



# 机器读心术之神经网络与深度学习 第1周

---

DATAGURU专业数据分析社区

## 法律声明

**【声明】本视频和幻灯片为炼数成金网络课程的教学资料，所有资料只能在课程内使用，不得在课程以外范围散播，违者将可能被追究法律和经济责任。**

**课程详情访问炼数成金培训网站**

**<http://edu.dataguru.cn>**

# 关注炼数成金企业微信



- 提供全面的数据价值资讯，涵盖商业智能与数据分析、大数据、企业信息化、数字化技术等，各种高性价比课程信息，赶紧掏出您的手机关注吧！



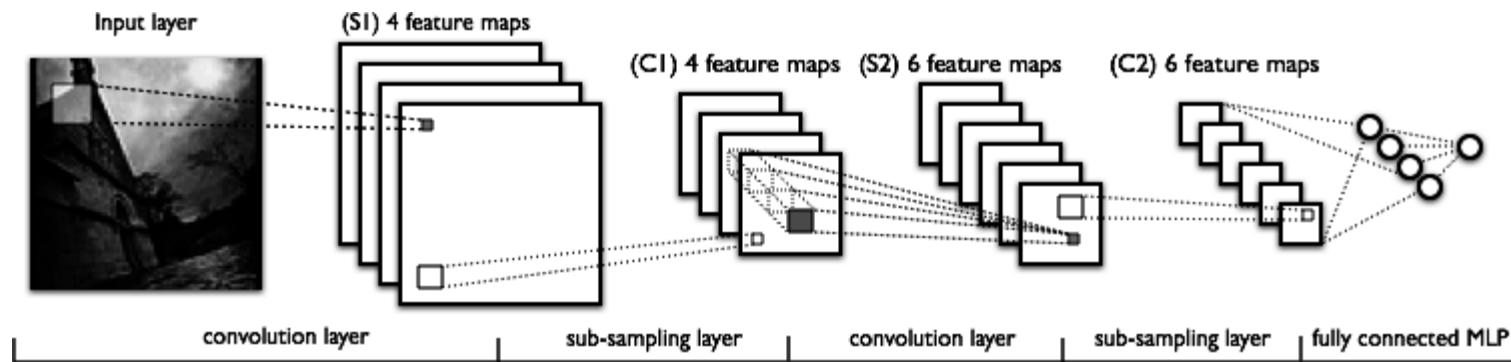
DATAGURU专业数据分析社区

# 课程内容

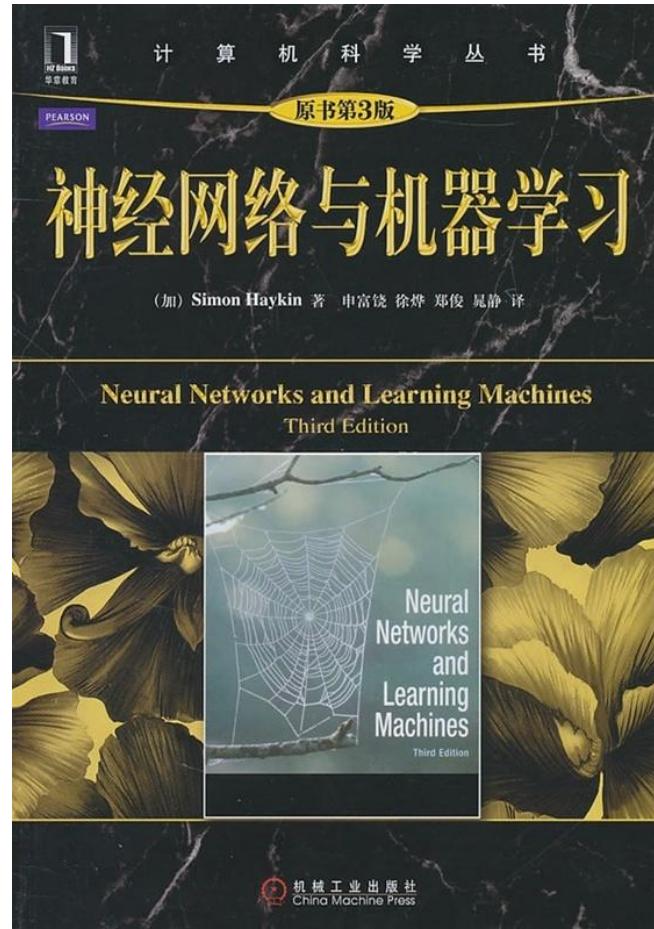
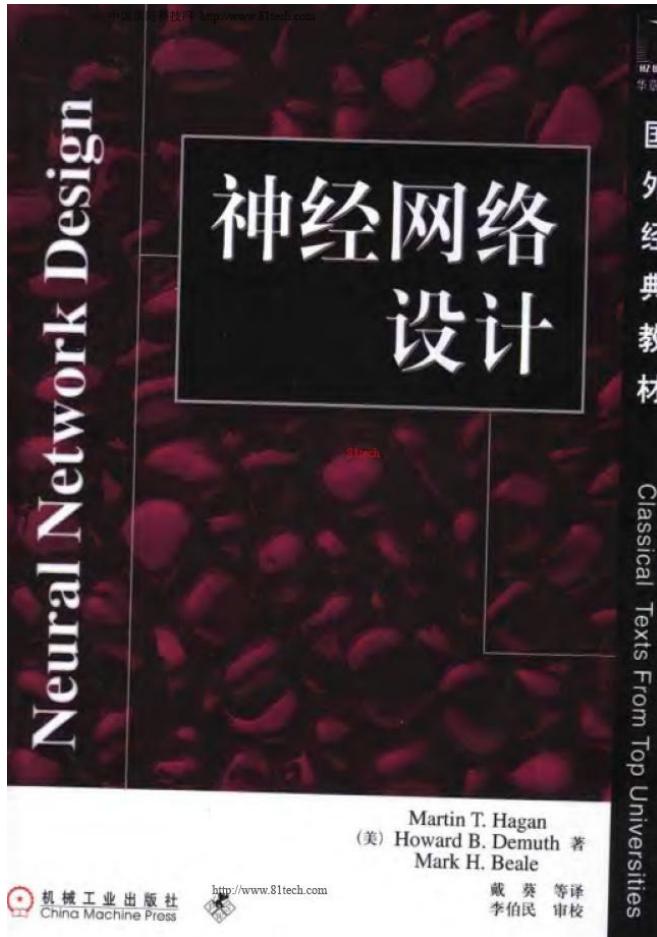
- 神经网络基础：单层感知器，线性神经网络，BP神经网络，Hopfield神经网络，径向基神经网络，PCA与SVM神经网络
- 神经网络进阶：自编码器，稀疏自编码器，玻尔兹曼机，受限玻尔兹曼机，递归神经网络，自组织竞争神经网络
- 深度学习网络：深度置信网络，卷积神经网络，深度残差网络
- 神经网络应用：应用于传统的数据挖掘与机器学习问题，手写体识别，图像识别，应用于自然语言处理，应用于人工智能：Alphago
- 算法为主的课程，需要数学和机器学习课程的基础，有部分编程练习（语言任选），有部分演示基于Matlab，R
- 深度学习部分落地实现：Caffe，Tensorflow
- 经过体系化整理的课程，有清晰的学习线路图和知识体系

