考

机器学习



让机器像人类一样看懂世界

视觉识别



让机器像人类一样运动

运动控制



让机器像人类一样听懂

语音识别



弱人工智能

在各个领域的应用较广泛，
能够模仿人脑的基本功能。

Google搜索引擎、智能手机、Alpha
Go
(感知、记忆、学习和决策)



强人工智能

学术界和产业界研究和发
展的目标。

需要结合情感、认真和推理等高阶智能，能够进行
思考，计划，解决问题、先后向思维、理解复杂概
念、快速学习和从经验中学习操作。



超级人工智能

其能力超过人类，且能不
断自我学习和进化。

不讨论，都超越人类了，还让人类去定义？



人工智能分类

Artificial intelligence classification

20/50



历史沿革

Historical evolution

技术萌芽

人工智能的基本假设是人类的思考过程可以机械化（三段式推理）——机械化的逻辑推理可以用数学的方式计算——二进制能够实现任何的计算，也就能模拟任何的推理。这就为人工智能发展奠定基础。这一时期的关键人物是图灵，1950年，图灵发表了一篇划时代的论文，文中预言了创造出具有真正智能的机器的可能性。

例如：知识分子都是应该受到尊重的，人民教师都是知识分子，所以，人民教师都是应该受到尊重的。

1956

第一次高潮期

1956年之后的十几年是人工智能的黄金年代，计算机被用于证明数学定理、解决代数应用题等领域。A. Newell和H. Simon研发的“逻辑理论家（Logic Theorist）”证明了《数学原理》中全部52条定理，其中某些证明比原著更加巧妙。人们几乎无法相信机器原来可以如此智能。这些成果让研究者对未来充满信心，认为完全智能的机器人二十年内就能出现。

1974

第一次低谷期

人们渐渐发现仅仅具有逻辑推理能力远远不够实现人工智能，许多难题并没有随着时间推移而被解决，很多AI系统一直停留在了玩具阶段。之前的过于乐观使人们期望过高，又缺乏有效的进展，许多机构逐渐停止了对AI研究的资助。人工智能遭遇了第一次低谷。



历史沿革

Historical evolution

1980

第二次高潮期

卡耐基梅隆大学为DEC公司设计了一个名为XCON的专家系统，取得巨大成功。许多公司纷纷效仿，开始研发和应用专家系统。专家系统依赖的知识工程因而也成为AI研究的焦点。与此同时，神经网络算法被提出和广泛应用。神经网络被广泛的应用于模式识别、故障诊断、预测和智能控制等多个领域。AI迎来了新一轮高潮。

1987

第二次低谷

从80年代末到90年代初，专家系统所存在的应用领域狭窄、知识获取困难、维护费用居高不下等问题逐渐暴露出来，十年前日本人宏伟的第五代计算机计划也宣告失败。AI遭遇了一系列财政问题，进入第二次低谷。

1993

第三次高潮期

随着计算机性能的高速发展，海量数据的累积和AI研究者的不懈努力，人工智能在许多领域不断取得突破性成果，掀起新一轮高潮。

1997

IBM的国际象棋机器人深蓝（Deep Blue）战胜国际象棋世界冠军卡斯帕罗夫，引起世界范围内轰动。

2006

Geoffrey Hinton提出深度学习。在接下来的若干年，借助深度学习技术，包括语音识别、计算机视觉在内的诸多领域都取得了突破性的进展。

2009

谷歌曾经成功预测H1N1流感爆发

2011

IBM的问答机器人Watson在美国问答节目《Jeopardy!》上击败两位人类冠军选手。

2012

微软就在“21世纪的计算”大会上展示了一个全自动同声传译系统，它将演讲者的英文演讲实时转换成与他的音色相近、字正腔圆的中文。

2016

AlphaGo在两个公开围棋网站上与中日韩数十位围棋高手进行快棋对决，连胜60局无一败绩，包括对当今世界围棋第一人柯洁连胜三局。

2017

百度的小度机器人在《最强大脑》中战胜人类“脑王”。搜狗的问答机器人汪仔在《一站到底》中战胜哈佛女学霸。

浪潮袭来，人工智能席卷各个领域，正在深刻地改变着人类社会的方方面面。



历史沿革

Historical evolution



1956年达特茅斯会议举行， Marvin Minsky、John McCarthy、Claude Shannon、Nathan Rochester、 Allen Newell、Herbert Simon等顶尖科学家汇聚一堂，共同确定了人工智能的名称和任务，标志着人工智能这门学科的正式诞生。与会的每一位科学家都在人工智能发展的第一个十年中做出了重要贡献。

（五位达特茅斯会议的主要参与人员在人工智能60周年纪念会上合影）

达特茅斯会议 五老星

Dartmouth Conference

AI浪潮

Artificial intelligence tide

2016年3月，人工智能Alpha Go与李世石的围棋大战引发了广泛关注，而这场对弈以Alpha Go的胜利告终更是激发了人们对人工智能的激烈探讨，直接助推了AI浪潮的爆发。

- 3月23号，微软聊天机器人“Tay”学会种族歧视言论；
- 9月28号，五大科技巨头组建超级AI联盟；
- 10月12号，美国白宫发布人工智能白皮书；
- 10月17号，微软AI的语音识别能力首次超过人类；
- 11月25号，人工智能改善癌症诊断疗法；
- 12月5号，亚马逊开展新零售业务Amazon Go；
- 12月19号，扎克伯格开发AI助手Jarvis；



AI驱动要素

Driving elements of artificial intelligence

人工智能之所以能在近年来掀起新一轮高潮，主要归根于三大驱动要素



算 法

Calculation method



大数据

Big data



计算力

Computing force

AI驱动要素

Driving elements of artificial intelligence



算法

Calculation method

机器学习技术一直是人工智能发展的核心推动力。而机器学习则需要借助一定的算法才能实现。

算法 (Algorithm) 是指解题方案的准确而完整的描述，是算法代表着用系统的方法描述解决问题的策略机制。



现实生活

解决问题：放假回家

解决算法：吃早餐——打车（公车）——车站（飞机场）



专业领域

解决问题：下图中一共有多少人？

解决算法1：

$N=0$ ：

$N=1+1+1\ldots$



AI驱动要素

Driving elements of artificial intelligence



算法

Calculation method

机器学习技术一直是人工智能发展的核心推动力。而机器学习则需要借助一定的算法才能实现。

算法 (Algorithm) 是指解题方案的准确而完整的描述，是算法代表着用系统的方法描述解决问题的策略机制。



现实生活

解决问题：放假回家

解决算法：吃早餐——打车（公车）——车站（飞机场）



专业领域

解决问题：下图中一共有多少人？

解决算法2：

```
N=0;  
N++;  
N=N+1;  
...
```



AI驱动要素

Driving elements of artificial intelligence



算法

Calculation method

机器学习技术一直是人工智能发展的核心推动力。而机器学习则需要借助一定的算法才能实现。

算法 (Algorithm) 是指解题方案的准确而完整的描述，是算法代表着用系统的方法描述解决问题的策略机制。



现实生活

解决问题：放假回家

解决算法：吃早餐——打车（公车）——车站（飞机场）



专业领域

解决问题：下图中一共有多少人？ 解决算法3：

$N=0$;

For each person in pic

Set $N = N + 1$;



AI驱动要素

Driving elements of artificial intelligence



算法

Calculation method

机器学习技术一直是人工智能发展的核心推动力。而机器学习则需要借助一定的算法才能实现。

算法 (Algorithm) 是指解题方案的准确而完整的描述，是算法代表着用系统的方法描述解决问题的策略机制。



现实生活

解决问题：放假回家

解决算法：吃早餐——打车（公车）——车站（飞机场）



专业领域

解决问题：下图中一共有多少人？ 解决算法4：

$N=0$;

For each 2 person in pic

Set $N= N+2$;

If 1 person remains

Set $N= N+1$;



