

# 第一讲：神经网络绪论

南京大学人工智能学院

申富饶

# 目录

CONTENTS

01. 什么是神经网络?
02. 为什么学习神经网络?
03. 神经网络研究的目标
04. 神经网络发展的历史
05. 神经网络的研究者在做什么研究
06. 神经网络的研究方法

01

# 什么是神经网络？

重点：一种模仿生物神经网络的结构和功能的数学模型或计算模型

# 神经网络的定义

人工神经网络（Artificial Neural Network, ANN），简称神经网络（Neural Network, NN），在机器学习和认知科学领域，是一种**模仿生物神经网络**（动物的中枢神经系统，特别是大脑）的结构和功能的数学模型或计算模型，用于对函数进行估计或近似。

神经网络由大量的人工神经元联结进行计算。大多数情况下人工神经网络能在外界信息的基础上改变内部结构，是一种**自适应系统**，通俗地讲就是具备**学习功能**。

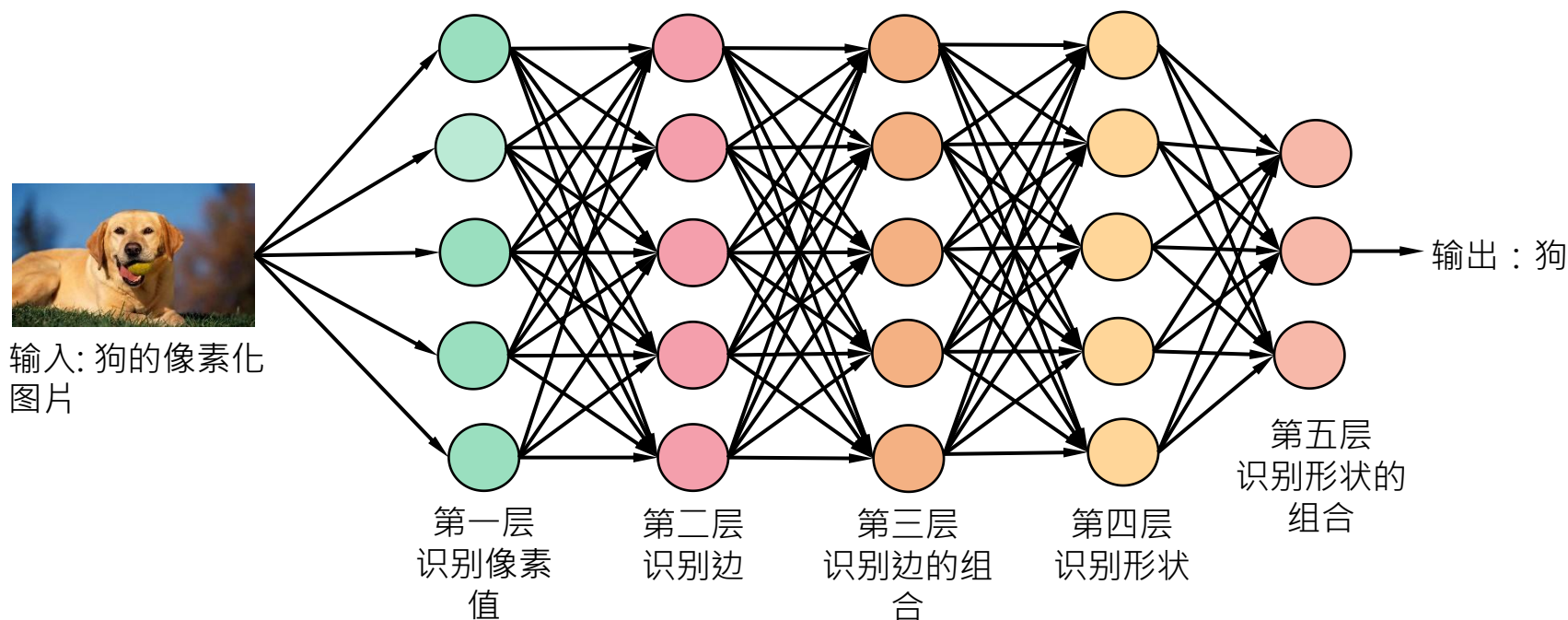
# 神经网络如何工作？

- 神经网络是一个具有相连节点层的计算模型
- 其分层结构与大脑中的神经元网络结构相似。
- 神经网络可**通过数据进行学习**
- 因此，可训练其识别模式、对数据分类和预测未来事件



# 神经网络如何工作？

- 神经网络将输入自底向上地抽象为多个层次的特征
- 神经网络的行为由神经元之间的连接方式以及连接的强度（**权重**）确定。
- 在训练期间，根据指定的学习规则自动调整相关权重，直到满足一定的结束条件为止。



## 02

## 为什么学习神经网络？

重点：神经网络独具特色和魅力，可以通过“学习”解决传统算法解决不了的问题



神经网络已经被用于解决各种各样的问题，例如机器视觉和语音识别。这些问题都是对于传统的机器学习算法来说非常困难的。

语义  
分割



A person riding a motorcycle on a dirt road.

Two dogs play in the grass.



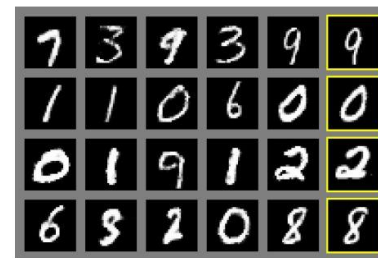
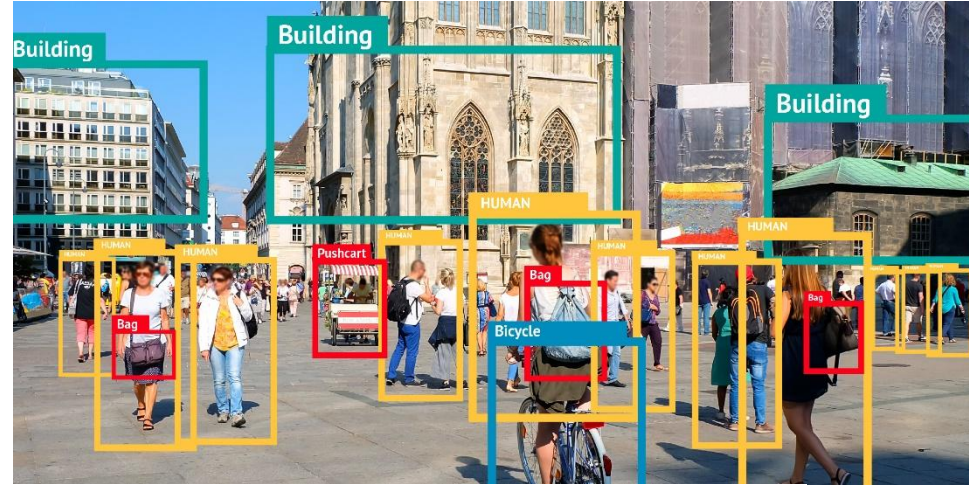
图文生成

A group of young people playing a game of frisbee.

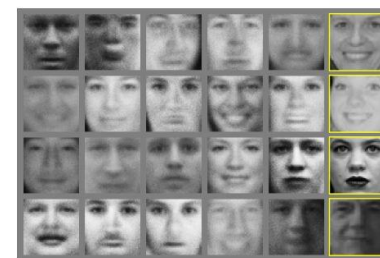
Two hockey players are fighting over the puck.



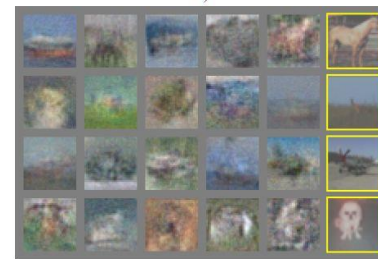
目标  
检测



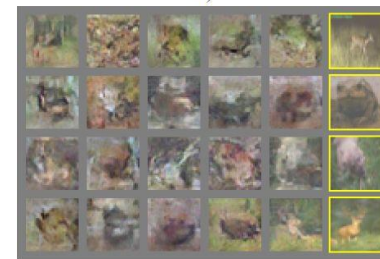
a)



b)



c)



d)

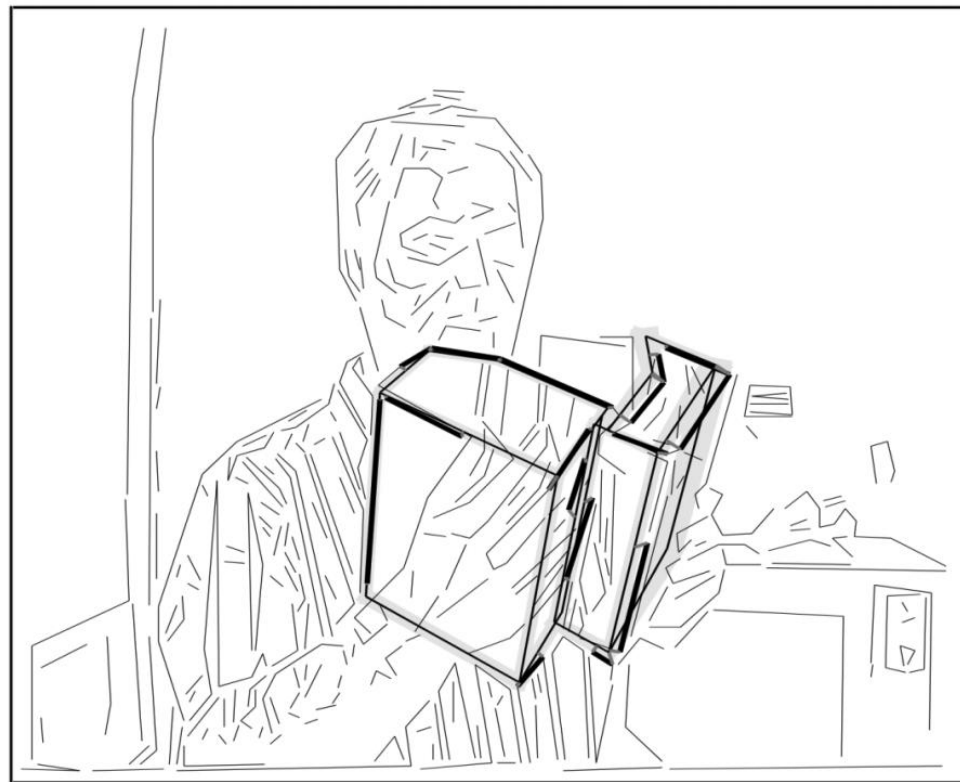
图像生成



# 传统方法如何解决计算机视觉问题？

通过我们对图像的先验知识，对图像中的特征进行提取。这个过程，被称为**特征工程**，常用的特征，有图像的亮度，纹理，对比度等。在提取到相关特征之后，才能对图像中的内容进行判断，做出例如分类，检测等操作。