# 课程目标

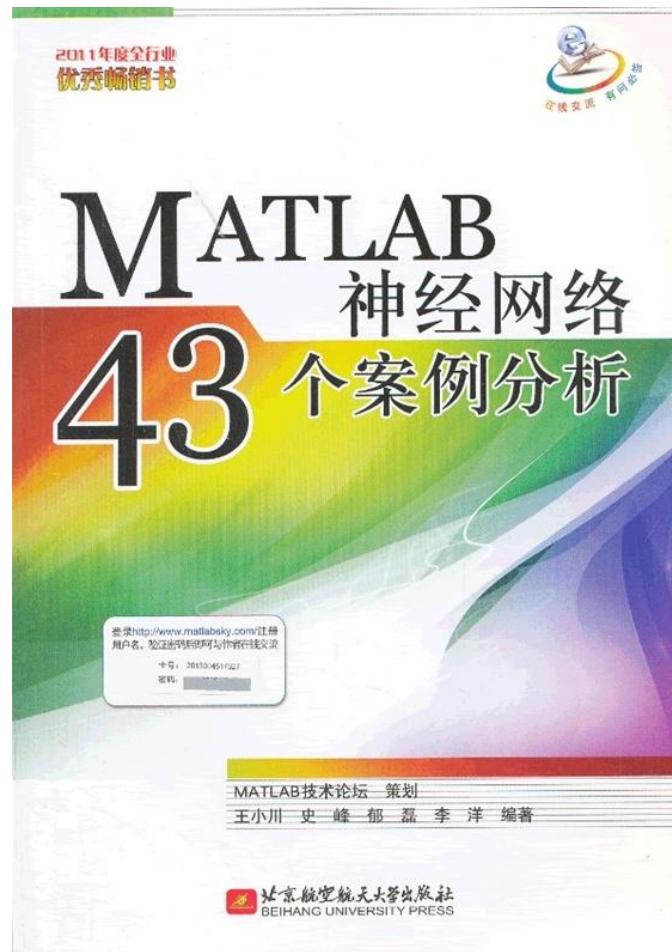
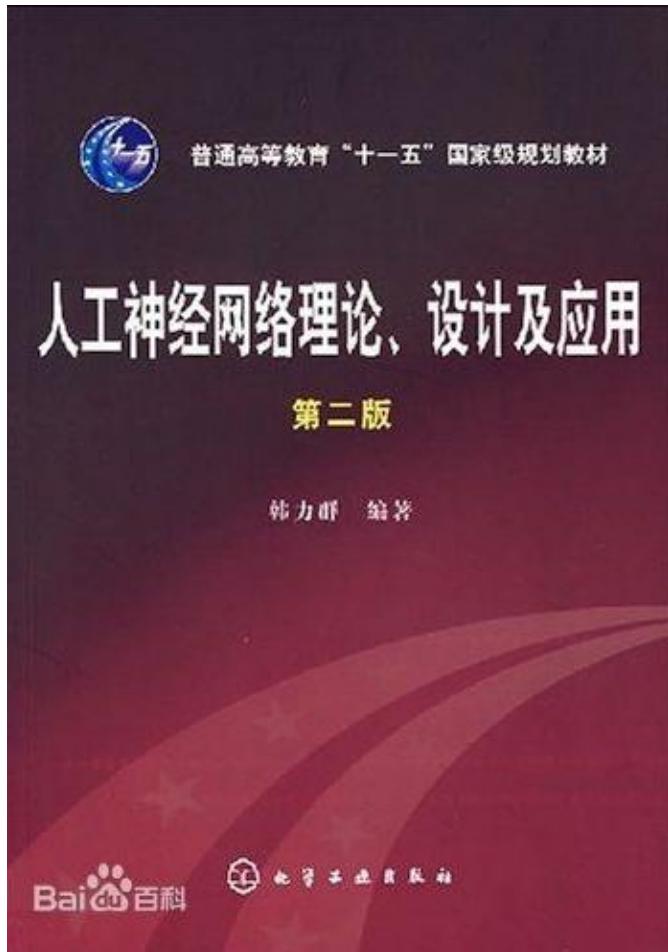
- 极度熟悉基础神经网络，熟悉深度学习网络的原理，技术细节，及应用场景
- 有能力在软件辅助下设计自己的神经网络系统解决实际问题
- 为创业者提供新的技术思路
- 卫星课程：《深度学习框架Caffe学习与应用》（<http://www.dataguru.cn/course-209.html>）