AI驱动要素

Driving elements of artificial intelligence



算法

Calculation method

机器学习技术一直是人工智能发展的核心推动力。而机器学习则需要借助一定的算法才能实现。



应用领域

传统的统计机器学习技术应用：搜索引擎、广告、商品推荐、内容过滤等。常见算法有朴素贝叶斯分类、逻辑回归、聚类、ELO、独立成分分析…



常见的推荐算法

算法一般会根据以下这三个维度来预估推荐

文本内容

个人信息

周围环境

AI驱动要素

Driving elements of artificial intelligence



算法

Calculation method

机器学习技术一直是人工智能发展的核心推动力。而机器学习则需要借助一定的算法才能实现。



应用领域

传统的统计机器学习技术应用：搜索引擎、广告、商品推荐、内容过滤等。常见算法有朴素贝叶斯分类、逻辑回归、聚类、ELO、独立成分分析…



常见的推荐算法

文本内容

系统获得一个文本内容时，会通过一套算法，迅速分析内容的主要特征

质量

关键词

兴趣标签

作者

相似文章

AI驱动要素

Driving elements of artificial intelligence



算法

Calculation method

机器学习技术一直是人工智能发展的核心推动力。而机器学习则需要借助一定的算法才能实现。



应用领域

传统的统计机器学习技术应用：搜索引擎、广告、商品推荐、内容过滤等。常见算法有朴素贝叶斯分类、逻辑回归、聚类、ELO、独立成分分析…



常见的推荐算法

个人信息

用户的点击内容、播放视频、浏览时间、点赞、评论、转发。后台量化成标签，系统会给用户账号贴上暂时标签，再结合用户基本信息刻画用户画像。

适用机型

性别

年龄

职业

兴趣



AI驱动要素

Driving elements of artificial intelligence



算法

Calculation method

机器学习技术一直是人工智能发展的核心推动力。而机器学习则需要借助一定的算法才能实现。



应用领域

传统的统计机器学习技术应用：搜索引擎、广告、商品推荐、内容过滤等。常见算法有朴素贝叶斯分类、逻辑回归、聚类、ELO、独立成分分析...



常见的推荐算法

周围环境

当一个人在不同的环境下，会想看不同的内容

季节变化

如果你是个吃货，那么在开海时期你就想看怎么挑海鲜，怎么做海鲜...

地理位置

如果你在农村老家，那么就不会给你推探店内容

AI驱动要素

Driving elements of artificial intelligence



大数据

Big data

深度学习的多层神经网络结构复杂，参数众多，需要大规模的数据才能训练生成有效的模型。

得益于互联网、社交媒体的发展，带宽大幅提高，存储硬件成本的降低，全世界的数据规模在爆发式增长，人类进入大数据时代。



计算力

Computing force

深度学习使用海量数据训练复杂的多层神经网络模型，需要强大的计算能力支撑。

传统的CPU向GPU转变进行模型训练，速率提升。各种定制芯片、分布式计算技术纷纷被用于加速深度学习。

人工智能应用分支

Artificial intelligence connotation

人工智能是计算机科学研究领域的一个重要分支，又是众多学科的一个交叉学科，它企图了解智能的实质，并生产出一种新的能以人类智能相似的方式做出反应的智能机器

机器学习

模式识别

数据挖掘

专家系统

分支方向

Branch field



机器学习

机器学习属于人工智能研究与应用的一个分支领域。机器学习的研究更加偏向理论性，使用计算机作为工具并致力于真实实时的模拟人类学习方式，并将现有内容进行知识结构划分来有效提高学习效率。

机器学习就是利用计算机从历史数据中找出规律。通过大量的人类标记过的数据来训练模型，让模型学习到样本和标签的关系。这里的数据可以是图像，语音，用户画像等等。

理论假设

二元分类：买水果过程



分支方向

Branch field



机器学习

机器学习已经有了十分广泛的应用，生物特征识别、搜索引擎、医学诊断、检测信用卡欺诈、证券市场分析、DNA序列测序、语音和手写识别、战略游戏和机器人运用。

1.线性回归。
2.逻辑回归。
3.决策树。

4.支持向量机。
5.贝叶斯模型。
6.正则化模型。

7.模型集成
8.神经网络。

这些预测模型中的每一个都基于特定的算法结构，参数都是可调的。训练预测模型涉及以下步骤：



机器学习

Machine learning



提出问题：判断猫能不能抓老鼠

问题剖析：凭借人生活经验，一只猫能不能抓老鼠主要看这只猫的颜色、胖瘦、能否爬树等等方面

采集数据：

颜色 (x_1)	胖瘦 (x_2)	能否爬树 (x_3)	能否抓老鼠 (y)
黑色	胖	不能	不能
白色	瘦	能	能
...	...	...	...

机器学习

Machine learning



向量化：把生活信息转化为可计算数据

颜色 (x_1)	胖瘦 (x_2)	能否爬树 (x_3)	能否抓老鼠 (y)
黑色 (1)	胖 (0)	不能 (0)	不能 (0)
白色 (0)	瘦 (1)	能 (1)	能 (1)
...	...	...	...

$$x_{11}, x_{12}, x_{13}, y_1 = (1, 0, 0, 0)$$

$$x_{21}, x_{22}, x_{23}, y_2 = (0, 1, 1, 1)$$

机器学习

Machine learning



模型选择：选择与问题解决相适应的模型

$$x_{11}, x_{12}, x_{13}, y_1 = (1, 0, 0, 0)$$

$$x_{21}, x_{22}, x_{23}, y_2 = (0, 1, 1, 1)$$

$$f(x_1, x_2, x_3) = a \cdot x_1 + b \cdot x_2^2 + c \cdot x_3^3$$

机器学习的目的就是解这个模型，把模型的权重给算出来

机器学习

Machine learning



模型训练：把向量化数据和随机权重值导入模型函数

$$x_{11}, x_{12}, x_{13}, y_1 = (1, 0, 0, 0)$$

$$x_{21}, x_{22}, x_{23}, y_2 = (0, 1, 1, 1)$$

$$f(x_1, x_2, x_3) = a \cdot x_1 + b \cdot x_2^2 + c \cdot x_3^3$$

$$f(x_{11}, x_{12}, x_{13}) = a \cdot x_{11} + b \cdot x_{12}^2 + c \cdot x_{13}^3 = \text{output}_1$$

$$\text{error}_1 = y_1 - \text{output}_1$$

$$a = \text{fx}(\text{error}_1)$$

.....

那么我们就可以通过这个误差值，再进行斜率和增量计算，由此来更新abc的值！
这样，我们就能够得到了abc的初始值，但这个值肯定也还存在误差，因为我们只做了一遍训练。我们再把许多的数据记录导入到现在的模型函数中，继续计算，继续更新abc的值，通过不断地试错，就能够得到一个解决问题的模型函数。

机器学习

Machine learning

问题：预测小强会不会去看电影？

如花	小倩	小明	小强
0	0	1	0
1	1	1	1
1	0	1	1
0	1	1	0
1	1	0	?

模型选择：

$$F_x = x_{11} * w_1 + x_{12} * w_2 + x_{13} * w_3$$

X			weight	output
x11	x12	x13	w1	o1
x21	x22	x23	w2	o2
x31	x32	x33	w3	o3
x41	x42	x43		o4

机器学习

Machine learning

```
#1. 导入Numpy类
from numpy import random, array, dot, exp

#2. 加载数据集
X = array([[0, 0, 1], [1, 1, 1], [1, 0, 1], [0, 1, 1]])
y = array([[0, 1, 1, 0]]).T

#3. 设置随机权重
random.seed(1)
weights = random.random((3, 1)) * 2 - 1

#4. 模型训练-循环

for it in range(5000):
    #利用矩阵点乘一次性计算4个z出来
    z = dot(X, weights)
    #使用Sigmoid函数计算最终的output
    output = 1/(1+exp(-z))
    #计算误差
    error = y - output
    #计算斜率
    slope = output * (1-output)
    #计算增量
    delta = error * slope
    #更新权重
    weights = weights + dot(X.T, delta)

print(weights)
```

1. 导入Numpy类 2. 加载数据集

如花	小倩	小明	小强
0	0	1	0
1	1	1	1
1	0	1	1
0	1	1	0

3. 设置随机权重，设置初始权重都相同，避免计算错误

$$F_x = x_{11} * w_1 + x_{12} * w_2 + x_{13} * w_3$$

`weights = random.random((3, 1)) * 2 - 1`

//生成3个都在1行的0-1的随机数

//确保权重值范围在-1—1内，用随机数乘2-1

机器学习

Machine learning

```
for it in range(5000):  
    #利用矩阵点乘一次性计算4个z出来  
    z = dot(X, weights)  
    #使用Sigmoid函数计算最终的output  
    output = 1/(1+exp(-z))  
    #计算误差  
    error = y - output  
    #计算斜率  
    slope = output * (1-output)  
    #计算增量  
    delta = error * slope  
    #更新权重  
    weights = weights + dot(X.T, delta)  
print(weights)
```