Edges, segmentation, and perception



D. Lowe. JCV, 1992

# 神经网络如何解决视觉问题？

## 深度学习

- 两层或三层连接的神经网络称为浅层神经网络。而当前主流视觉应用中的深度神经网络可以有数十层，甚至数百层。
- 深度卷积神经网络能够逐层学习从低级到高级的图像特征，而不用人工进行特征工程。

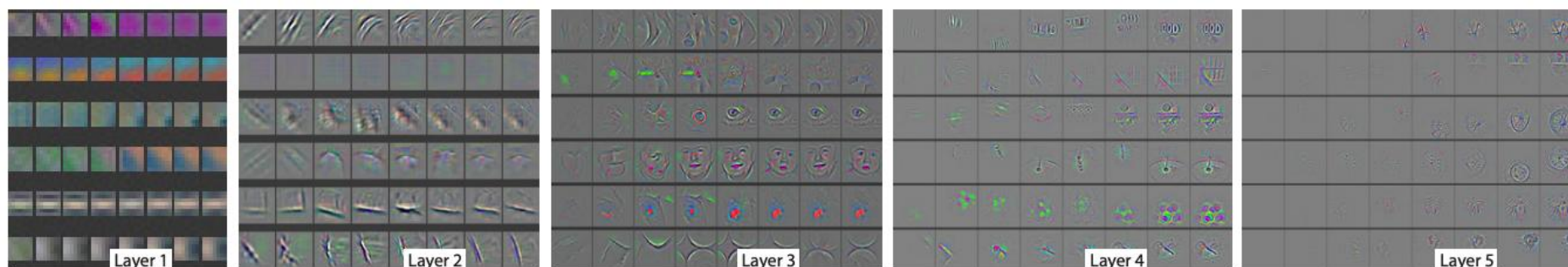


Figure 4. Evolution of a randomly chosen subset of model features through training. Each layer's features are displayed in a different block. Within each block, we show a randomly chosen subset of features at epochs [1,2,5,10,20,30,40,64]. The visualization shows the strongest activation (across all training examples) for a given feature map, projected down to pixel space using our deconvnet approach. Color contrast is artificially enhanced and the figure is best viewed in electronic form.

# 与传统机器学习相比，神经网络的优势

## 特征提取的高效性

- 机器学习算法：要求事先确定特征欠拟合、过拟合。在特征工程上需花大量时间和精力
- 神经网络：不需要做大量的特征工程，可以直接利用数据，让它自己训练，自我“修正”

## 数据格式的简易性

- 传统机器学习分类问题中，需对数据进行如量纲的归一化，格式的转化等，在神经网络里不一定需要对数据做额外处理

## 参数数目的少量性

- SVM：需要调整的参数有核函数、惩罚因子、松弛变量等，需要对背后理论知识的深入了解
- 基本的三层神经网络(输入-隐含-输出)：只需初始化时给每一个神经元随机赋予一个权重 $w$ 和偏置项 $b$ ，对于调参的背后理论知识并不需要过于精通，需要的先验知识更少。

# 神经网络适合什么任务？

编码型神经网络特别适合执行**大规模复杂数据的模式识别**，用以识别语音、视觉和控制系统中的对象或信号并对其分类。它们还可以用于执行时序预测和建模。

## 示例：

- 电力公司准确预测其电网上的负载，以确保可靠性，并优化他们运行的发电机的效率。
- ATM通过读取支票上的账号和存款金额的可靠方式接受银行存款。
- 病理学家依靠癌症检测应用的指导，根据细胞大小的均匀度、肿块密度、有丝分裂及其他因素将肿瘤分类为良性或恶性。

# 神经网络适合什么任务？

解码型神经网络特别适合**生成式智能**任务，用以产生以假乱真的人工数据，涵盖语言、图像、视频、音频等数据类型。

示例：

- 基于神经网络语言模型的机器翻译。
- 基于生成模型的超分辨率重建，用于实现低清影像的高清化。
- AI绘图、作曲、视频生成系统。
- ChatGPT以及诸多AI公司跟进研发的大语言模型。



## 03

## 神经网络研究的目标

重点：借助神经科学、脑科学与认知科学的研究成果，研究大脑信息表征、转换机理和学习规则，建立模拟大脑信息处理过程的智能计算模型，最终使机器掌握人类的认知规律

# 为什么研究神经网络？

总的来说，发展人工神经网络或模型有两个目的：

首先，是为了弄清生物神经系统的运作机制；

其次，是在此基础上制造具有类似功能的信息处理系统。

- **科学目标**是理解智能，即从原理上研究神经网络具有智能的原因，了解智能的形成机理；
- **工程目标**要求我们对智能建立相应的计算模型，将智能移植到机器上，从而在工程上实现智能。

# 如何模拟智能？

**Norbert Wiener: 就其控制行为而言，所有的技术系统都模拟生物系统，然而，没有任何一种生物系统是模拟技术系统的**

**模拟人的智能行为，是技术系统或人工系统模拟生物系统的最高形式**

**使机器具有智能，是技术系统或人工系统模拟生物系统的最高目标**

# 脑功能和神经网络

## 脑：

- 接受信息→处理信息→输出信息
- 控制各器官、融合信息、记忆、联想、识别、推理、决策等
- 能产生意识

## 人工神经网络：

- 接受信息→处理信息→输出信息
- 控制、信息融合、联想记忆、模式识别、专家系统等
- 能否产生意识？

**对脑的研究可以促进人工神经网络的研究**

**对人工神经网络的研究，也可以反过来促进对大脑的研究**

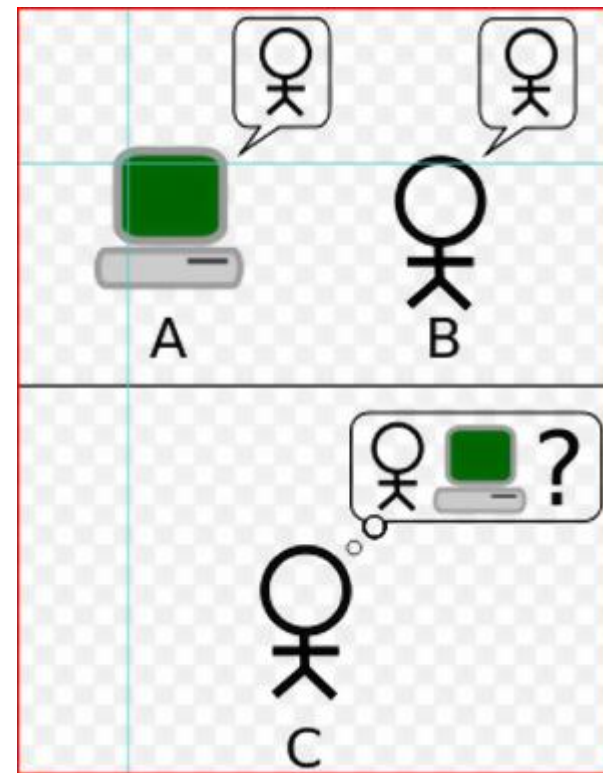
# 人工神经网络能产生意识吗？

## 什么是意识？

- 意识是人脑对大脑内外表象的觉察
- 人的头脑对于客观物质世界的反映，也是感觉、思维等各种心理过程的总和

## 人工神经网络

- 能通过网络结构与权重计算来自外部输入信息的特征表示
- 基于特征表示完成识别、推理等智能任务



人工神经网络在广义上已经具有了某种“意识”



## 04

## 神经网络发展的历史

- 历史大事件概述
- 重要人物、团队
- 重要会议、期刊

# 神经网络的发展历史

前言：深度学习 “海啸”

Part1:起源(1950s-1980s)

Part2:神经网络的兴盛(1980s-2000s)

Part3:深度学习(2000s-2020s)

Part4:生成式智能(2020s)

# 前言：深度学习“海啸”

“Deep Learning waves have lapped at the shores of computational linguistics for several years now, but 2015 seems like the year when the full force of the tsunami hit the major Natural Language Processing (NLP) conferences.”

“在过去的几年中，深度学习的浪潮也曾几次冲击计算机语言学的海岸，但在2015年，深度学习乘海啸之势涌入自然语言学习（NLP）的会议。”

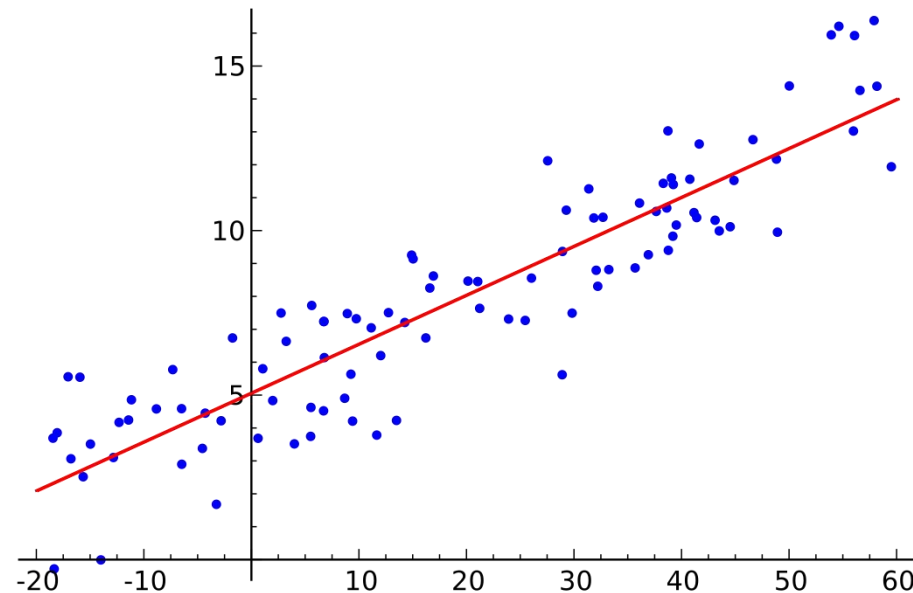
-Dr.ChristopherD.Manning, Dec2015

**不仅在NLP领域，在之前提到计算机视觉（CV）等其他人工智能领域，深度学习，这个神经网络中的佼佼者，都是科研和应用领域的重中之重。**

# Part1:起源(1950s-1980s)

## 传统机器学习方法的年代

- 从传统线性回归过度到监督学习
- 从直接求解合适的平面去拟合点，到利用训练集( $X$ 是输入， $Y$ 是输出)和测试集去“学习”一个平面来拟合点。



# Part1:起源(1950s-1980s)

## 感知机年代的出现

- 1943年，神经元的MP模型在论文《神经活动中所蕴含思想的逻辑活动》中被首次提出，创建该模型的是来自美国的心理学家McCulloch以及另一位数学家Pitts。继而得名MP神经元。
- MP神经元模型，利用数学模型简化了生物学神经元，但是在现阶段，他并不具有“学习”的功能