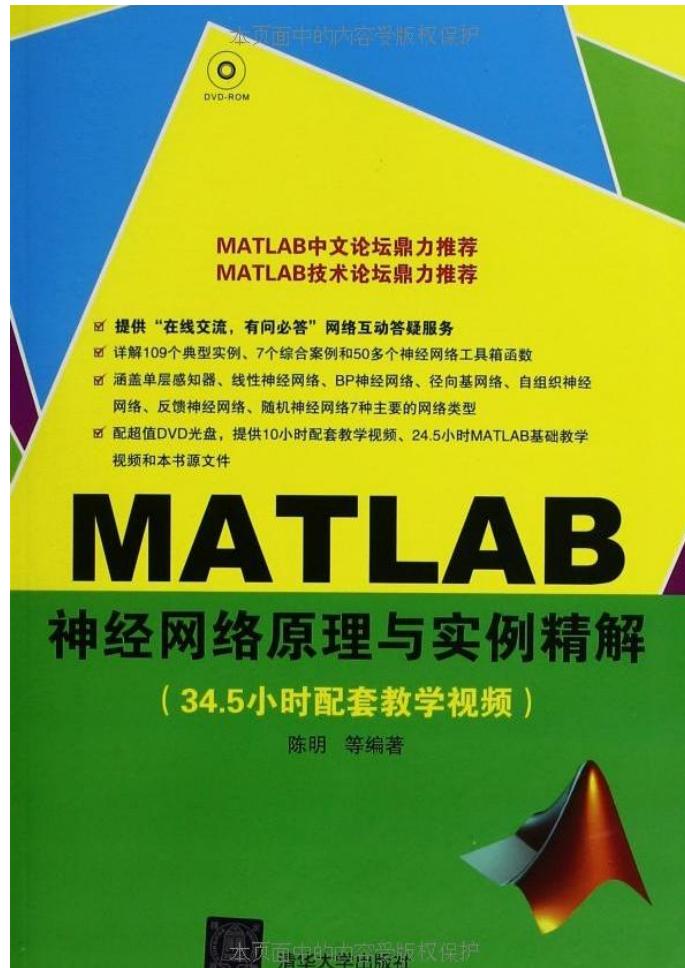
# 参考书（基础部分）



# 参考书（基础部分）



# 参考书（基础部分）

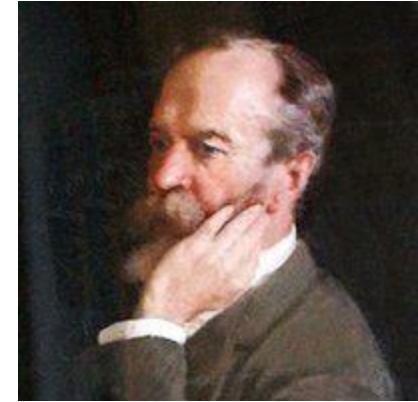


# 人工神经网络激荡的70年

- “思维的本质”一直是人类思索的本源问题之一，研究、模仿人类大脑和神经系统的工作机制，并且使用机器去仿真，获得具有人一般的逻辑能力，联想能力，推理能力，抽象能力的机器，然而机器却是永不疲倦，具有比人类更快的运转速度，且适合大量复制。
- 神经网络的研究可以追溯到19世纪末期，美国心理学家William James的《心理学原理》，首次详细论述人脑结构与功能，对学习，联想记忆相关基本原理作了开创性研究，此为启蒙时期。
- [http://baike.baidu.com/link?url=5YIrkfLMne3XgESOp0H9SnGQCvMhEidnK\\_N165VvLodz5N-EdIZwFdaXJmn\\_C6UsORq1dHVJiDZ7MquNBgDT91GkRchzcHKyUEoTstQ4V5kYZOPbT--Mq0qMHLfFhTmv1FcIicuEDF8fIYxcJDGlq](http://baike.baidu.com/link?url=5YIrkfLMne3XgESOp0H9SnGQCvMhEidnK_N165VvLodz5N-EdIZwFdaXJmn_C6UsORq1dHVJiDZ7MquNBgDT91GkRchzcHKyUEoTstQ4V5kYZOPbT--Mq0qMHLfFhTmv1FcIicuEDF8fIYxcJDGlq)

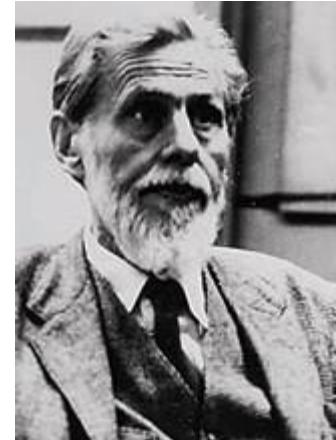
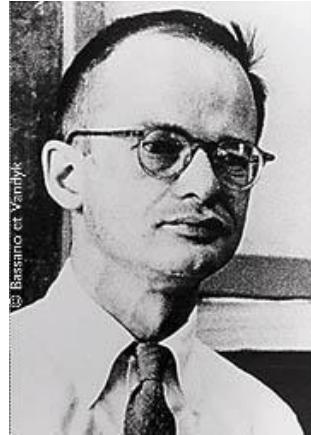
# William James的名言

- 播下一个行动，收获一种习惯；播下一种习惯，收获一种性格；播下一种性格，收获一种命运。
- 人类本质中最殷切的需求是渴望被肯定。
- 人的难题不在于他想采取何种行动，而在与他想成为何种人。
- 我这一代人最大的发现是人类可以通过改变态度来改变生活。
- 邪不压正，正不压疯



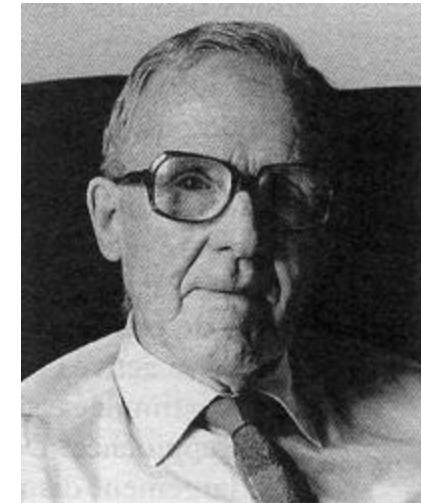
# M-P模型

- 1943年，精神病学家和神经元解剖学家McCulloch和数学天才Pitts在生物物理学期刊发表文章提出神经元的数学描述和结构。
- 神经元遵循“全或无”原则。
- 证明了只要有足够的简单神经元，在这些神经元相互连接并同步运行的情况下，可以模拟任何计算函数（M-P模型）。
- 上述开创性工作被认为是人工神经网络的起点。