- 4.模型训练，开展循环，训练5000次
- 5.引用模型函数，计算初步结果z
- 6.把结果z换算成0-1范围内的数，保证所有计算的结果都统一是一样的数值范围，方便后期的迁移引用。Sigmoid函数
- 7.计算误差、斜率、增量
- 8.更新权重

```
[[ 8.95950703]  
 [-0.20975775]  
 [-4.27128529]]
```

如花
小倩
小明

机器学习

Machine learning

按照现阶段主流分类来看，机器学习主要分为四类：

经典机器学习

（简单数据的特征分析）

强化学习

（没有数据但有数据交互环境）

深度学习

（针对复杂无特征的数据）

集成方法

（追求数据的质量）

监督学习

无监督学习

模型总是需要一个导师，即对应特征的标注，这样的话机器就可以基于这些标注学习进行进一步分类。

基于兴趣去分类用户、基于语言和主题分类文章、基于类型而分类音乐以及基于关键词分类电子邮件。

根据未知特征对目标进行分割，而由机器去选择最佳方式。

回归分析

分类分析

聚类

降维

关联

聚类是一种没有预先定义类的分类。

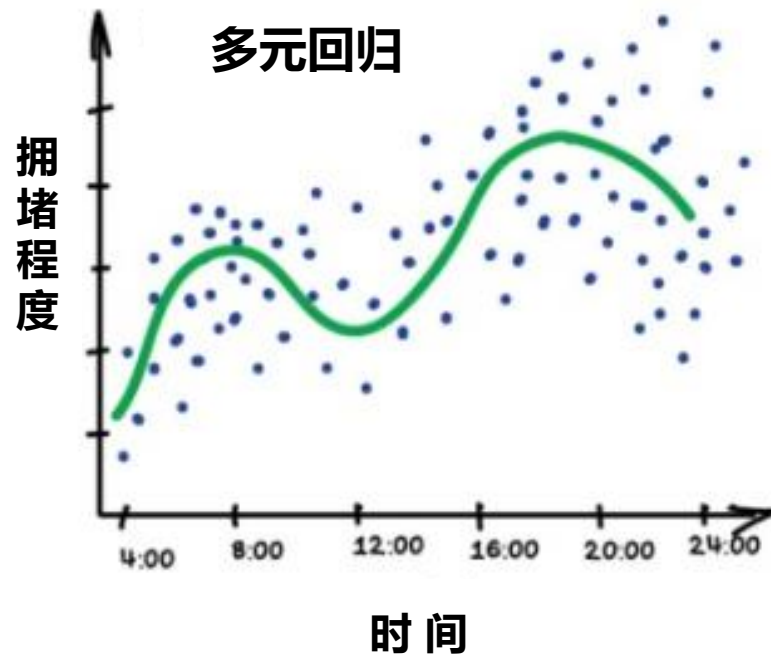
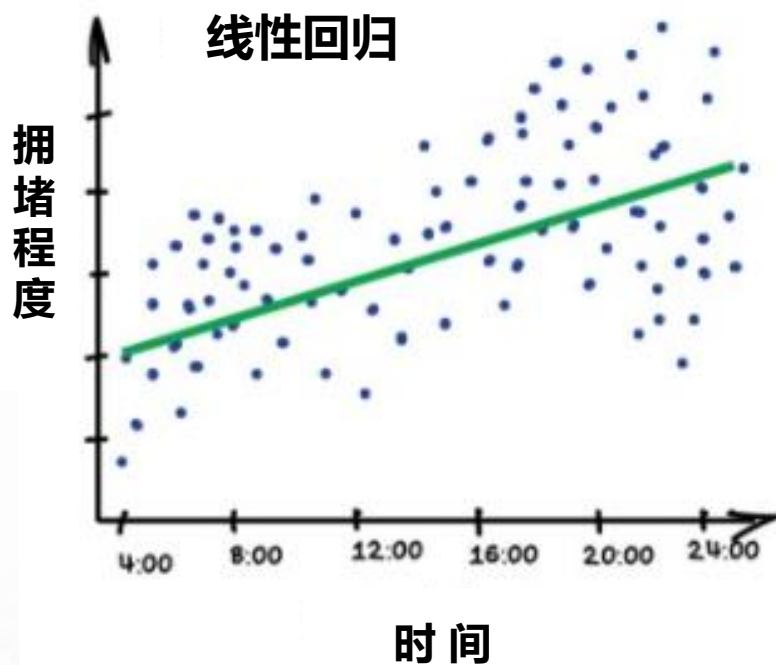
机器学习

Machine learning

回归分析

- 1.Linear Regression线性回归
- 2.Logistic Regression逻辑回归
- 3.Polynomial Regression多项式回归

.....