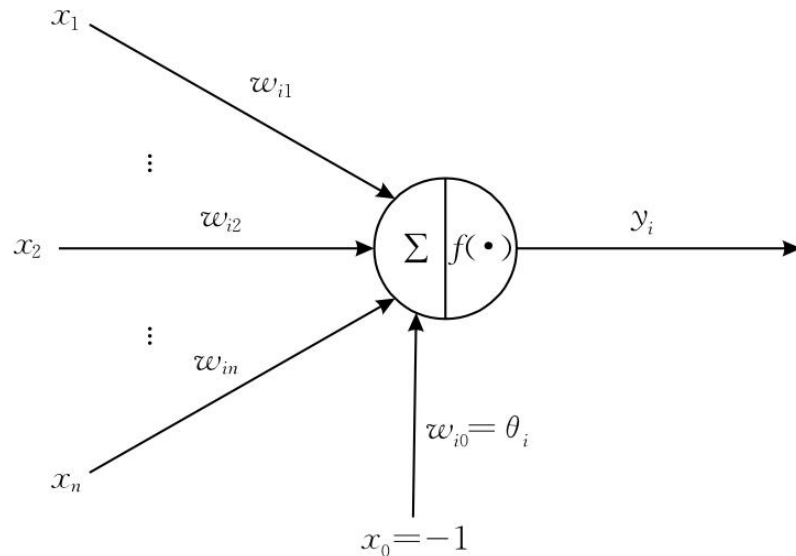


图 1 神经元的 M-P 模型示意

图中,  $x_i (i=1, 2, \dots, n)$  表示来自与当前神经元相连的其他神经元传递的输入信号,  $w_{ij}$  代表从神经元  $j$  到神经元  $i$  的连接强度或权值,  $\theta_i$  为神经元的激活阈值或偏置,  $f$  称作激活函数或转移函数. 神经元的输出  $y_i$  可以表示为如下形式:

$$y_i = f\left(\sum_{j=1}^n w_{ij} x_j - \theta_i\right) \quad (1)$$



# Part1:起源(1950s-1980s)

- 1949年, 在《行为的组织》一书中心理学家 Donald Hebb 对神经元之间连接强度的变化规则进行了分析, 并基于此提出了著名的Hebb学习规则
- 受启发于巴甫洛夫的条件反射实验, Hebb认为如果两个神经元在同一时刻被激发, 则它们之间的联系应该被强化
- Hebb规则隶属于无监督学习算法的范畴, 其主要思想是根据两个神经元的激发状态来调整其连接关系, 以此实现对简单神经活动的模拟

$$w_{ij}(t+1) = w_{ij}(t) + \alpha y_j(t) y_i(t) \quad (2)$$

其中,  $w_{ij}(t+1)$  和  $w_{ij}(t)$  分别表示在  $t+1$  和  $t$  时刻时, 神经元  $j$  到神经元  $i$  之间的连接强度, 而  $y_i$  和  $y_j$  则为神经元  $i$  和  $j$  的输出.

# Part1:起源(1950s-1980s)

- 继Hebb学习规则之后，神经元的有监督Delta学习规则被提出，用以解决在输入输出已知的情况下神经元权值的学习问题
- 该算法通过对连接权值进行不断调整以使神经元的实际输出和期望输出到达一致，其学习修正公式如下

$$w_{ij}(t+1) = w_{ij}(t) + \alpha(d_i - y_i)x_j(t) \quad (3)$$

其中， $\alpha$  为算法的学习速率， $d_i$  和  $y_i$  为神经元  $i$  的期望输出和实际输出， $x_j(t)$  表示神经元  $j$  在  $t$  时刻的状态（激活或抑制）。

# Part1:起源(1950s-1980s)

- 1958年, Rosenblatt等人成功研制出了代号为Mark I的感知机(perceptron), 这是历史上首个将神经网络的学习功能用于模式识别的装置, 标志着神经网络进入了新的发展阶段
- 感知机是二分类的线性判别模型, 旨在通过最小化误分类损失函数来优化分类超平面, 从而对新的实例实现准确预测
- 假设输入特征向量空间为 $x \in R^n$ , 输出类标空间为 $y = \{-1, +1\}$ , 则感知机模型如下

$$y = f(x) = \text{sign}(w \cdot x + b) \quad (4)$$

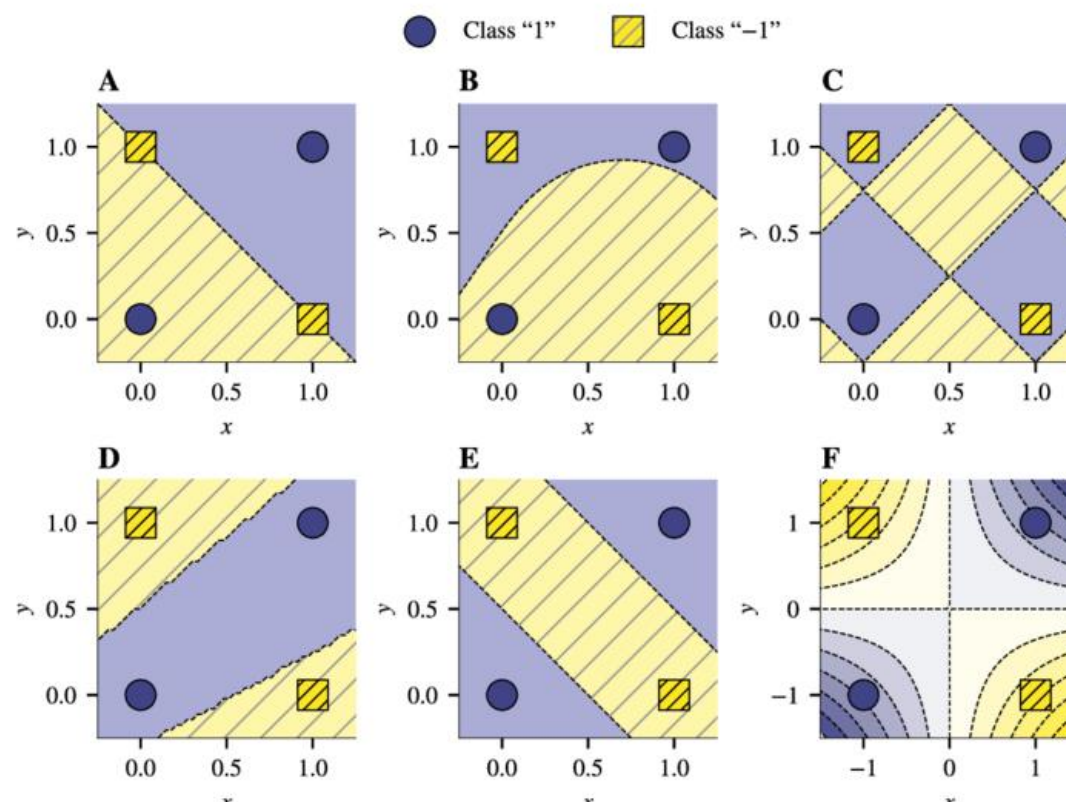
其中,  $w$  和  $b$  为神经元的权值向量和偏置;  $w \cdot x$  表示  $w$  和  $x$  的内积;  $\text{sign}$  为符号函数:

$$\text{sign}(x) = \begin{cases} +1, & x \geq 0 \\ -1, & x < 0 \end{cases} \quad (5)$$

# Part1:起源(1950s-1980s)

- Widrow等人设计的自适应线性元件 Adaline和由Steinbuch等人设计的被称为学习矩阵的二进制联想网络及其硬件实现
- 1969年Minsky和Papert从数学的角度证明了单层神经网络具有有限的功能，甚至在面对简单的“异或”逻辑问题时也无能为力。
  - 神经网络的寒冬 (AI Winter)

parameters. Perhaps surprisingly, even when generalizing this expression to  $\sigma(f(x) + g(y))$ , where  $f, g$  are arbitrary univariate functions, it is not possible to solve the XOR problem (Fig 1A to 1C; see appendix S1 for a proof).