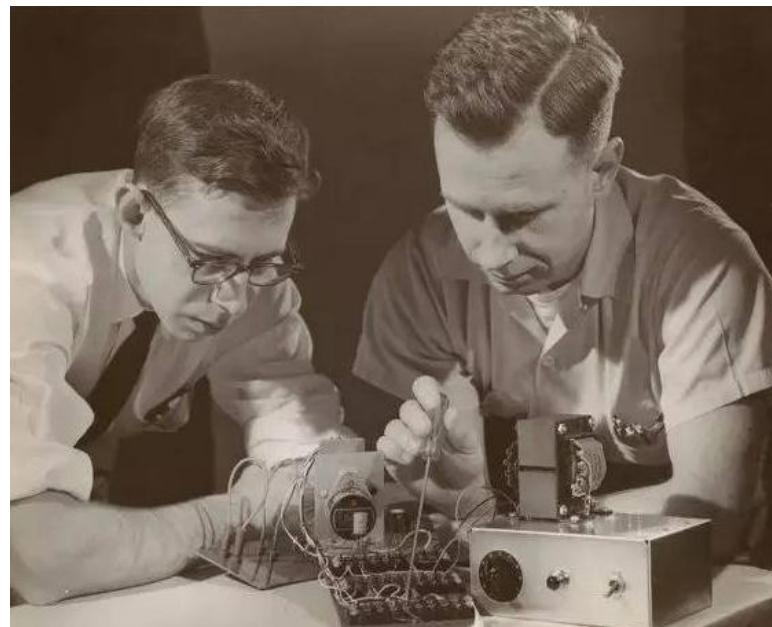


# Hebb学习规则

- 1949年，生理学家Hebb出版了《行为组织学》。描述了神经元权值的Hebb调整规则。
- 提出“连接主义”。
- 引入“学习假说”，即两个神经元之间重复激活，将使其连接权值加强。
- 成为学习系统和自适应系统的灵感源泉。



- 1957年，Rosenblatt提出感知器的概念。掀起了神经网络研究第一次热潮
- 提出感知器收敛定理。
- Widrow和Hoff引入最小均方误差准则（LMS）
- 整个60年代，感知器是如此流行，人们认为它可以完成任何事情，只要很多感知器连在一起，就可以完成大脑的工作



*Psychological Review*  
Vol. 65, No. 6, 1958

## THE PERCEPTRON: A PROBABILISTIC MODEL FOR INFORMATION STORAGE AND ORGANIZATION IN THE BRAIN<sup>1</sup>

F. ROSENBLATT

*Cornell Aeronautical Laboratory*

If we are eventually to understand the capability of higher organisms for perceptual recognition, generalization, recall, and thinking, we must first have answers to three fundamental questions:

1. How is information about the physical world sensed, or detected, by the biological system?
2. In what form is information stored, or remembered?
3. How does information contained in storage or in memory influence

and the stored pattern. According to this hypothesis, if one understood the code or "wiring diagram" of the nervous system, one should, in principle, be able to discover exactly what an organism remembers by reconstructing the original sensory patterns from the "memory traces" which they have left, much as we might develop a photographic negative, or translate the pattern of electrical charges in the "memory" of a digital computer. This hypothesis is appealing in its

# Minsky的打击

- 1969年，Minsky和Papert出版《感知器》。指出：单层感知器存在致命缺陷，简单的神经网络只能运用于线性问题的求解，连易或这样的问题都难以解决。求解非线性问题的网络应具有隐层，但理论上还不能证明感知器扩展到多层是有意义的。单层感知器的局限性在多层情况下也不能被完全解决。
- 由于Minsky在业界的地位，这些悲观观点极大影响了人工神经网络研究，经费萎缩，大量研究人员流失，启蒙时期结束，此后陷入10年低潮时期。



# Marvin Minsky

- MIT人工智能实验室联合创始人，人工智能之父。美国科学院和美国工程院院士。他曾出任美国人工智能学会AAAI的第三任主席（1981-1982）。
- 1969年获得了计算机科学领域的最高奖图灵奖。
- [http://blog.sina.com.cn/s/blog\\_842bf5cd010168yt.html](http://blog.sina.com.cn/s/blog_842bf5cd010168yt.html)
- 主要著作：

《计算：有限与无限的机器》(Computation : Finite and Infinite Machines, Prentice-Hall , 1967)

《语义信息处理》(Semantic Information Processing , MIT Pr. , 1968)

《感知器》(Perceptrons , MIT Pr. , 1969 ; expanded edition , 1988)

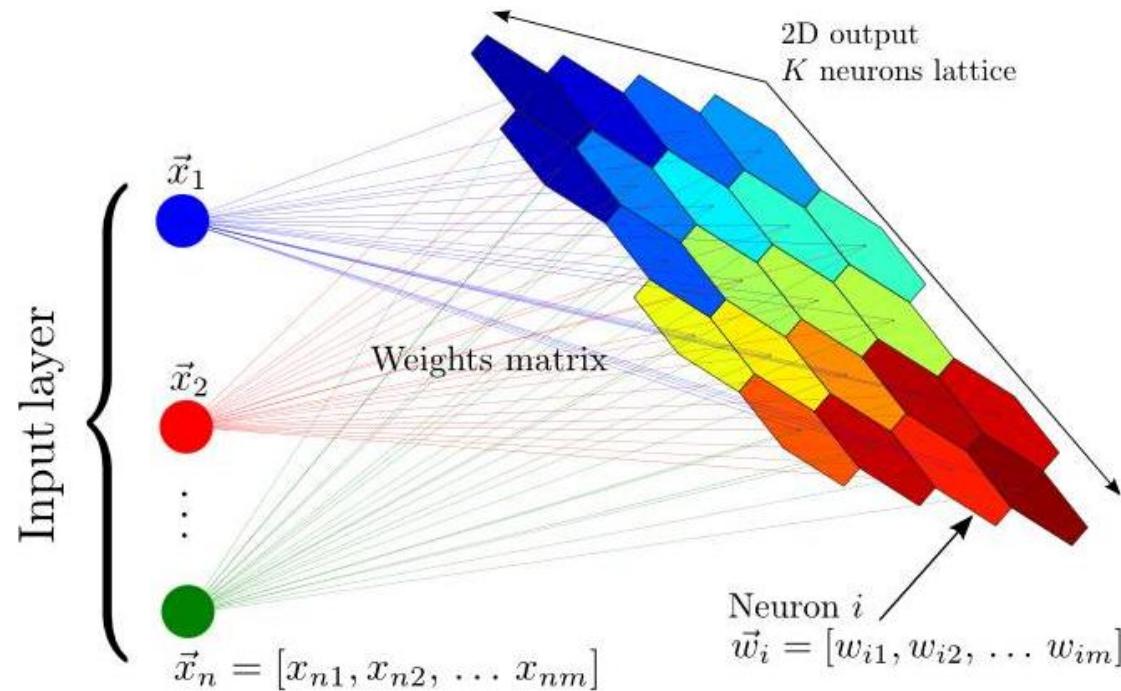
《表示知识的框架》(A Framework for Representating Knowledge, McGraw-Hill ,1975)