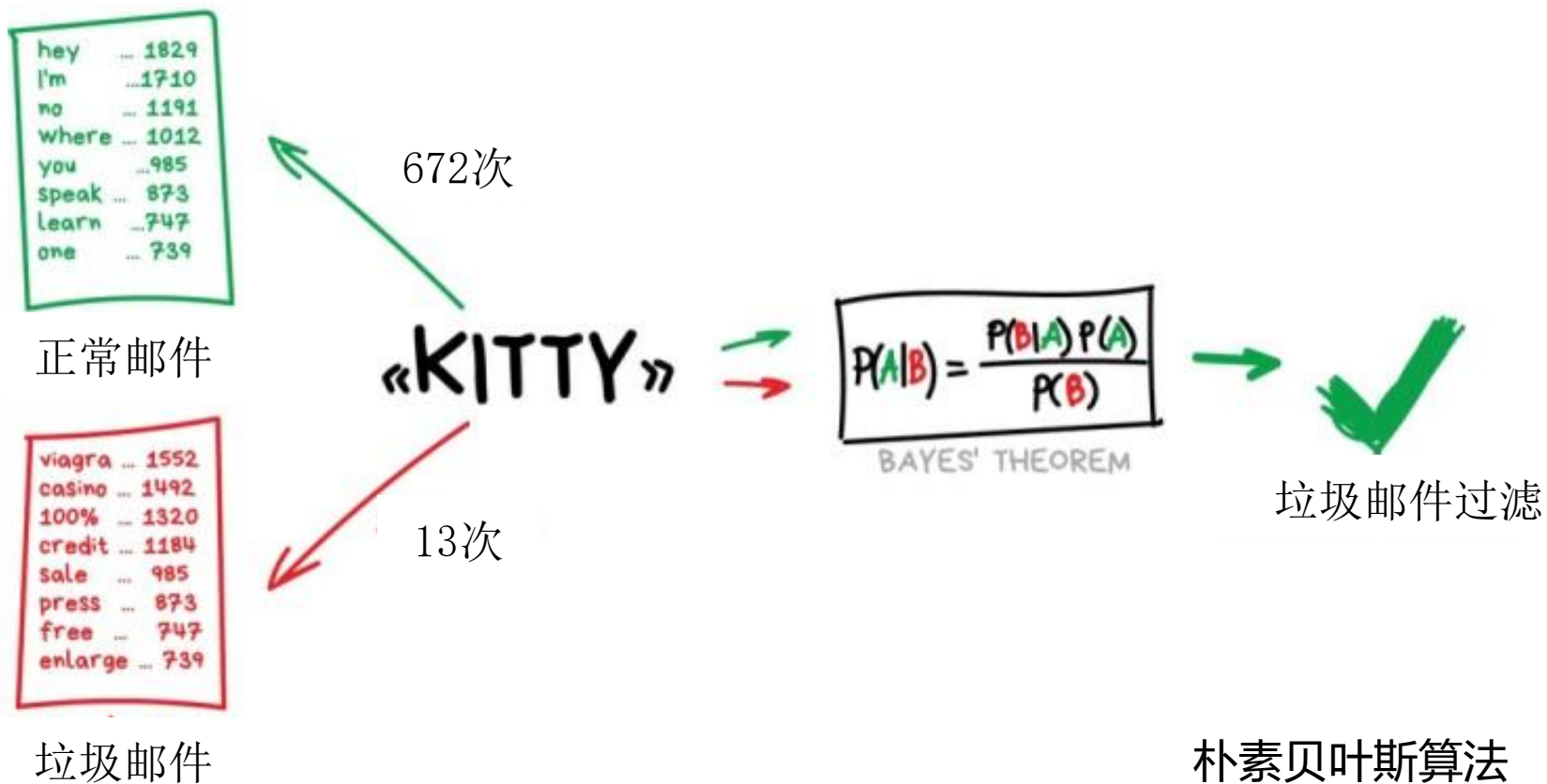


预测交通堵塞的回归分析

机器学习

Machine learning

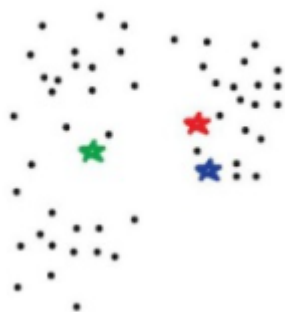
分类分析



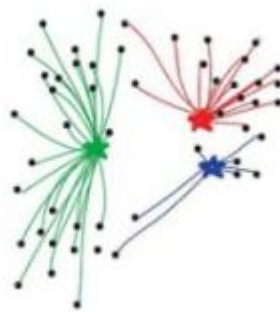
机器学习

Machine learning

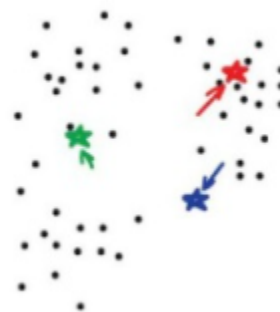
聚类分析



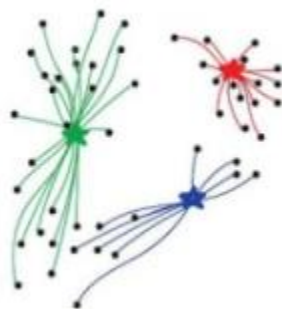
1. 随意摆放奶茶店



2. 看看买家如何选择?