# Part1:起源(1950s-1980s)

- 1982年美国加州理工学院的Hopfield提出了**连续的和离散的Hopfield神经网络模型**，并采用全互联型神经网络尝试对非多项式复杂度的旅行商问题(Traveling Salesman Problem, TSP)进行了求解，促进了神经网络研究再次进入蓬勃发展期。
- Hopfield强调工程实践的重要性，他**利用电阻、电容和运算放大器等元件组成的模拟电路**实现了对网络神经元的描述，把最优化问题的目标函数转换成Hopfield神经网络的能量函数，通过网络能量函数最小化来寻找对应问题的最优解。Hopfield网络是一种循环神经网络，从输出到输入有反馈连接。



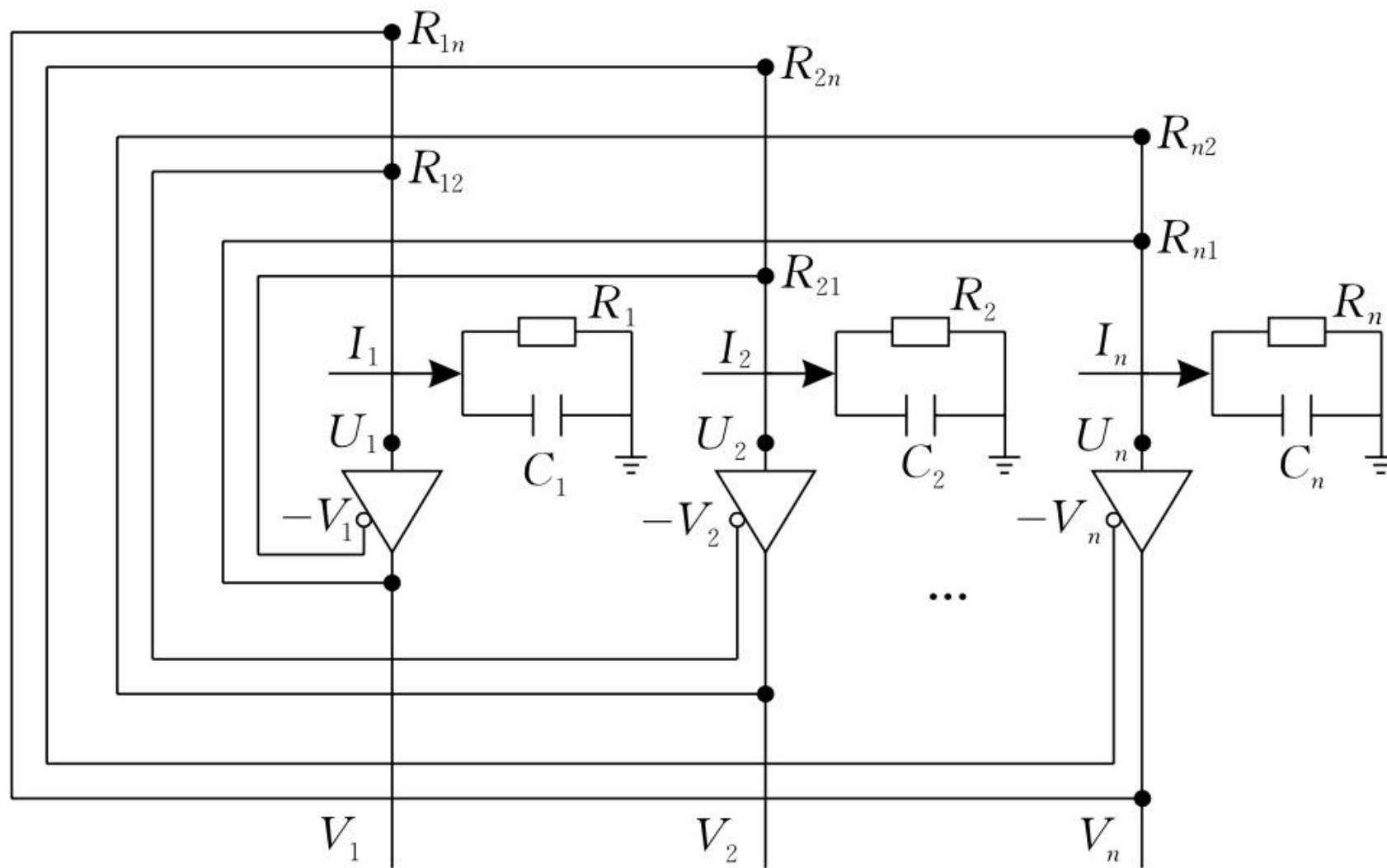
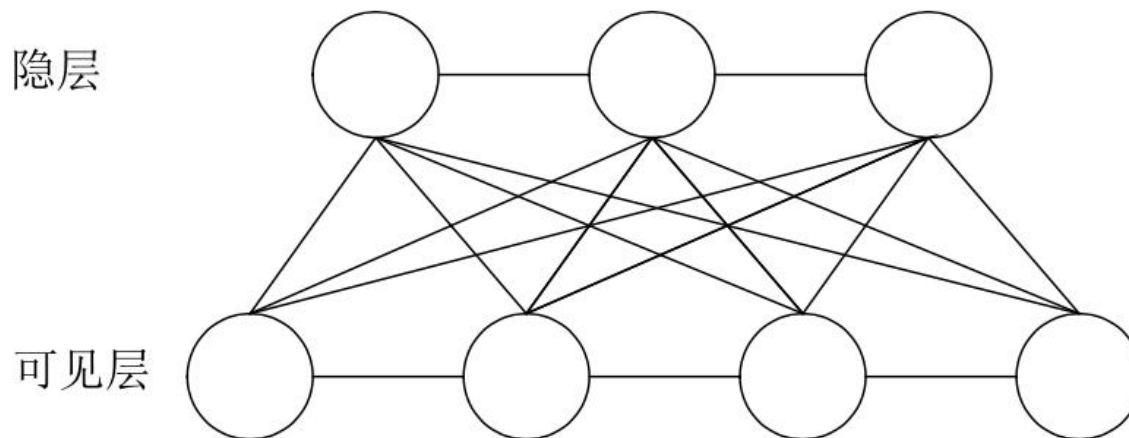


图 2 典型的 Hopfield 网络模型

# Part1:起源(1950s-1980s)

- 1983年, “隐单元”的概念首次被Sejnowski和Hinton提出, 并且他们基于此设计出了玻尔兹曼机(Boltzmann Machine, BM)和受限玻尔兹曼机(Restricted Boltzmann Machine, RBM)。
- 玻尔兹曼机是一种由随机神经元全连接组成的反馈神经网络, 包含一个可见层和一个隐层。网络中神经元的输出只有两种状态(未激活和激活, 用二进制0和1表示), 其取值根据概率统计规则决定。玻尔兹曼机具有较强的无监督学习能力, 可以从数据中学习到的复杂知识规则, 然而也存在着训练时间过长的的问题。



# Part1:起源(1950s-1980s)

- 此外，不仅难以准确计算BM所表示的分布，得到服从BM所表示分布的随机样本也很困难.基于以上原因，对玻尔兹曼机进行了改进，引入了限制玻尔兹曼机(Restricted Boltzmann Machine, RBM)。
- 相比于玻尔兹曼机，RBM的网络结构中层内神经元之间没有连接，尽管RBM所表示的分布仍然无法有效计算，但可以通过Gibbs采样得到服从RBM所表示分布的随机样本。
- Hinton于2002年提出了一个学习的快速算法(对比散度)，只要隐层单元的数目足够多时，RBM就能拟合任意离散分布。
- RBM已被用于解决不同的机器学习问题，比如分类、回归、降维、高维时间序列建模、语音图像特征提取和协同过滤等方面。
- 同时，作为深度学习初期主要框架之一的深度信念网也是以RBM为基本组成单元的。**这一阶段的神经网络已经从起初的单层结构扩展到了双层，隐含层的出现使得网络具有更强的数据表示能力。**

# Part1:起源(1950s-1980s)

- 1974年，Werbos在他的博士论文里提出了用于神经网络学习的BP(Back Propagation)算法，才为多层神经网络的学习训练与实现提供了一种切实可行的解决途径。
- 1986年由Rumelhart和McClland为首的科学家小组对多层网络的误差反向传播算法进行了详尽的分析，进一步推动了BP算法的发展。
- BP网络的拓扑结构包括输入层、隐层和输出层，它能够在事先不知道输入输出具体数学表达式的情况下，通过学习来存储这种复杂的映射关系。其网络中参数的学习通常采用反向传播的策略，借助最速梯度信息来寻找使网络误差最小化的参数组合。

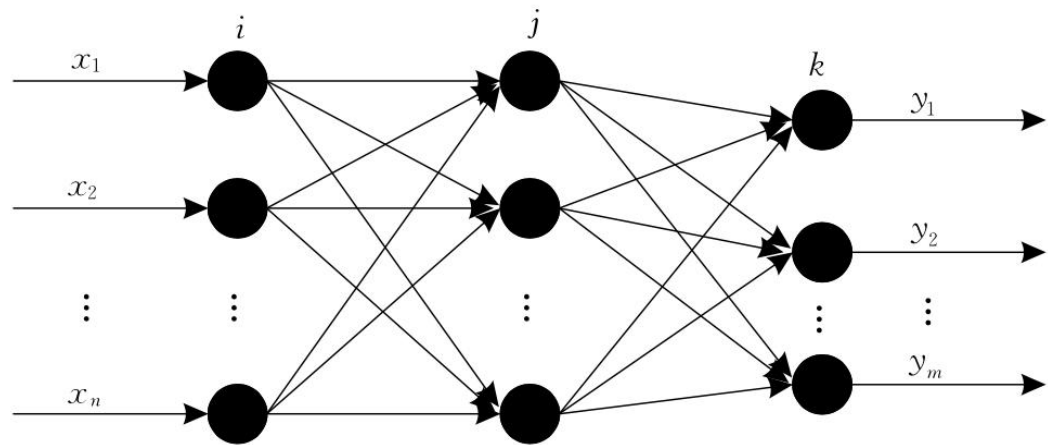


图 4 3 层 BP 网络示意

输入层-隐藏层-输出层

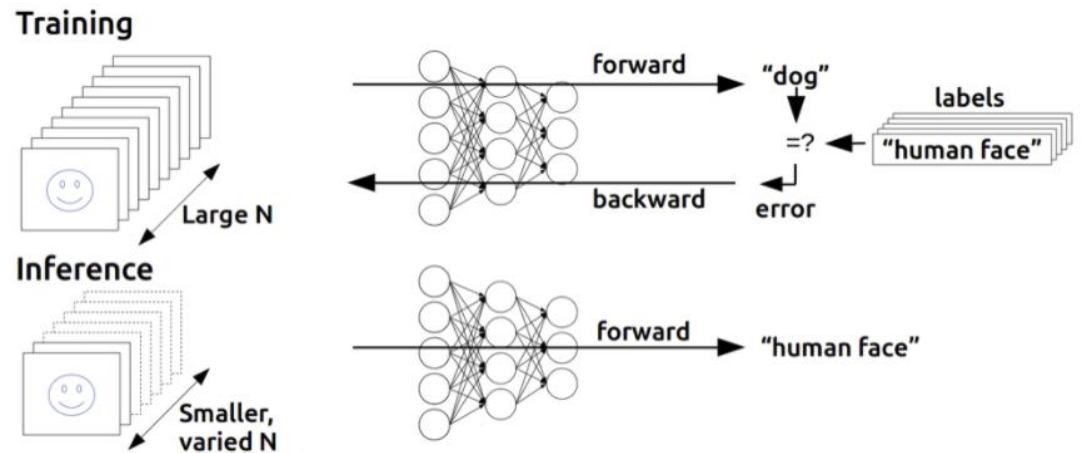


Figure 1: Deep learning training compared to inference. In training, many inputs, often in large batches, are used to train a deep neural network. In inference, the trained network is used to discover information within new inputs that are fed through the network in smaller batches.