《心智的社会》(The Society of Mind , Simon & Schuster ,1986)

《机器人学》( Robotics, Anchor Pr./Doubleday, 1985 )

# 低潮时期的工作

- Kohonen提出SOM（自组织映射），自组织网络
- Grossberg夫妇提出自适应共振理论，提出三个ART系统
- 福岛邦彦的“新认知机”



# 复兴时期

- 60年代末遭受的质疑均在80年代被攻克
- 1982年，美国加州工学院物理学家J.J.Hopfield提出了 Hopfield神经网格模型（<http://alex-wbf.blog.sohu.com/131663695.html>）
- 1986年，Rumelhart, Hinton, Williams发展了BP算法（多层感知器的误差反向传播算法）。Rumelhart和McClelland出版了《Parallel distribution processing: explorations in the microstructures of cognition》。迄今，BP算法已被用于解决大量实际问题。



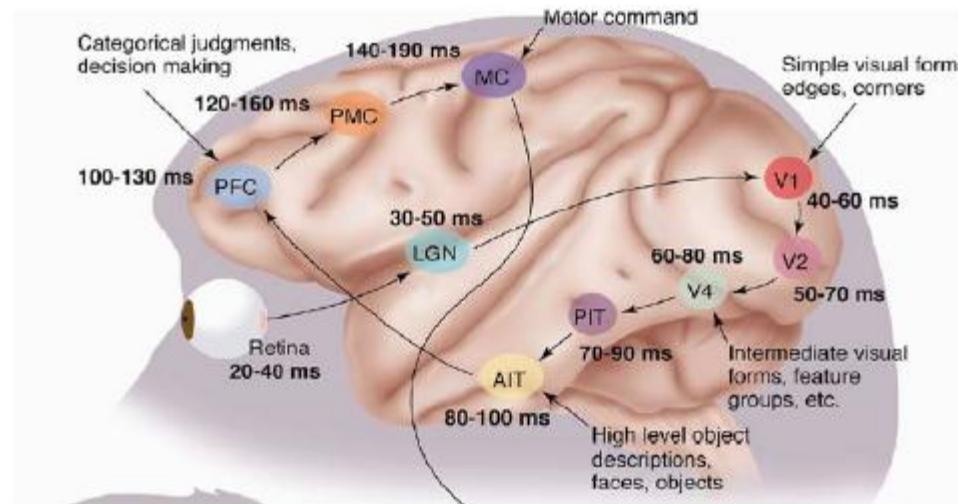
# Geoffrey Hinton

- <http://www.cs.toronto.edu/~hinton/>
- 多伦多大学计算机科学系教授，机器学习与人工智能顶级专家，诸多神经网络原创性研究的发起者和关键技术推手
- 目前领导Google Brain项目



# 更多层的神经网络？

- 三层BP神经网络的学习算法能推广到更多层吗？（学习信号越来越弱）
- 求解的问题通常不是凸的，意味着有多个极值点
- 梯度下降法的弱点导致容易陷入局部极值点
- 只有w权值取得相当接近理想值时，才能得到较好的结果
- 浅层神经网络无法模拟人类复杂的行为
- 文本分析，自然语言处理，图像识别等领域需要能深层次学习的神经网络



# 深层学习的突破

## ■ Hinton的革命性突破

materials are identical for all configurations. The blue bars in Fig. 1 summarize the measured SHG signals. For excitation of the *LC* resonance in Fig. 1A (horizontal incident polarization), we find an SHG signal that is 500 times above the noise level. As expected for SHG, this signal closely scales with the square of the incident power (Fig. 2A). The polarization of the SHG emission is nearly vertical (Fig. 2B). The small angle with respect to the vertical is due to deviations from perfect mirror symmetry of the SRRs (see electron micrographs in Fig. 1). Small detuning of the *LC* resonance toward smaller wavelength (i.e., to 1.3- $\mu\text{m}$  wavelength) reduces the SHG signal strength from 100% to 20%. For excitation of the Mie resonance with vertical incident polarization in Fig. 1D, we find a small signal just above the noise level. For excitation of the Mie resonance with horizontal incident polarization in Fig. 1C, a small but significant SHG emission is found, which is again po-

## Reducing the Dimensionality of Data with Neural Networks

G. E. Hinton\* and R. R. Salakhutdinov

High-dimensional data can be converted to low-dimensional codes by training a multilayer neural network with a small central layer to reconstruct high-dimensional input vectors. Gradient descent can be used for fine-tuning the weights in such “autoencoder” networks, but this works well only if the initial weights are close to a good solution. We describe an effective way of initializing the weights that allows deep autoencoder networks to learn low-dimensional codes that work much better than principal components analysis as a tool to reduce the dimensionality of data.