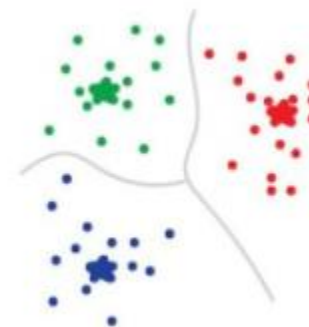
3. 把奶茶店迁移到人气中心



4. 再次观察并移动



5. 重复一百万次



6. 达到高内聚低耦合标准

模式识别

Artificial intelligence connotation



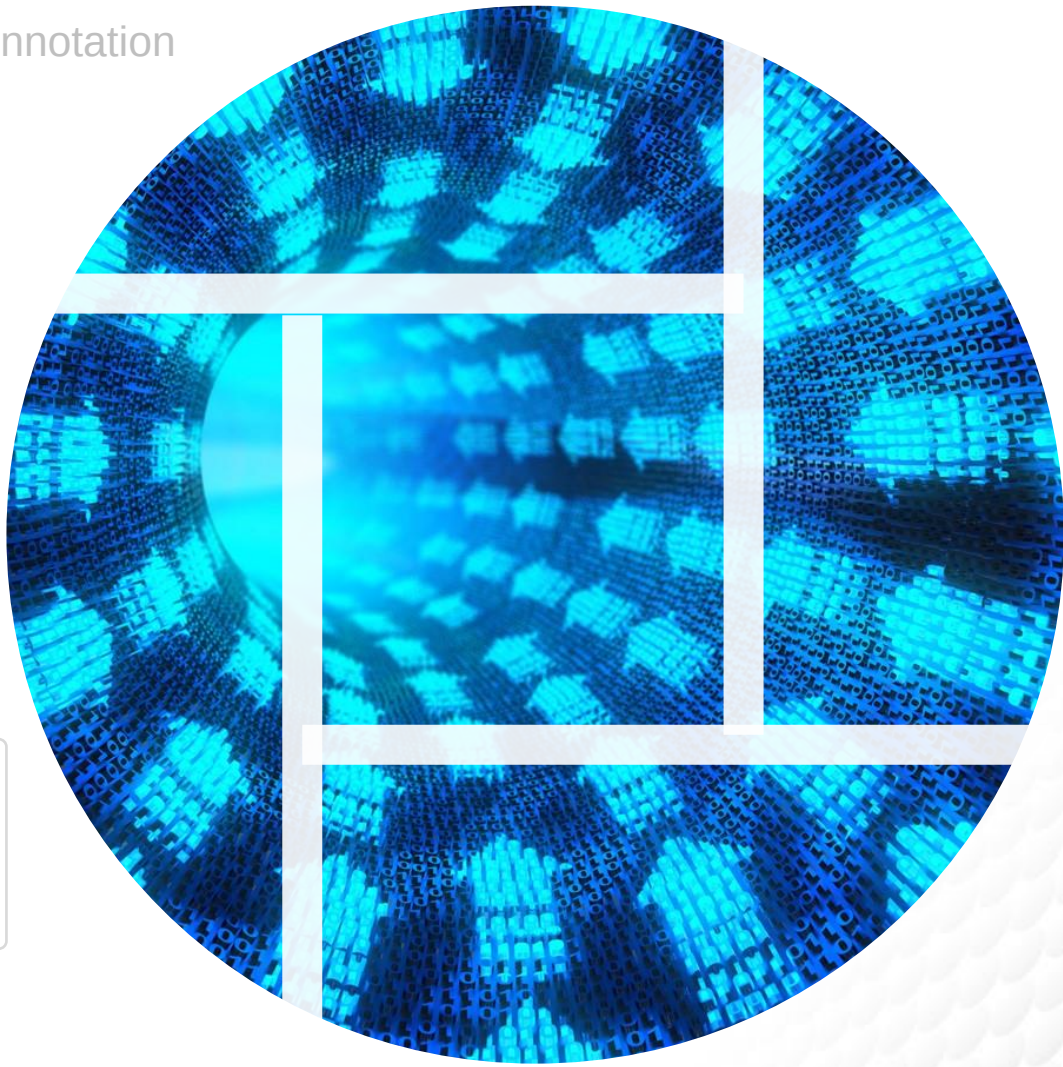
模式识别

指用机器模拟人的感知过程实现对感知数据(图像、视频、语音等)的模式分析与识别。

模式分类是模式识别的主要任务和核心研究内容：识别物品的视频

应用领域

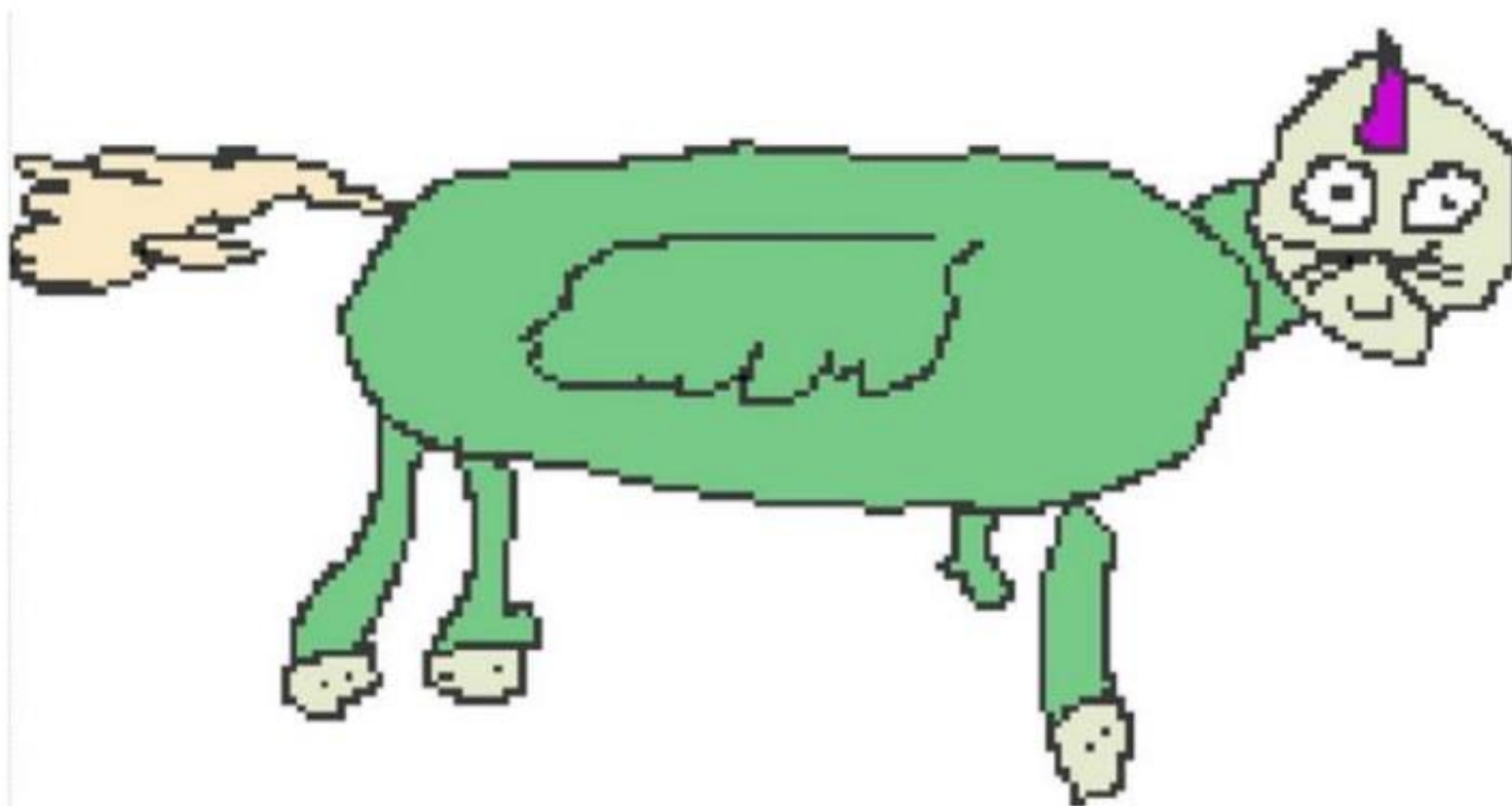
模式识别可用于文字和语音识别、人脸识别、指纹识别、遥感和医学诊断等方面。



模式识别

Artificial intelligence connotation

问题：这是一只猫嘛？



模式识别

Artificial intelligence connotation

问题：这是一只猫嘛？

提取特征

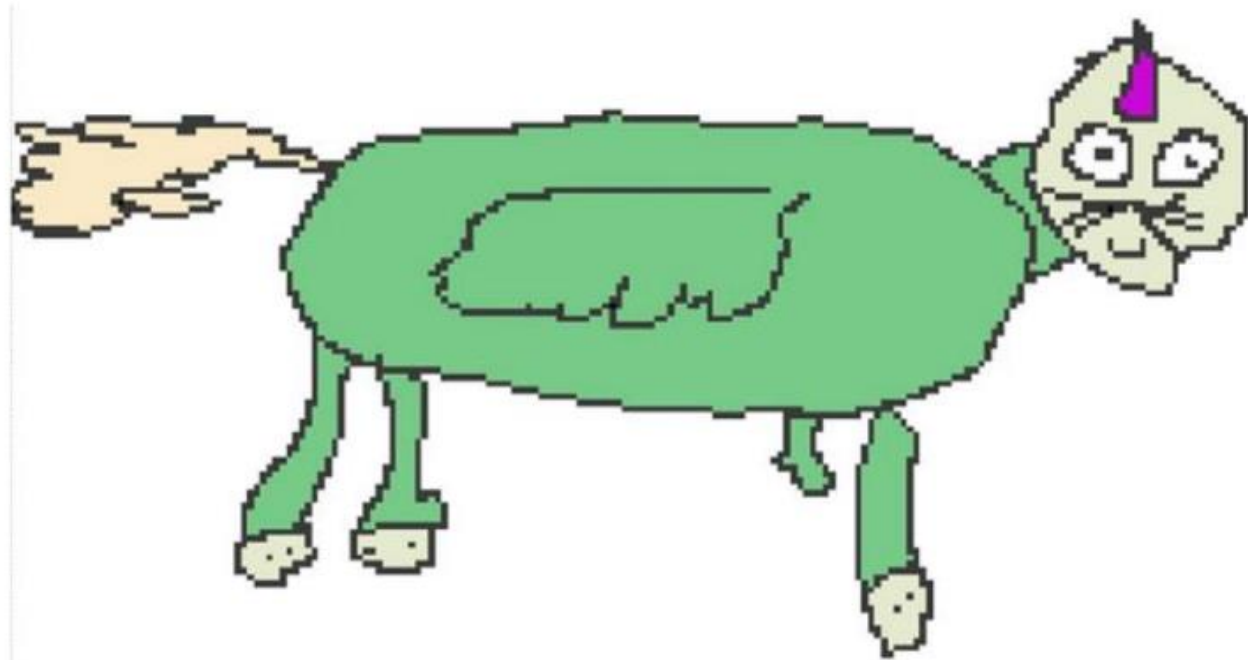
角

眼睛

4条腿

翅膀

尾巴



模式识别

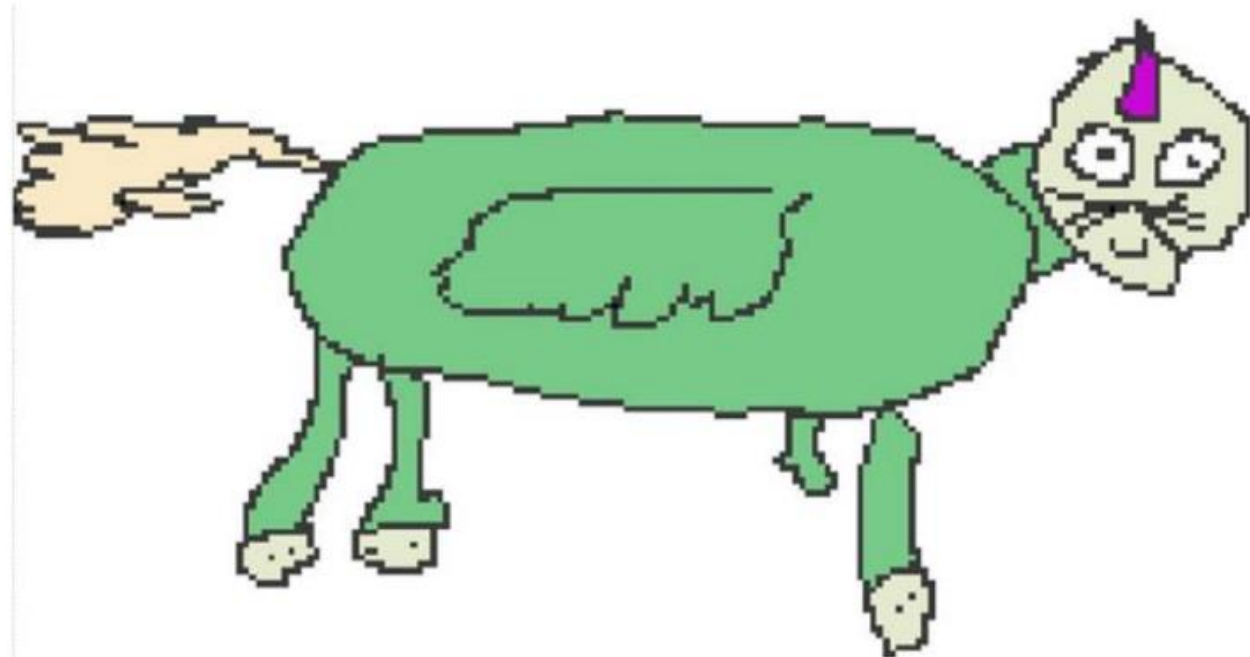
Artificial intelligence connotation

问题：这是一只猫嘛？

确定特征

你并不确定是不是这些特征，所以你要给这些特征分配概率。比如“尾巴”可能不是尾巴，而是一朵云，它是尾巴的概率大概是50%。

角	83%
眼睛	95%
4条腿	66%
翅膀	72%
尾巴	50%



模式识别

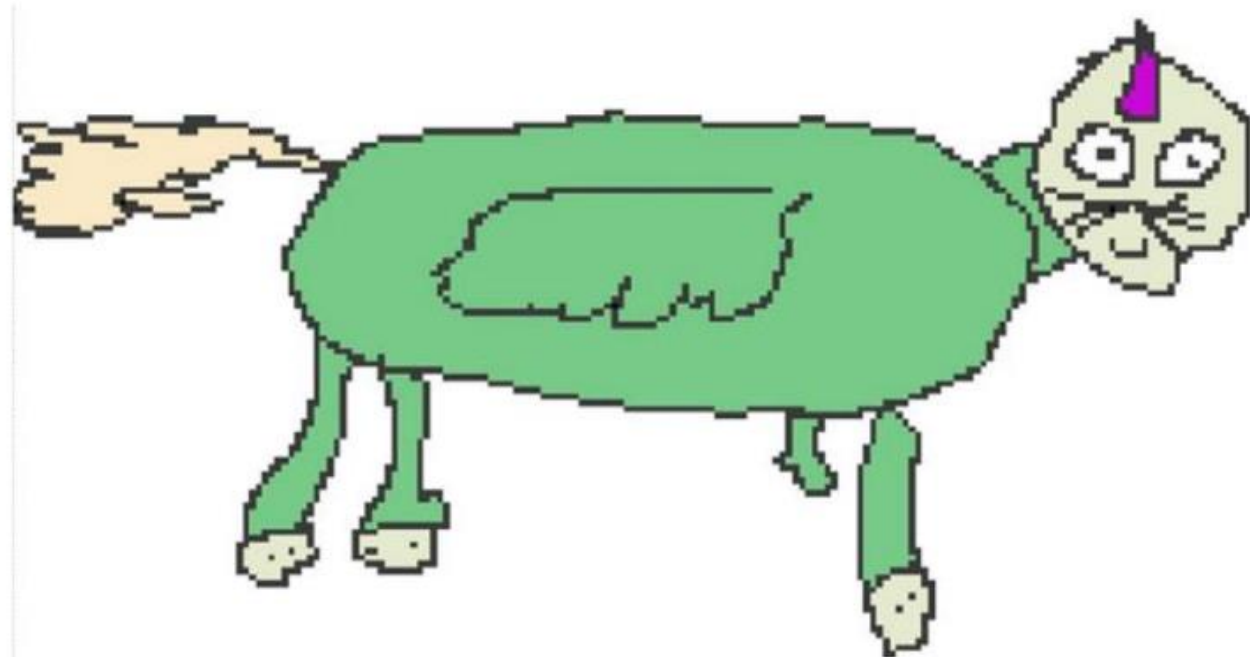
Artificial intelligence connotation

问题：这是一只猫嘛？

分配权重

以上这些特征能够验证“这是一只猫”的假设了吗？这些特征同样重要吗？比如说，“角”这一特征，猫有吗？没有，那这一特征不重要！“2个眼睛”重要吗？一般重要，猫的一只眼睛瞎了也是猫；“4条腿”重要吗？相对重要，翅膀呢？不重要；尾巴呢？重要。这就需要给予每个特征分配权重，来验证假设。

角	83%	-87%
眼睛	95%	42%
4条腿	66%	69%
翅膀	72%	-64%
尾巴	50%	35%



$$Y = 0.83 \times (-0.87) + 0.95 \times 0.42 + 0.66 \times 0.69 + 0.72 \times (-0.64) + 0.5 \times 0.35 = -0.1535$$

笑脸识别·自动拍照

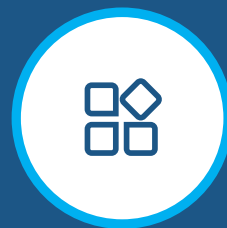
假如把嘴角上扬，牙齿微微露出，以及脸上的某块肌肉在动，定义成笑脸的特征。



模式识别

Moshishibie

我告诉机器具有这些特征的脸就是笑脸，否则不是笑脸，那么机器就对下一张脸进行识别，检查下一张脸有没有那些特征，有的话就识别为笑脸，拍照，没有的话就不拍照。



机器学习

Jiqixuexi