训练：通过反传更新参数  
推理：利用训练好的参数得出结论

# Part1:起源(1950s-1980s)

- 此后于1989年，Cybenko、Funahashi、Hornik等人相继对BP神经网络的非线性函数逼近性能进行了分析，并证明了对于具有单隐层、传递函数为sigmoid的连续型前馈神经网络**可以以任意精度逼近任何复杂的连续映射**。
- 继BP之后，为模拟生物神经元的局部响应特性，Broomhead和Lowe于1988年将径向基函数引入到了神经网络的设计中，形成了**径向基神经网络RBF (Radial basis functions)**。后来，Jackson和Park分别于1989年和1991年对RBF在非线性连续函数上的一致逼近性能进行了论证。
- RBF神经网络是一种3层的前向网络，其基本工作原理是：利用RBF构成的隐藏层空间对低维的输入矢量进行投影，将数据变换到高维空间中去，以使原来线性不可分的问题能够变得线性可分。

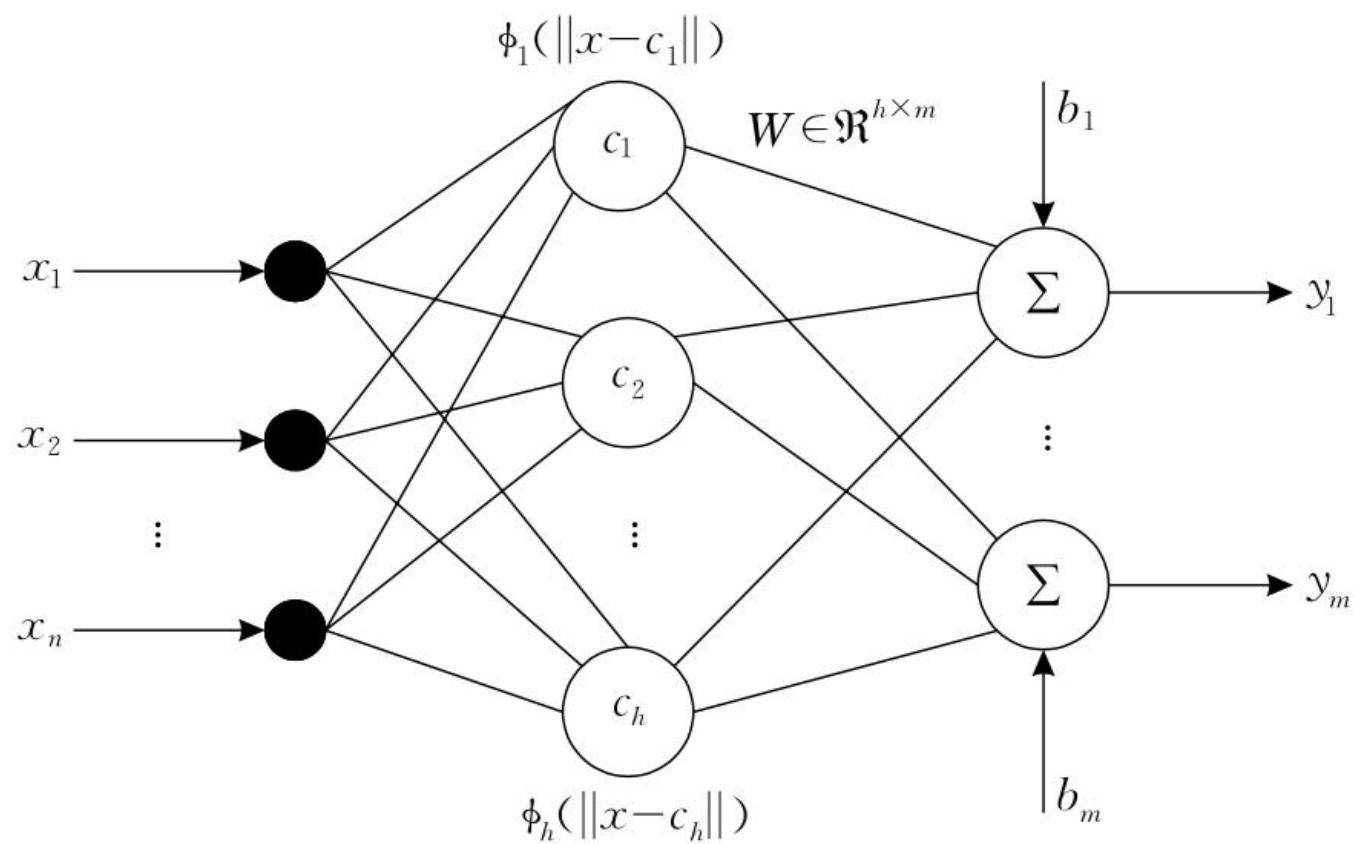
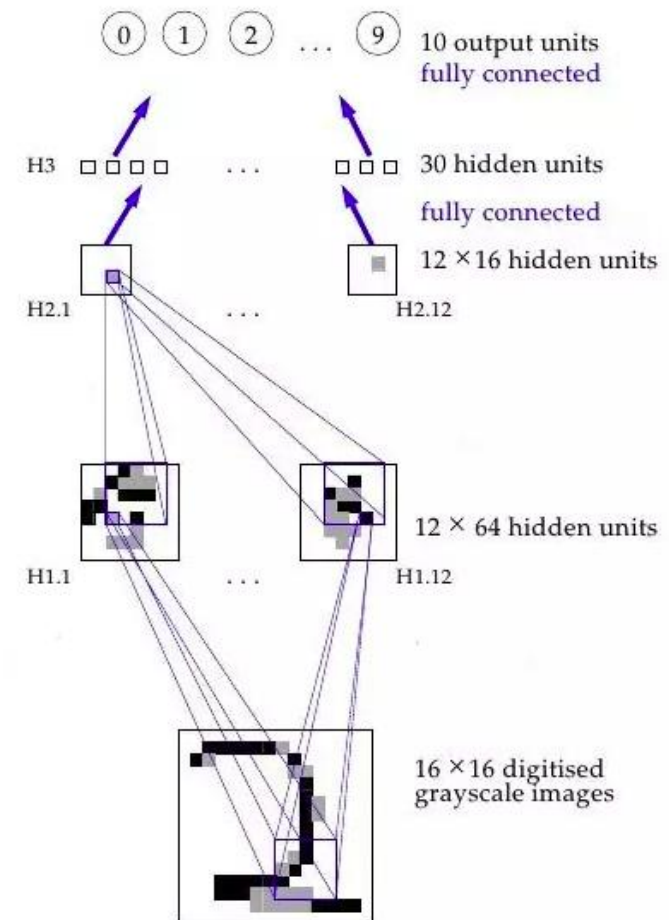


图 5 RBF 神经网络基本结构示意图



# Part2:神经网络的兴盛(1980s-2000s)

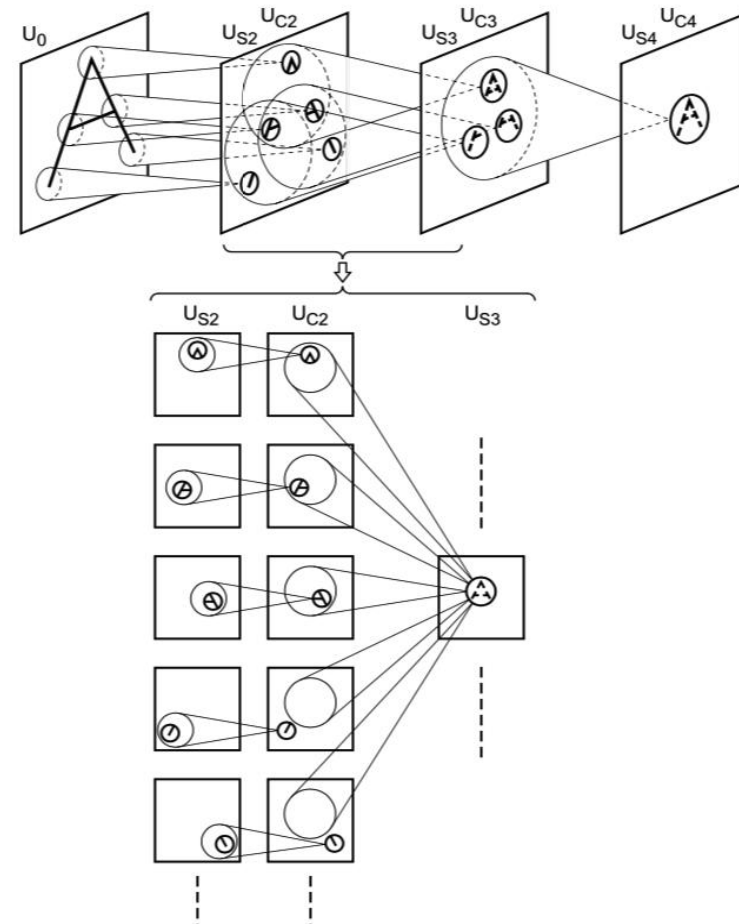
- 1989年在“Multilayer feedforward networks are universal approximators”一文中，作者给出了数学证明，即**多层结构可以使神经网络在理论上拟合任意函数**，当然，也包括异或（XOR）
- 同年，YannLeCun和AT&T Bell Labs中的其他研究者，将理论落在了实际问题中。他们利用多层神经网络和反传成功让**计算机识别了手写邮政编码**，发表了“Backpropagation Applied to Handwritten ZipCode Recognition”一文，奠定了现代计算机视觉的基础。



A visualization of how this neural net works. (Source)

## Part2:神经网络的兴盛(1980s-2000s)

- 在YanLeCun的论文中，除了反传的应用，他们提出对神经网络的改进：**卷积** (convolution)。卷积通过“**权值共享**”大大加速了神经网络的学习过程。
- “权值共享”的概念在1986年时被Rumelhart, Hinton, Williams等人详细论证了。在此之前，福岛邦彦 (Kunihiko Fukushima) 在1980年的“Neocognitron”中，提到过类似的概念。