Dimensionality reduction facilitates the classification, visualization, communication, and storage of high-dimensional data. A simple and widely used method is principal components analysis (PCA), which

finds the directions of greatest variance in the data set and represents each data point by its coordinates along each of these directions. We describe a nonlinear generalization of PCA that uses an adaptive, multilayer “encoder” network

# 深度学习

- 深度学习是机器学习研究中的一个新的领域，其动机在于建立、模拟人脑进行分析学习的神经网络，它模仿人脑的机制来解释数据，例如图像，声音和文本。深度学习是无监督学习的一种。
- 深度学习的概念源于人工神经网络的研究。含多隐层的多层感知器就是一种深度学习结构。深度学习通过组合低层特征形成更加抽象的高层表示属性类别或特征，以发现数据的分布式特征表示。
- 深度学习的概念由Hinton等人于2006年提出。基于深信度网(DBN)提出非监督贪心逐层训练算法，为解决深层结构相关的优化难题带来希望，随后提出多层自动编码器深层结构。此外Lecun等人提出的卷积神经网络是第一个真正多层结构学习算法，它利用空间相对关系减少参数数目以提高训练性能。
- 炼数成金深度学习专题：<http://f.dataguru.cn/topic-deep-learning.html>

# Andrew Wu

- Cousera创始人。原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。
- 参与创建Google Brain项目
- 2014年5月19日，百度宣布任命吴恩达博士为百度首席科学家，全面负责百度研究院。包括百度大脑计划。这是中国互联网公司迄今为止引进的最重量级人物。

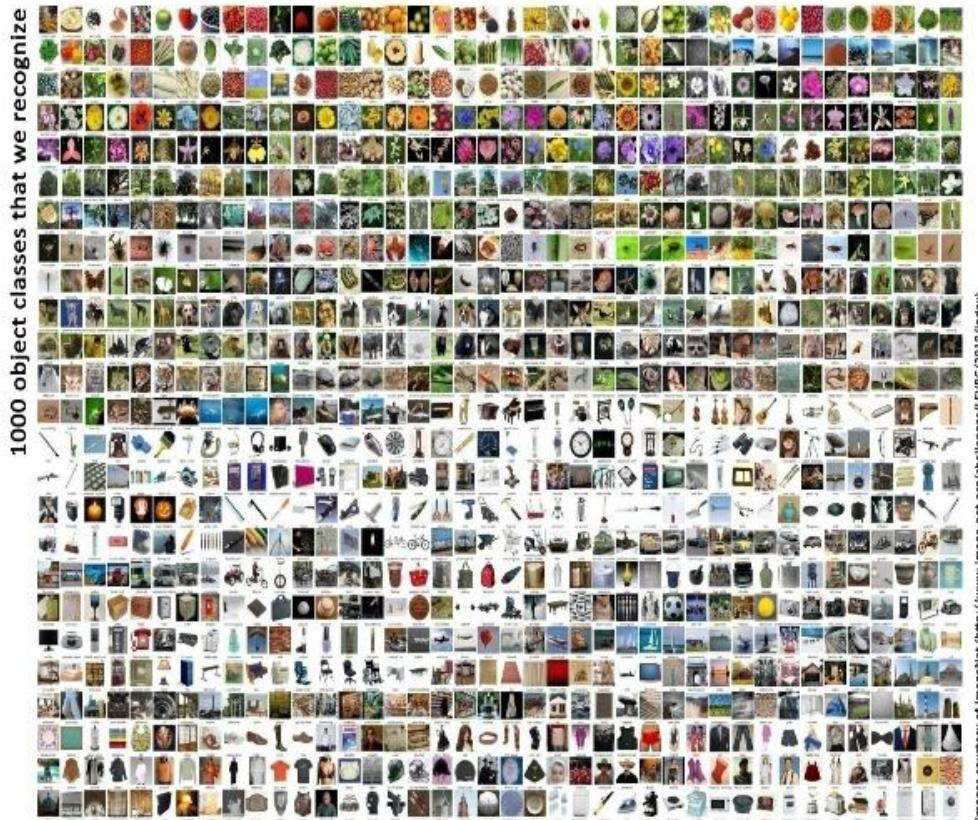


- 谷歌大脑是“Google X实验室”一个主要研究项目。是谷歌在人工智能领域开发出的一款模拟人脑的软件，这个软件具备自我学习功能。Google X部门的科学家们通过将1.6万台电脑的处理器相连接建造出了全球为数不多的最大中枢网络系统，它能自主学习，可以称之为“谷歌大脑”。
- Google Brain运用深度学习的研究成果，使用1000台电脑创造出多达10亿个连接的“神经网络”，让机器系统学会自动识别猫，成为国际深度学习领域广为人知的案例之一。
- 通过应用这个神经网络，谷歌的软件已经能够更准确的识别讲话内容，而语音识别技术对于谷歌自己的智能手机操作系统Android来说已经变得非常重要，而同样这一技术也可以用于谷歌为苹果iPhone开发的应用程序。
- 通过神经网络，能够让更多的用户拥有完美的、没有错误的使用体验。随着时间的推移，谷歌的其他产品也能随之受益。例如谷歌的图像搜索工具，可以做到更好的理解一幅图片，而不需要依赖文字描述。
- 而谷歌无人驾驶汽车、谷歌眼镜也能通过使用这一软件而得到提升，因为他们可以更好的感知真实世界中的数据。

# 得人工智能者得天下



- 余凯的PPT：<http://pan.baidu.com/share/link?shareid=136269&uk=2267174042>
- Geoffrey Hinton挑战图像识别新极限



72%, 2010

74%, 2011

85%, 2012

# Yann LeCun

- Yann LeCun是人工智能尤其是深度学习领域最知名的学者之一，1960年生于法国巴黎。在居里大学读博士期间，提出过后向传播学习算法。
- 在多伦多大学随Hinton读博士后。1988-2002年加盟贝尔实验室，期间研发了卷积神经网络（Convolutional Neural Networks）与曾广泛用于手写识别和OCR的图变换网络方法。当然，图形压缩技术DjVu可能更为大家所熟知。
- 2003年加入纽约大学，从事广度与深度兼具的各类研究，涉及机器学习、计算机视觉、移动机器人和计算神经学。他甚至还开发了一种开源的面向对象编程语言Lush，比Matlab功能还要强，当然他也是一位Lisp高手。
- 获得了IEEE计算智能学会的神经网络先驱奖，将在2014年北京计算智能大会上颁奖。
- 2013年底Facebook CEO Mark Zuckerberg在NIPS会议的演讲期间宣布Facebook成立新的人工智能实验室，宣称实验室将具有“宏大长远的目标，要给人工智能带来重大突破”，Yann LeCun担任负责人。LeCun仍将兼任纽约大学的教授，继续在该校从事研究和教学工作。与此同时，Facebook和纽约大学数据科学中心将在数据科学、机器学习和人工智能领域展开合作。