我没有告诉机器笑脸的特征，只是给了机器很多，笑脸与非笑脸的照片，并告诉机器哪张是笑脸（有监督学习），哪张不是笑脸（注意：这里并没有告诉机器笑脸的特征，仅仅告诉机器是不是笑脸），然后机器自己去挖掘笑脸和非笑脸的特征，新的照片来了之后，机器可以判断是不是笑脸，决定拍不拍照。

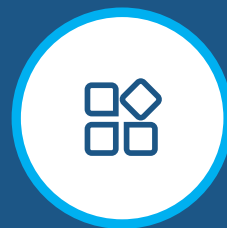
人脸识别·开门



模式识别

Moshishibie

我先告诉机器，A同学脸的特征是大眼睛，长睫毛，樱桃嘴巴等等，当扫描人脸的时候，机器开始检查，这张脸有没有那些特征，有的话，就开门，没的话就不开。



机器学习

Jiqixuexi

我给了机器很多照片，告诉，哪些是A同学的脸，哪些不是（supervised learning），机器就自己去找特征了，（这里的特征只有机器自己知道），再来一张脸，机器就一句自己找到的特征去判断是不是A同学了。

模式识别就是明确告诉了机器某些特征，而机器学习就是让机器自己去找特征，再去判断。

应用领域

Application field



语言识别

Language recognition

应用领域

Application field

该领域其实与机器人领域有交叉，设计的应用是把语言和声音转换成可进行处理的信息：与语音开锁、语音邮件以及未来的计算机输入等方面。



对输入的信息进行识别、转码、解构、分析、集成理解，再从知识库调去输出信息



过程类似，但无需表达完整或精确，也可依据语气、表情等判断意思。



OCR



应用领域

Application field



图像识别

Image recognition

利用计算机进行图像处理、分析和理解，以识别各种不同模式的目标和对象的技术，例如人脸识别、汽车牌号识别、换脸等。



原理



Deepfakes

最早的AI换脸算法

应用领域

Application field

人工智能的应用的未来发展也一定是集成的，就是把不用方向的AI算法集成，综合解决实际问题。

AI机器人

Artificial intelligence robot

它能够理解人的语言，用人类语言进行对话，并能够用特定传感器采集分析出现的情况调整自己的动作来达到特定的目的。

What can I help you with?



优势

理解输入信息、回应操作要求、情感互动

不足

情感互动浅、上下文理解不足...

索菲亚——
一位“机器人公民”



分支方向

Branch field



数据挖掘



数据挖掘又称数据库中知识发现（概念、规则、模式，规律和约束），是指从数据库的大量数据中揭示出隐含的、有潜在价值的信息的非平凡过程。

数据挖掘的特征：

大数据支持

从大量的结构化、非结构化数据中提取价值信息

数据挖掘不是一种用来证明假说的方法，而是用来构建各种各样的假说的方法

构建假说

结果是A和B有关系，但不知道什么关系（不证明假说）

支持学习

数据挖掘本质上是机器学习的基础

分支方向

Branch field

数据挖掘

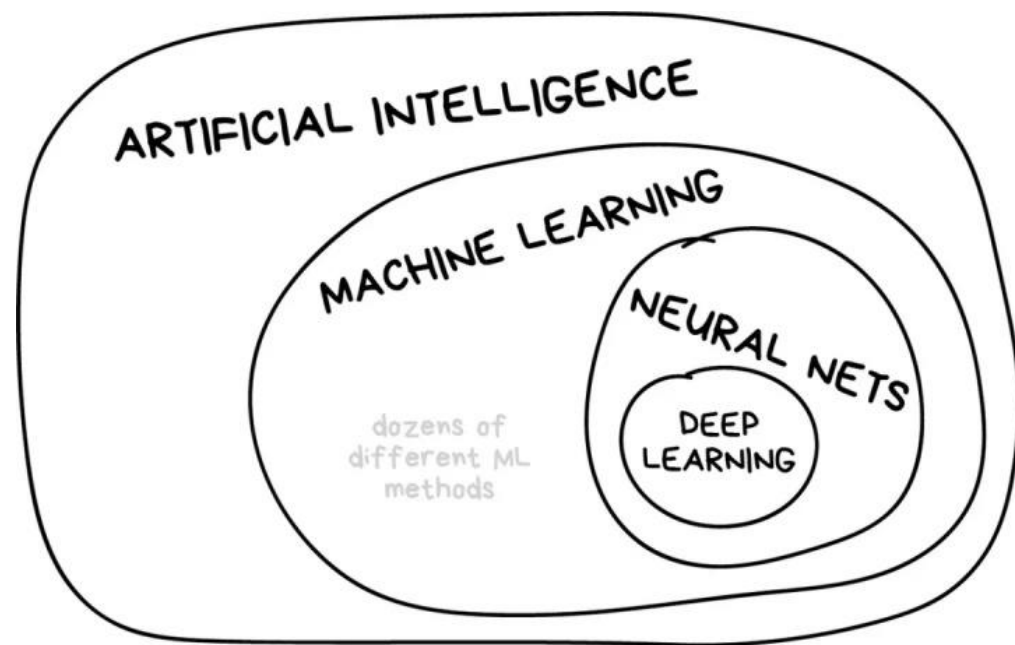
概念宽泛，大数据分析是，Excel统计也算。

机器学习

基本目标是学习一个 $x \rightarrow y$ 的函数（映射），来做分类或者回归的工作。之所以经常和数据挖掘合在一起讲是因为现在好多数据挖掘的工作是通过机器学习提供的算法工具实现的，