## Part2:神经网络的兴盛(1980s-2000s)

- 自编码器(Autoencoder)是一种无监督的特征学习网络，它利用反向传播算法，让目标输出值等于输入值，其结构如图所示。对于一个输入  $x \in R^n$ ，首先将其通过一个特征映射得到对应的隐藏层表示  $h \in R^m$ ，隐藏层表示接着被投影到输出层  $y \in R^n$ ，并且希望输出与原始输入尽可能相等。
- 自编码器试图学习一个**恒等函数**，当  $m < n$  时，可以实现对信号的压缩表示，获得对输入数据有意义的特征表示。通常隐层权值矩阵和输出层权值矩阵互为转置，这样大大减少了网络的参数个数。

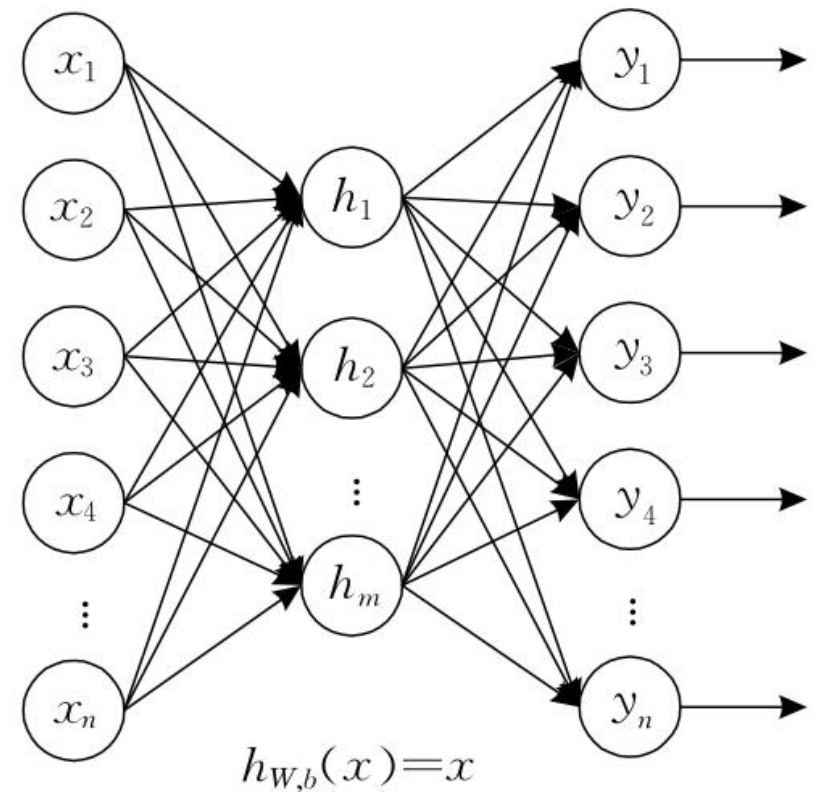


图 6 自编码器示意

## Part2:神经网络的兴盛(1980s-2000s)

- 深度信念网络 (Deep Belief Network, DBN) 是由Hinton在2006年提出, 它是一种生成模型, 通过训练神经元之间的权重, 可以让整个神经网络按照最大概率来生成训练数据
- DBN是由多层RBM堆叠而成的, 神经元可以分为显性神经元和隐性神经元, 显性神经元用于接受输入, 隐性神经元用以提取特征, 最顶上的两层连接是无向的用以组成联想记忆单元

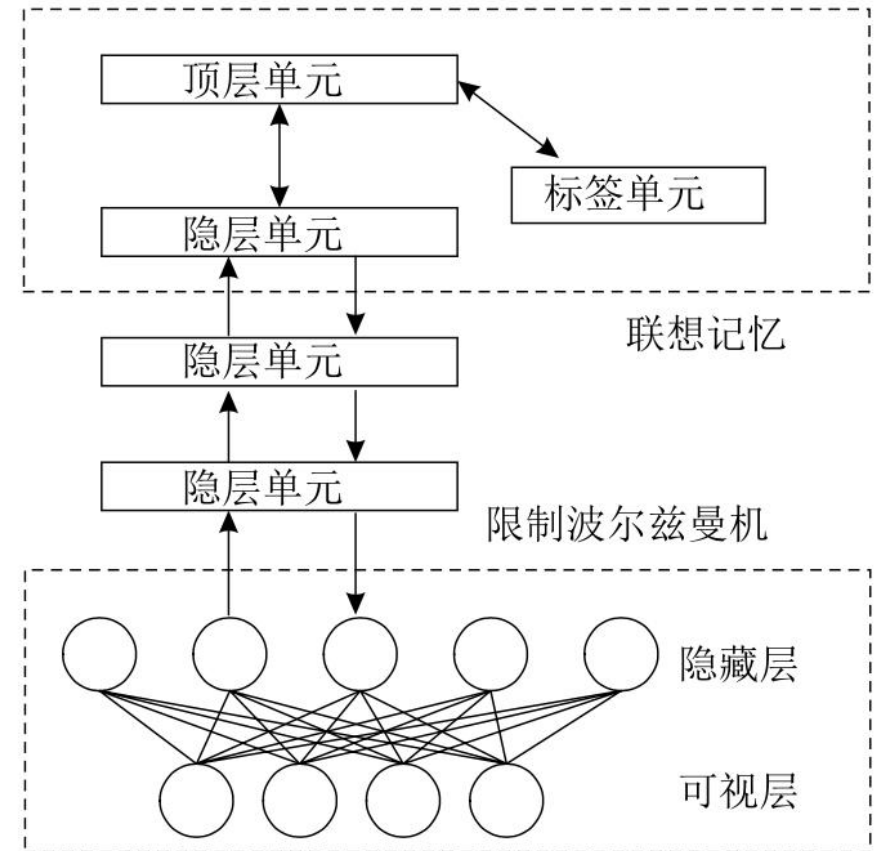
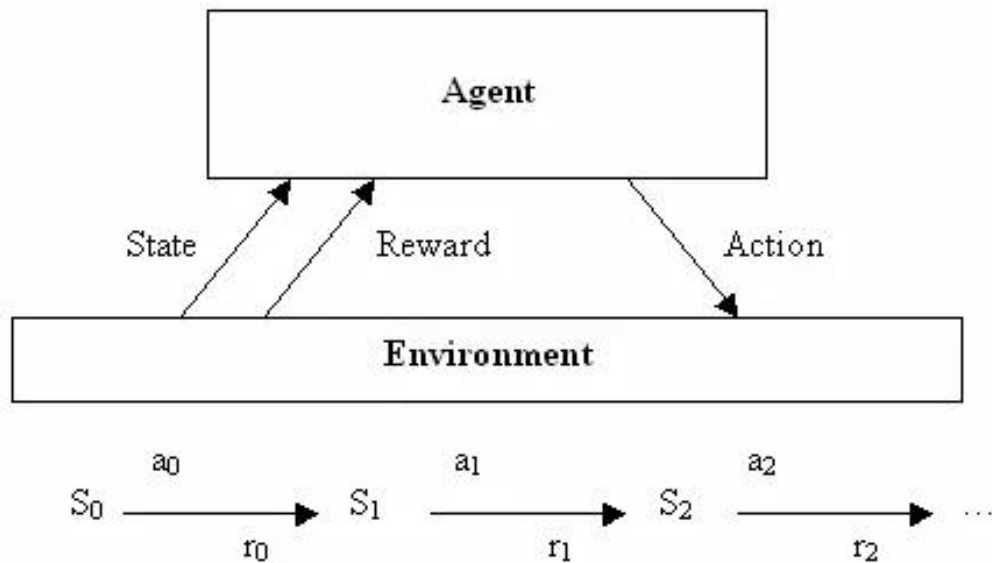


图 9 DBN 结构示意图

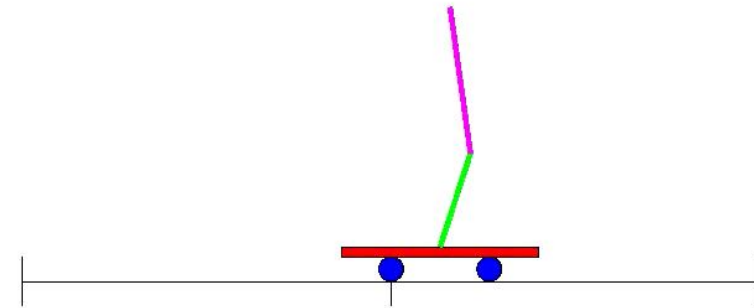
## Part2:神经网络的兴盛(1980s-2000s)

- **强化学习 (reinforcement learning)** 的出现让神经网络的作用再次扩展，可以通过 agent 和环境 (environment) 的交互，达到做出决策的效果。



Goal: learn to choose actions that maximize:  
 $r_0 + \gamma r_1 + \gamma^2 r_2 + \dots$ , where  $0 \leq \gamma < 1$

Reinforcement learning. (Source)



The double pendulum control problem - a step up from the single pendulum version, which is a classic control and reinforcement learning task.