# AlphaGo



DATAGURU专业数据分析社区

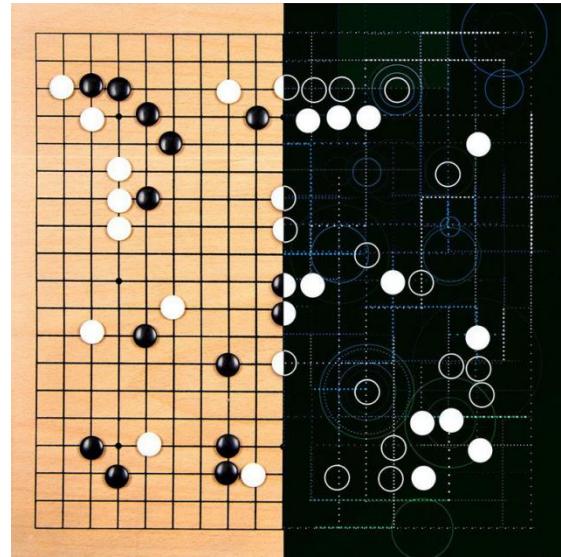
# Demis Hassabis与Google Deepmind



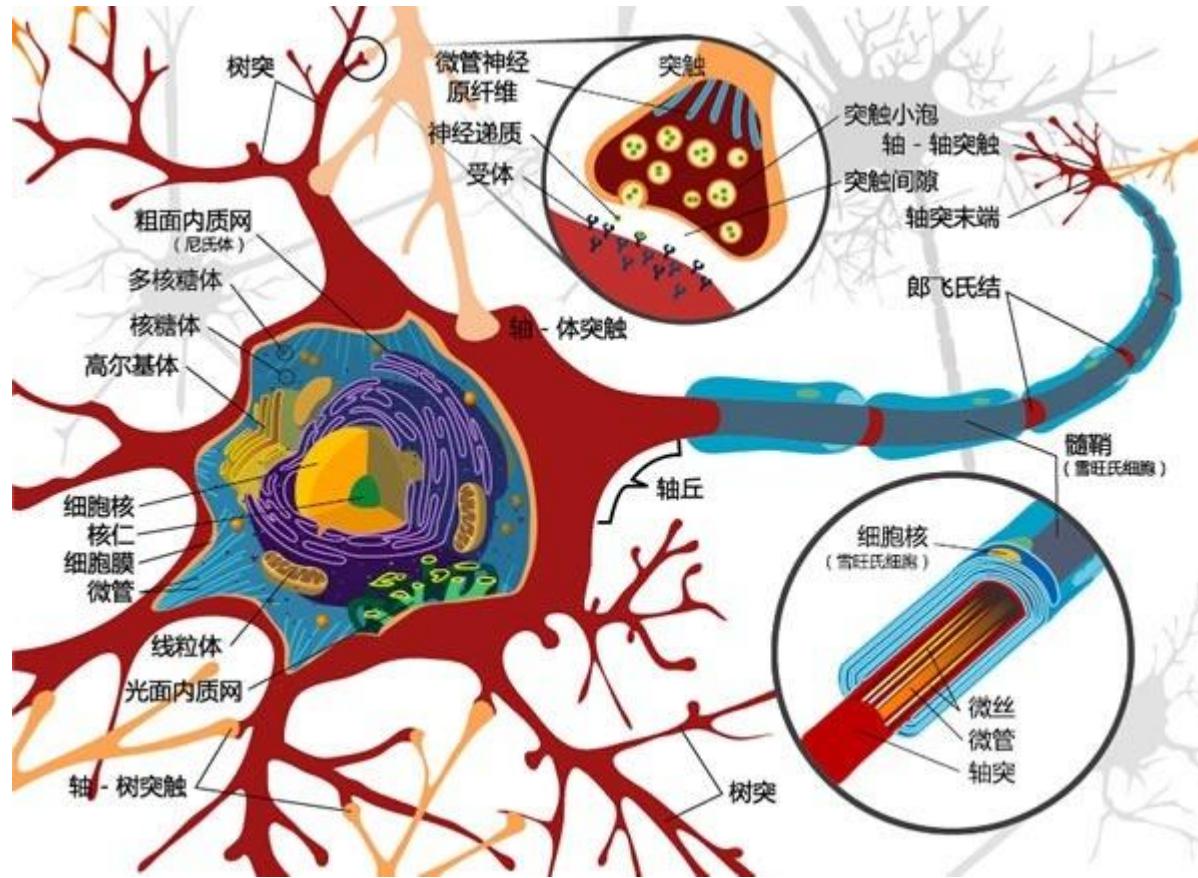
DATAGURU专业数据分析社区

# 深度学习引发人工智能革命

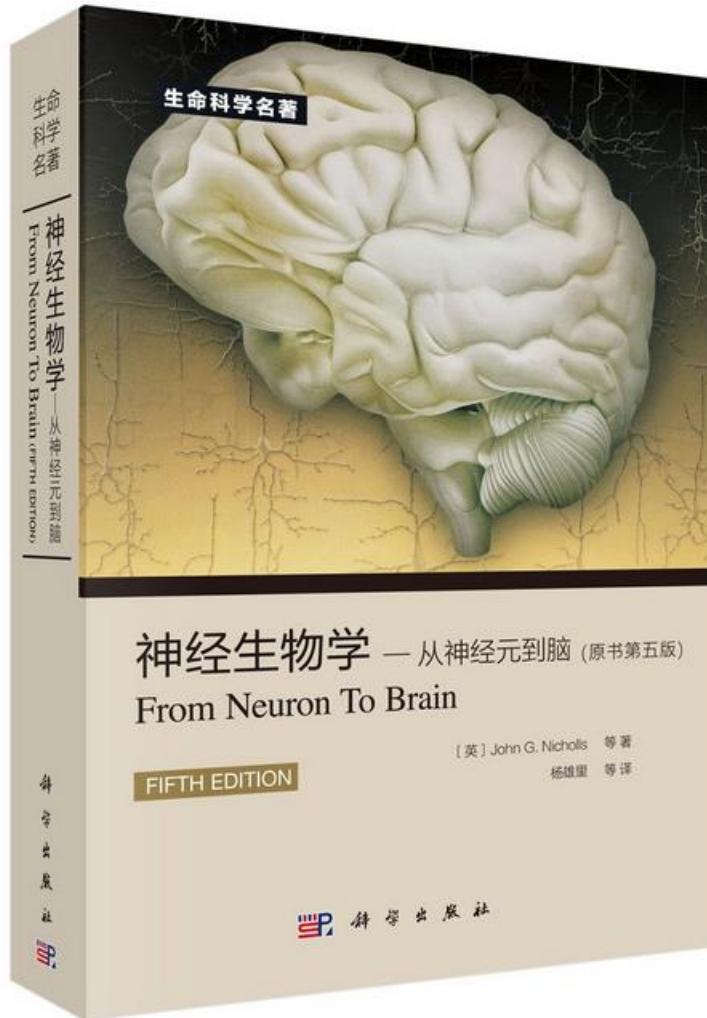
- 深度学习正在引发一场深刻的技术革命，这是人类首次如此接近思维的本质。