深度学习

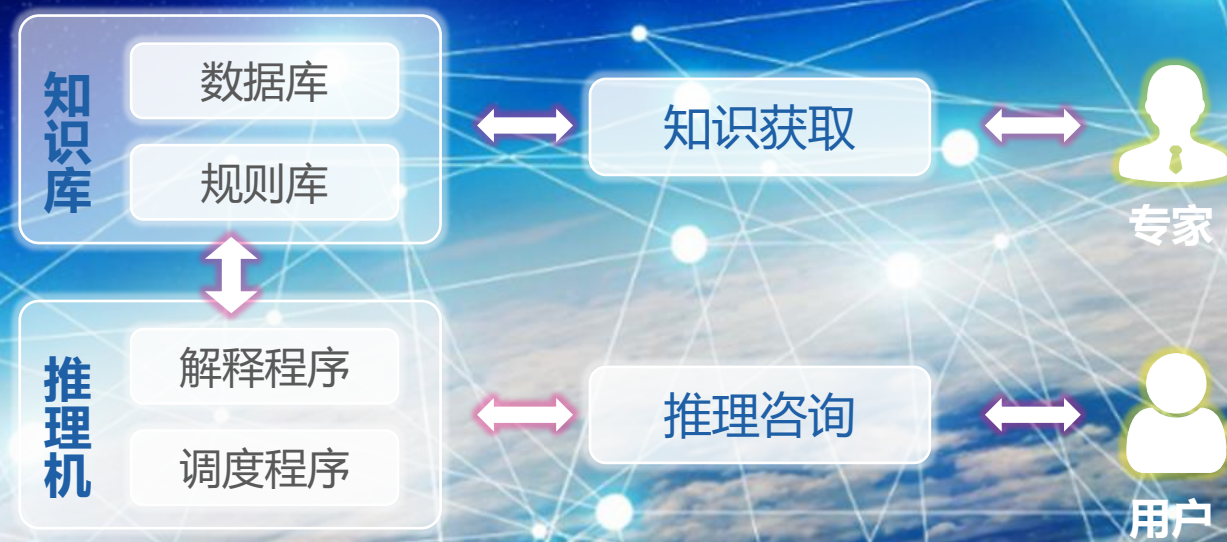
本身是神经网络算法的衍生



数据挖掘是个很宽泛的概念，数据挖掘常用方法大多来自于机器学习这门学科，深度学习是机器学习一类比较火的算法，本质上还是原来的神经网络。

应用领域

Application field



专家系统 Expert system

具有专门知识和经验的计算机智能程序系统，后台采用的数据库相当于人脑，具有丰富的知识储备，采用数据库中的知识数据和知识推理技术来模拟专家解决复杂的问题。



传统程序 = 数据结构+算法
专家系统 = 知识+推理

智慧医疗

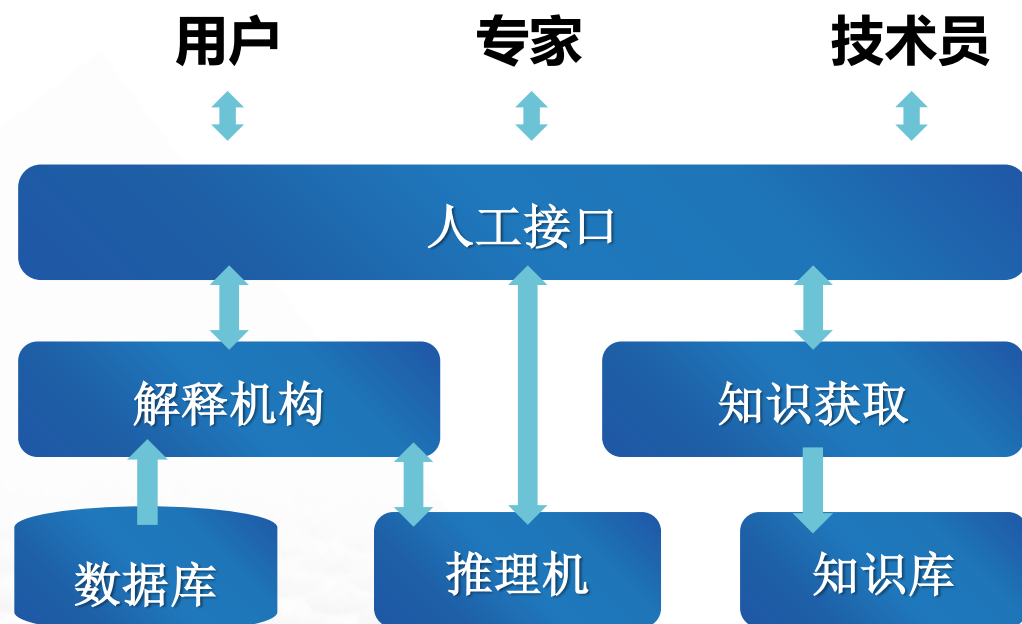
Intelligent medical treatment



医学专家系统

Medical Expert Systems

将医学专家的理论知识、经验教训总结等输入计算机系统，并运用人工智能的方法进行归类、分析、整合。而后运用这些知识模拟人类医学专家或医生的诊断和治疗过程，对具体病例进行分析、推理。最终做出诊断并给出切实可行的治疗建议，以减轻医生工作量或给医生的日常诊疗工作提供辅助性意见。



医学专家系统

Medical Expert Systems

之所以有医学专家系统是因为医学是科学，是能够最简单的逻辑判断来推理的。比如：

根据血常规中白细胞、中性粒细胞、淋巴细胞以及单核细胞的变化，可以判断患者的病情属于病毒还是细菌感染。

- 如果血常规显示中性粒细胞或白细胞计数升高，一般考虑细菌感染的可能性较大,临床上多建议使用头孢类、喹诺酮类或青霉素类的广谱抗生素进行治疗。
- 如果单核细胞或淋巴细胞计数升高，多提示身体内存在病毒感染。治疗方面需要使用抗病毒和清热解毒的中成药进行治疗。

检验项目		结果	提示	单位	参考区间	检验方法
检验组合：血常规						
1	白细胞计数	WBC	8.4	10 ⁹ /L	3.5--9.5	
2	中性粒细胞比值	NE%	51.5	%	50--70	
3	淋巴细胞比值	LY%	31.7	%	20--40	
4	单核细胞比值	MO%	10.0	↑ %	3--8	
5	嗜酸性粒细胞比值	EO%	6.4	↑ %	0.5--5.0	
6	嗜碱性粒细胞比值	BA%	0.4	%	0--1	
7	中性粒细胞绝对计数	NE#	4.3	10 ⁹ /L	1.8--6.3	
8	淋巴细胞绝对计数	LY#	2.7	10 ⁹ /L	1.1--3.2	
9	单核细胞绝对计数	MO#	0.8	↑ 10 ⁹ /L	0.1--0.6	
10	嗜酸粒细胞绝对计数	EO#	0.50	10 ⁹ /L	0.02--0.52	
11	嗜碱粒细胞绝对计数	BA#	0.00	10 ⁹ /L	0--0.06	
12	红细胞计数	RBC	7.23	↑ 10 ¹² /L	4.3--5.8	
13	血红蛋白浓度	HGB	152	g/L	130--175	
14	红细胞压积	HCT	0.490		0.4--0.5	
15	平均红细胞体积	MCV	67.7	↓ fL	82--100	
16	红细胞分布宽度标准差	RDW-SD	36.3	↓ fL	37.0--54.0	
17	红细胞分布宽度	RDW-CV	15.3	↑ %	10--15	
18	平均RBC血红蛋白含量	MCH	21.1	↓ Pg	27--34	
19	平均RBC血红蛋白浓度	MCHC	311	↓ g/L	316--354	
20	有核红细胞比值	NRBC%	0.10	%	0--0.3	
21	有核红细胞绝对计数	NRBC#	0.01			
22	血小板计数	PLT	276	10 ⁹ /L	125--350	
23	血小板压积	PCT	0.239	%	0.19--0.282	
24	平均血小板体积	MPV	8.6	fL	7.5--11	
25	血小板分布宽度	PDW	16.8	fL	11--17.1	