## Part2:神经网络的兴盛(1980s-2000s)

- 为了解决出现在自然语言以及音频处理中的长序列输入问题，循环神经网络 (Recurrent Neural Network, RNN) 应运而生，通过将输出再一次输入当前神经元来赋予神经网络以“记忆” (memory)

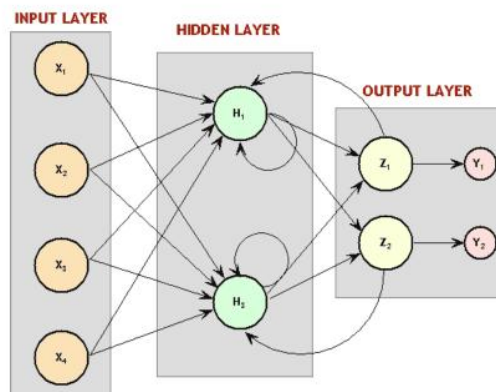
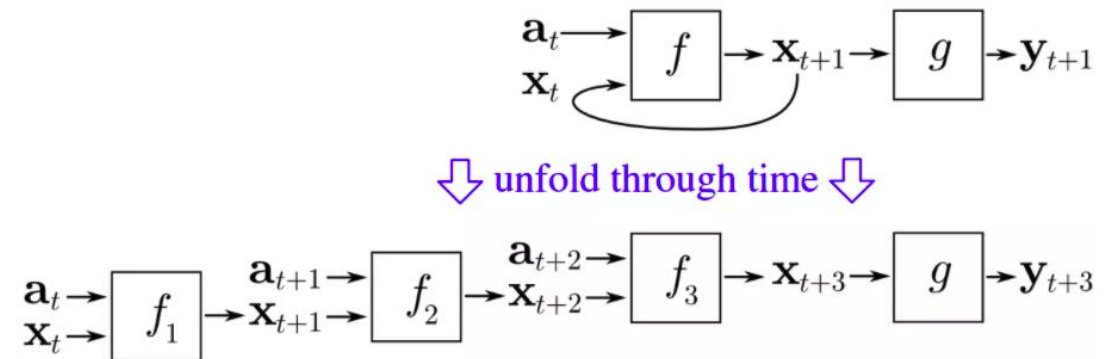


Diagram of a Recurrent Neural Net. Recall Boltzmann Machines from before? Surprise! Those were recurrent neural nets. (Source)



The wonderfully intuitive backpropagation through time concept. (Source)

## Part3:深度学习(2000s-2020s)

- 2010年后，深度学习作为神经网络研究的一个分支，在计算机视觉、自然语言处理等方向上大放异彩。2014-16年诞生的一系列代表性成果使人工智能进入了深度学习的时代。
- 深度学习发展的驱动力是“三驾马车”—— **数据、算力、算法。**

**Deep Learning =**

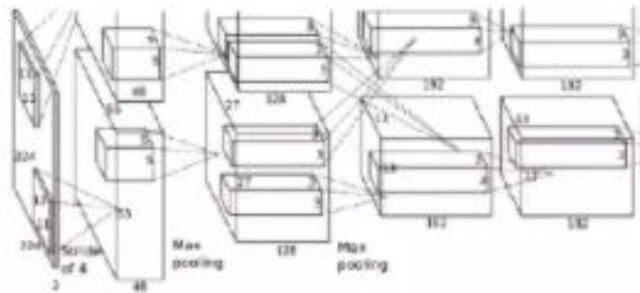
**Lots of training data + Parallel Computation + Scalable, smart algorithms**

The Deep Learning “Computer Vision Recipe”



Big Data: ImageNet

+



Deep Convolutional Neural Network

+



Backprop on GPU

=



Learned Weights



# 神经网络发展的历史：总结

## 1940s:萌芽期

- MP网络、HebbLearning

## 1950s,1960s:第一个黄金时代

- Perceptrons、ADALINE

**坚持自己的科学理念：不盲从、不放弃**

## 1970s:安静的年代

- AssociativeMemory、Brain-State-in-a-Box(BSB)

## 1980s:复苏

- Backpropagation（反向传播）、Hopfieldnets、Neocognition、Boltzmannmachine、Hardware

**今天：深度神经网络、大语言模型.....**

# 神经网络发展的历史：总结

## 诞生

1943-1956

控制论  
符号推理

## 高潮

1956-1974

搜索式推理  
自然语言处理  
反向传播算法

1970

在三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。

——马文·闵斯基

1967

一代之内.....创造 ‘人工智能’ 的问题将获得实质上的解决。

——马文·闵斯基

1965

二十年内，机器将能完成人能做到的一切工作。

——艾伦·纽厄尔和赫伯特·西蒙

1958

十年之内，计算机将成为国际象棋世界冠军。

十年之内，计算机将发现并证明一个重要的数学定理。

——艾伦·纽厄尔和赫伯特·西蒙

## 低谷/冬天

1974-1980

运算能力不够  
复杂性指数级爆炸  
没有常识  
“中文房间” 问题

## 繁荣

1980-1987

专家系统五代机  
神经网络

## 低谷/冬天

1987-1993

PC性能超过专用的AI硬件  
五代机没有实用的成果

## 发展

1993-2006

1997深蓝  
贝叶斯网络  
隐马尔科夫模型  
信息论的发展

## 突破

2007-

深度学习  
视觉/翻译在比赛中超过人类

# 神经网络重要人物及团队

# Geoffrey Hinton



- Geoffrey Hinton, 被称为“神经网络之父”、“深度学习鼻祖”。2013年, Hinton 加入谷歌并带领一个AI团队, 他将神经网络带入到研究与应用的热潮, 将“深度学习”从边缘课题变成了谷歌等互联网巨头仰赖的核心技术, 并将BP算法应用到神经网络与深度学习。
- Hinton在2017年的NIPS会议上提出的胶囊网络, 基于一种新的结构, 通过与现有的卷积神经网络(CNN)相结合, 在一些图像分类的数据上取得了非常优越的性能, 成为了2018年的发展新趋势。

# Yann LeCun



- Yann LeCun 是美国工程院院士、Facebook 前人工智能研究院院长、纽约大学Silver教授，同时还兼职于科学数据中心，数学科学交流学院，神经科学中心，以及电子工程计算机系。他于2003年加入纽约大学，之后还在普林斯顿的NEC研究院短暂任职。
- Lecun创立的卷积网络模型，被广泛地应用于计算机视觉和语音识别应用里，也因此他被称为卷积网络之父，是公认的世界人工智能三巨头之一。

# Yoshua Bengio



- Yoshua Bengio, 蒙特利尔大学的终身教授，同时是蒙特利尔大学机器学习研究所（MILA）的负责人，加拿大统计学习算法学会的主席，是CIFAR项目的负责人之一，负责神经计算和自适应感知器等方面。
- Bengio在蒙特利尔大学任教之前，是AT&T贝尔实验室&MIT的机器学习博士后。他的主要贡献在于他对循环神经网络（RNN, Recurrent Neural Networks）的一系列推动，包括经典的neural language model, gradient vanishing 的细致讨论, word2vec的雏形, 以及machine translation。
- Bengio是Deep Learning一书的合著者，且Bengio的“A neural probabilistic language model”论文开创了神经网络语言模型先河，其思路影响了之后的很多基于神经网络开展的NLP研究工作。



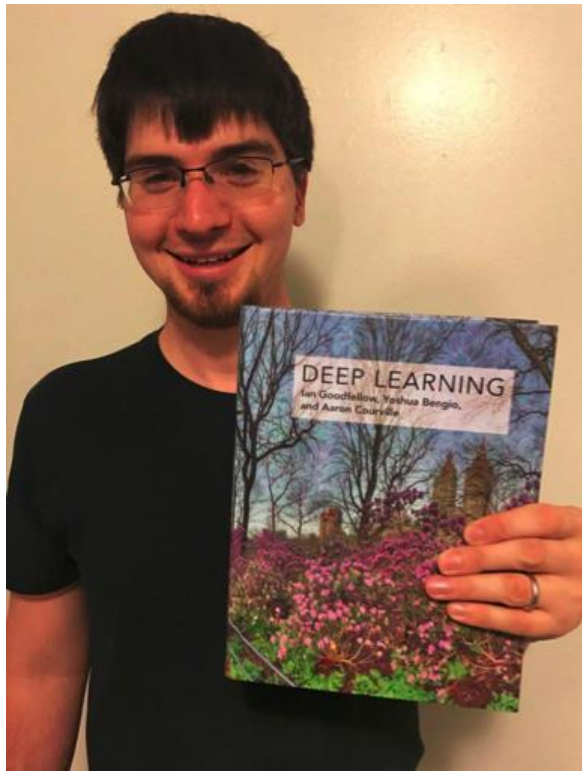
# 吴恩达 Andrew Ng



- 吴恩达，华裔美国人，是斯坦福大学计算机科学系和电子工程系副教授，人工智能实验室主任。吴恩达是人工智能和机器学习领域国际上最权威的学者之一。
- 吴恩达也是在线教育平台Coursera的联合创始人。2018年5月，吴恩达团队在MURA数据集上发起了一项深度学习挑战赛，这个数据集是他们团队在2018年1月开源的一个骨骼 X 光片的大型数据集，总共有40561 份多视图放射线影像。
- 2019年年初，他们斯坦福团队又在Nature Medicine上发表了一项研究，开发了一种深度神经网络，可基于单导程 ECG 信号分类 10 种心率不齐以及窦性心律和噪音，性能堪比心脏病医生，准确度高达83.7%，超过了人类心脏病医生的78.0%。

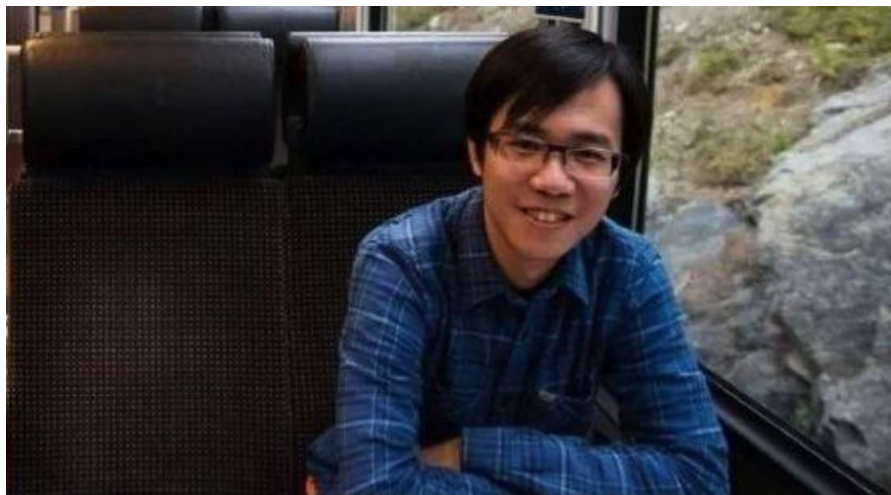


# Ian Goodfellow



- Ian Goodfellow, 人工智能领域的顶级专家，因提出了生成对抗网络（GANs）而闻名，被誉为“GANs之父”。他从斯坦福大学获得计算机科学学士、硕士学位，在蒙特利尔大学获得博士学位。
- 毕业后，Goodfellow加入Google，成为Google Brain研究团队的一员。然后他离开谷歌加入新成立的OpenAI研究所。Ian Goodfellow 在OpenAI短暂工作后，于2017年3月从OpenAI重回谷歌。Goodfellow最出名的是发明了生成性对抗网络，这是Facebook经常使用的机器学习方法。他也是Deep Learning教科书的主要作者。2017年，Goodfellow被麻省理工学院技术评论评为35位35岁以下的创新者之一。

# 何恺明



- 何恺明，2003年广东9名高考状元之一，本科就读于清华大学。博士毕业于香港中文大学多媒体实验室，研究生导师为汤晓鸥。
- 何恺明与他的同事开发了深度残余网络（ResNets），目前是计算机视觉领域的流行架构。ResNet也被用于机器翻译、语音合成、语音识别和AlphaGo的研发上。
- 2009年，何恺明成为首获计算机视觉领域三大国际会议之一CVPR “最佳论文奖”的中国学者。何恺明作为第一作者获得了CVPR 2009，CVPR 2016和ICCV 2017（Marr Prize）的最佳论文奖，并获得了ICCV 2017最佳学生论文奖。

# Open AI

- OpenAI于2015年12月在美国旧金山成立，创始人有埃隆·马斯克、萨姆·奥尔特曼等，其理念是确保人工智能技术的发展能够安全地造福全人类。
- 主要成就包括
  - GPT系列：从GPT-1到ChatGPT，以及最新的GPT-4，能够用自然语言进行问答。
  - DALL·E & CLIP：DALL·E根据文本描述生成创意图像，CLIP根据图像生成语言描述。
  - Codex：为代码自动完成工具 GitHub Copilot 提供支持的人工智能。
  - Sora：根据文本提示创建最长60秒的视频，可以深度模拟真实物理世界。
- OpenAI是微软的合作伙伴，微软为其提供云计算资源，共同开发AI技术。微软将ChatGPT的技术集成到了Microsoft Bing搜索引擎、Edge浏览器、Microsoft 365办公软件。

# 谷歌：Deep Mind人工智能实验室

- DeepMind位于英国伦敦，是由人工智能程序师兼神经科学家戴密斯·哈萨比斯(Demis Hassabis)等人联合创立，是前沿的人工智能企业，其将机器学习和系统神经科学的最先进技术结合起来，建立强大的通用学习算法。
- 最初成果主要应用于模拟、电子商务、游戏开发等商业领域。谷歌于2014年收购了该公司。
- 2016年3月，DeepMind开发的AlphaGo程序以4：1击败韩国围棋冠军李世石，成为人工智能领域的里程碑事件，使 DeepMind 成为 AI 领域的明星。
- 2018年2月27日，Deepmind提出了命名为“独角兽（Unicorn）”的智能体架构，它展示出优秀的持续学习能力。
- 2020年年底，在国际蛋白质结构预测赛中，DeepMind的AlphaFold 2摘得桂冠，并破解了一个困扰人类50年的难题：预测蛋白质如何折叠。AlphaFold开发者获2024诺贝尔化学奖。
- 2023年4月20日，谷歌谷歌母公司Alphabet表示合并旗下两个主要的人工智能研究部门——Google Brain（谷歌大脑）和DeepMind，并在23年底发布了多模态大模型Gemini。
- 24年底谷歌发布了最先进的视频生成大模型Veo2,能够以4K分辨率生成2分钟短视频。

# FAIR (Facebook's Artificial Intelligence Research)

- 2013年，Yann Lecun创立了 Facebook 人工智能研究院（FAIR），旨在通过开放研究推进人工智能的发展，并惠及所有人。FAIR 的目标是理解智能的本质，以创造真正的智能机器。自此以后，FAIR不断发展，并成长为一个国际研究组织，在门洛帕克、纽约、巴黎、蒙特利尔、特拉维夫、西雅图、匹兹堡、伦敦都设有实验室。
- FAIR组织致力于人工智能研发的各个方面，从基础研究到应用研究和技术开发。FAIR团队经常早早地发布前沿研究成果，并尽可能地开源研究代码、数据集和工具。FAIR团队研发的PyTorch是研究者最常用的深度学习框架。FAIR公布了开源大模型LLaMA系列，为推动大语言模型的研究和应用开发做出贡献。

# 神经网络的重要期刊



# 期刊

- **Journal of Machine Learning Research(JMLR)**
- **IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)**
- **IEEE Transactions on Neural Networks and Learning Systems (TNNLS)**
- **Neural Networks**
- **Neural Computation**
- **Pattern Recognition**
- **Neurocomputing**

# 会议

- **Neural Information Processing System (NeurIPS)**
- **International Conference on Machine Learning (ICML)**
- **International Conference on Learning Theory (COLT)**
- **International Conference on Learning Representations (ICLR)**
- **Artificial Intelligence and Statistics (AISTATS)**
- **Conference on Computer Vision and Pattern Recognition (CVPR)**
- **International Joint Conferences on Artificial Intelligence (IJCAI)**
- **Association for the Advancement of Artificial Intelligence (AAAI)**
- **The IEEE International Conference on Data Mining (ICDM)**

05

## 神经网络的研究者在 做哪些研究

重点：在理论研究和应用之间寻求平衡

# 实践

- 对神经元的优化：IC神经元
- 对神经网络架构的优化：ResNet引入残差，减缓梯度消失问题
- 对损失函数的优化：目标检测中引入前景和背景的权重超参，使得学习过程更加注重前景
- 针对知识表示的改进：显式或隐式求解目标分布
- 学习过程的优化：强化学习中引入了agent和环境的交互
- 学习任务的创新：元学习，从学习数据分布到学习神经网络学习的过程

**大部分针对实践的研究，都是在强调神经网络应用于特定任务后的优化问题。**

# 理论：破解神经网络的“黑盒”效应

- 利用数学公式证明某种神经网络架构优化过程的合理性
  - 循环神经网络单元理论上可以无限叠加
- 神经网络可解释性
  - 神经网络的神经元之间的互动和输出之间的联系
- 优化方法的可收敛性
  - 批次梯度下降是可以收敛的
- 优化方法的复杂度上下界
  - 二分查找的时间复杂度为  $O(\log_2(N))$

理论研究一般是为了解释某种在实际应用中表现优异的方法是否在数学上具有解释性

**目标：用理论研究来指导优化现有的神经网络模型，破解“黑盒”效应**



## 06

# 神经网络的研究方法

- 如何提出问题
- 数据从何而来
- 形成假设
- 设计研究
- 通过模型分析数据并得出结论
- 形成论文以报告研究结论
- 以此研究为基础进一步思考更加开放的问题

# 如何提出问题

## 从现实世界的所存在的问题出发，去提出科研问题

- 预测：如何用当前新冠疫情的感染人数去预测之后的感染情况？
- 分类：如何让计算机自动分拣出正常邮件和垃圾邮件？
- 识别：在一段视频中，如何自动化地识别有意义的帧？

## 从已有的解决方法中，发现不足，提出优化类的科研问题

- 现有的算法识别车牌号码准确率高达99.98%，但模型过大，无法部署使用，如何解决这个问题？

# 数据从何而来

神经网络的基础是“学习”，学习的原材料就是数据，调研任何一个科研项目时，应当首先考虑数据是否可以获得

需思考：

- 数据是否反应真实分布？
- 数据中的噪音是否能够遮盖寻找的模式？
- 数据是否完整，有无缺失，需要标注等情况？
- 例子：如何用当前新冠疫情的感染人数去预测之后的感染情况？
- 第一时间的数据是否能够获得？是否涉密？
- 数据来自区域还是全国，它是否能够反映疫情的总体趋势？

# 形成假设

从数据类型出发思考输入和输出，并提出对模型的要做的任务做出假设

需思考：

- 数据类型是什么？文本，图像，视频，音频等
- 有无标注？需要用到有监督学习，还是无监督学习？
- 希望的输出是什么样的？类别，数值或是矩阵？
- 输出和输入的关系是什么？
- 数据大约有多少，是否能足够训练所假设的模型？
  - 数据过少，要扩充样本还是压缩模型？
  - 数据过多，需要筛选数据还是增大模型？
  - 数据样本过大，应该如何切分？

# 设计研究

可以选择在现有模型中改进，也可以选择另辟蹊径，针对问题，开发新的模型架构。

需思考：

- 我所选择的模型是否有部署要求？是否需要实时性而不宜过大？
- 我所选择的模型应该优化什么方面？架构，神经元，损失函数还是训练流程？
- 我优化后的模型，应该在什么方面超越当前常用的模型？

# 通过模型分析数据并得出结论

通过对选定数据集的训练，得到当前模型的性能数据。通过调节超参数等，可以将模型优化到最佳状态。得到的性能数据需要与其他模型进行比较，并且分析优势和劣势，及其背后的原因。

需思考：

- 我对模型的调整是否有理论支撑，有没有数学意义？
- 我的模型如果只对特定的数据集表现优异，原因是什么？
- 当前最佳的模型和我的模型相比，为什么优于/劣于我的模型？



# 形成论文以报告研究结论

当进行完对比实验以后，我们对自己的模型就有了完整的评估。此时，应当对我们的工作进行总结，即将以上的几部分总结成论文的形式。

- 在论文的写作过程中，可能还需要再次修改模型，增加实验等。
- 在论文完成后，和指导老师商议，应该投稿到什么类型的论文或期刊。让更多的研究者可以参考自己的工作成果。

# 以此研究为基础进一步思考更加开放的问题

论文并不是研究的终点，好的研究应该是发散性的，甚至是成系统的。从一个问题的解决能引出另一个问题的提出。从当前论文中总结出的未来可改进的方向中，可以继续精进当前的研究。

也可以跳脱当前论文的限制，而思考在更高层次上如何解决遇到的问题。

- 例如：
- 当前文章解决了数据切分的问题，在下一个工作中，可以用切分的数据去做预测，并改进模型，达到更优秀的效果。
- 当然，也可以彻底地解决需要切分的问题，即研究新的硬件设备以支持完整数据的存储。

07

## 思考题

- 你认为将来可能怎样研究大脑？
- 从受精卵到幼儿期，一个人在什么时候变得有意识？而我们怎么知道的？
- 对大脑的研究能促进神经网络的研究吗？
- 对神经网络的研究能帮助人们更好地发现大脑的奥秘吗？解决动物实验的伦理问题？
- 从应用上看，可以通过大脑直接控制机械吗？
- 反过来，对于大脑不能正常工作的例如植物人，可以通过人工神经网络来控制他的肢体吗？
- 阐述生物神经系统的运行流程是怎样的？与现代人工神经网络执行流程有什么关系？

# Q&A

Questions and Answers