医学专家系统

Medical Expert Systems

MYCIN

MYCIN系统是一种用于医学诊治、治疗感染性疾病的计算机专家咨询系统，它能够帮助医生对住院的血液感染患者进行诊治并选用抗生素类药物进行治疗。

MYCIN系统通过如下四个过程来选择抗菌药物：

1

初步诊断

判断患者是否受细菌感染是一个复杂问题，往往要考虑其他临床因素（体液、血液…），才能最后断定患者是否可能具有传染病。

2

病菌分类

根据细菌的染色特性和形态等，并结合病史及临床情况对细菌的可能类别作出判断。

3

判断作用

通常医生只知道哪些药物对抑制某些细菌可能有效，所以当病菌初步确定之后，还要判断哪些抗生素对该细菌起作用。

4

最佳处方

初步选定药物之后，再结合患者的具体情况，如对某抗生素是否过敏，性别及肾功能是否影响药物的选择，最后选定最佳处方。



人工智能在放射科领域的应用

放射科的主要业务：拍照、分析、初步诊断。每个环节都需要人的参与才能够完成。技术是提高效率的主要手段，人工智能技术有没有可能取代人在放射科中的作用？



智慧医疗

Intelligent medical treatment

图像识别

人工智能辅助影像诊断包括图像识别和深度学习两个应用分支

深度学习



按不同症状地图像特征输入给AI得到不同症状地图像模型，然后拿新的医学图像输入到模型中，看是否能匹配上之前学习的特征，由此得出诊断。

从大量的医学图像中找到病状特征，自动判断是否存在病状，得出诊断。

放射科领域成为医疗人工智能进军的突出领域有如下优势在于：

- | | | |
|---|----------------|---|
| 1 | 数据获取更容易 | 医院放射科一般会存储患者历次影像学资料(CT、MRI、MRA、DSA、PET-CT 等)，具有调取容易、费时少和数据量大等特点。 |
| 2 | 数据处理难度小 | 一份完整的住院病历至少要包括多个方面，而在放射科领域的数据仅仅是一张CT片或者核磁共振片 |
| 3 | 数据更重要 | 放射科影像信息可以在客观上直观反映患者病情，是临床医生做出诊断和临床决策的重要且较直接的依据。 |

临床决策支持系统

Clinical decision support system

“

是指通过构建的医学知识库、模型库、数据库和数据仓库为基础，利用数据挖掘技术和联机分析技术，对临床数据进行综合分析处理和挖掘，并输出决策结果的决策支持系统，辅助治疗。

It refers to the decision support system which uses data mining technology and online analysis technology to comprehensively analyze, process and mine clinical data and output decision results on the basis of medical knowledge base, model base (Model Base), database and data warehouse.

— CDSS —

临床决策支持系统

Clinical decision support system

基于知识库

大多数CDSS由三部分组成，即知识库、推理机和人机交流接口部分，知识部分依赖包括编译信息的规则与联系，通常采用IF-THEN规则来存储和管理知识。

例如，某一系统用来研究药物之间单位相互作用，规则是如果服用X药物与Y药物，那么（THEN）服用者需要注意或者警惕什么。如果采用另一种方式，高级使用者可以编辑相关知识库里的规则，从而用于其它新药的研究。推理机部分是知识库的知识与患者信息整合、比较、分析的引擎。人机交流接口则允许将结果显示给使用者，同时也可以作为系统输入。

基于非知识库

基于非知识库的CDSS系统多采用人工智能的形式，这种人工智能能在近年的CDSS研发中被称为机器学习，可以允许计算机从既往经验中或是其他临床资料中获得知识。CDSS中常用的机器学习算法包括逻辑回归、支持向量机、人工神经网络、深度学习等。

人工智能的综合应用：临床支持智能机器人

智慧医疗

Intelligent medical treatment



临床决策支持系统

Clinical decision support system



解决诊断问题

临床决策支持系统可以作为医务人员临床诊断的辅助工具，根据患者症状及其他相关临床资料提出诊断建议。

当遇到复杂、罕见的病例或者医务人员缺乏诊断经验时，系统可以根据患者资料提供可能的辅助诊断建议。



订制治疗方案

临床决策支持系统作为医务人员制订临床治疗方案的辅助工具，根据患者诊断结果从知识库中选出治疗方案。

临床决策支持系统功能

临床决策支持系统

Clinical decision support system



合理开药

临床决策支持系统能够确保医务人员合理用药，通过适当的剂量使用适当长的时间，同时就可能发生的药物不良反应提出警报，减少医疗事故和药物错误，提高整个治疗的质量。



自动评价

临床决策支持系统能够根据治疗方案自动进行成本效益评价和决策提示，为医务管理人员提供关于疾病、人群、医师及成本变化等趋势的信息。

临床决策支持系统功能

临床决策支持系统

Clinical decision support system

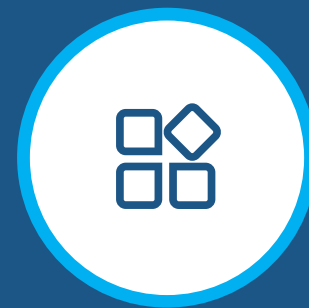
临床决策支持系统依据应用目的可分两类



了解病症

Understand the disease

是关于诊断和预后决策的，通过从医学学科中吸取知识，目的是减少患者当前或将来的状况的不确定性



制订方案

Make a plan

依据诊断结果理性科学的方式提出治疗方案：做什么检查...用什么方式告诉患者。

临床决策支持系统

Clinical decision support system

系统根据工作方式可以分为被动系统、半自动系统和自动系统三类



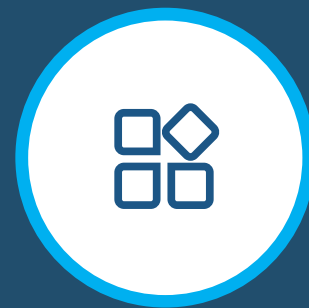
被动系统

Passive system



半自动系统

Semi-automatic system



自动系统

Automatic system

临床决策支持系统

Clinical decision support system

系统根据工作方式可以分为被动系统、半自动系统和自动系统三类



被动系统

Passive system

多数决策支持系统是被动型的，医生必须对系统明确提出问题，描述患者的情况，然后等待系统的建议。根据系统所提供的信息和用户要求，被动系统还可以进一步分成两类。

咨询系统

用户提供患者状况信息，输入数据，依据信息提供诊疗建议

评议系统

用户提供治疗方案，系统对方案作出评价提出意见。

临床决策支持系统

Clinical decision support system

系统根据工作方式可以分为被动系统、半自动系统和自动系统三类



半自动系统

Semi-automatic system

不需要临床医生发出请求，系统根据现有的患者数据自动激活并提供建议。该系统一般是与电子病历系统整合的临床决策支持系统，可以进一步划分为两类：

自动提示

该系统可监视医务人员的活动，帮助医务人员避免重复检查和处方错误。

报警系统

该系统可监视患者状态信号的变化，可以提示医生异常值或异常的变动

临床决策支持系统

Clinical decision support system

系统根据工作方式可以分为被动系统、半自动系统和自动系统三类



自动系统

Automatic system

能够在不需要通过医生干预下自动决策。目前，自动系统广泛应用于麻醉、呼吸机控制和给药系统等领域，它通过借助医疗检测仪，定时测定某些参数，决定是否要实施当前治疗。但导致参数异常的原因除了患者病情变化外，设备或传感器失常也会导致参数异常。

智慧医疗

Intelligent medical treatment





AI局限性

Limitations of AI

- 人工智能成为近年来最受关注的高新技术产业，被许多国家提到了战略发展的高度，国内外的IT巨头也纷纷在此领域布局。
- IT巨头在人工智能领域的布局

IT巨头	产业布局	AI技术
百度	汽车领域无人驾驶、基于智能手机的APP应用与插件	语音识别、视觉识别、自动驾驶、深度学习
阿里巴巴	智能家具、物联网	云平台、可视化人工智能平台
腾讯	智能硬件	视觉、智能计算与搜索实验室
IBM	认知商业计算平台、医疗领域	IBM Watson（超级计算机系统）
Facebook	语音助手、VR生态	视觉DeepFace技术、语音识别、自然语义、神经网络训练+机器学习
Apple	汽车领域无人驾驶、SIRI语音助手	自然语言、可视化地图
Microsoft	语言助手（微软小冰、Tay、Cortana小娜）、VR	语音、视觉、自然语言、分布式机器学习
Google	无人驾驶、智能家居、VR、图像识别	TensorFlow数据库



《医学信息学》

本课至此·谢谢

医学信息学是以信息学、信息管理和信息技术为依托，研究医学领域中的信息现象和信息规律，用于医学决策和管理的一门交叉学科。学科以生命信息现象和人类社会活动中的医学信息为研究对象。

谭积斌