象手写体识别，脸像识别这类系统，过去的思路是从业务背景中线提取特征，然后产生若干辨识逻辑，再形成算法编程实现，但对于像imagenet那样要对上百万的图片进行上千个分类识别的问题，以往的技术就傻眼了，别说逻辑，连特征的提取都因为过于复杂而没办法进行。现在流行的深度学习网络的方法，把逻辑隐藏在成百上千万的神经网络权值里，让特征被自动识别与提取，却能得出让人吃惊的高准确率。给出通用的框架，通过大量学习数据训练出合适的权值，权值就是逻辑，这是未来的方向，那种先设计算法敲代码的日子该一去不复返了，以后甚至程序员的工作都由机器全部完成也不是没有可能。



## ■ 人类神经系统原理

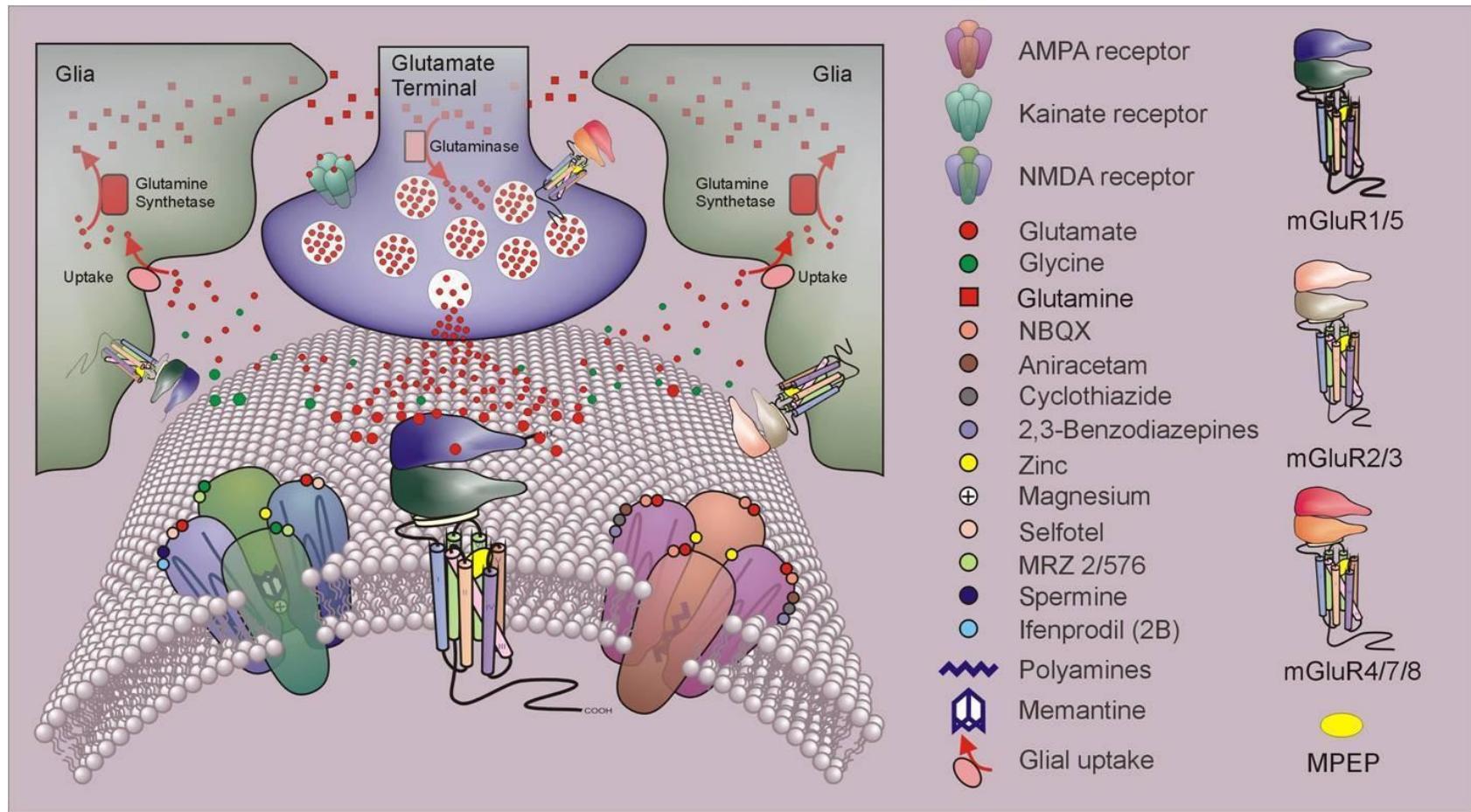


# 参考书

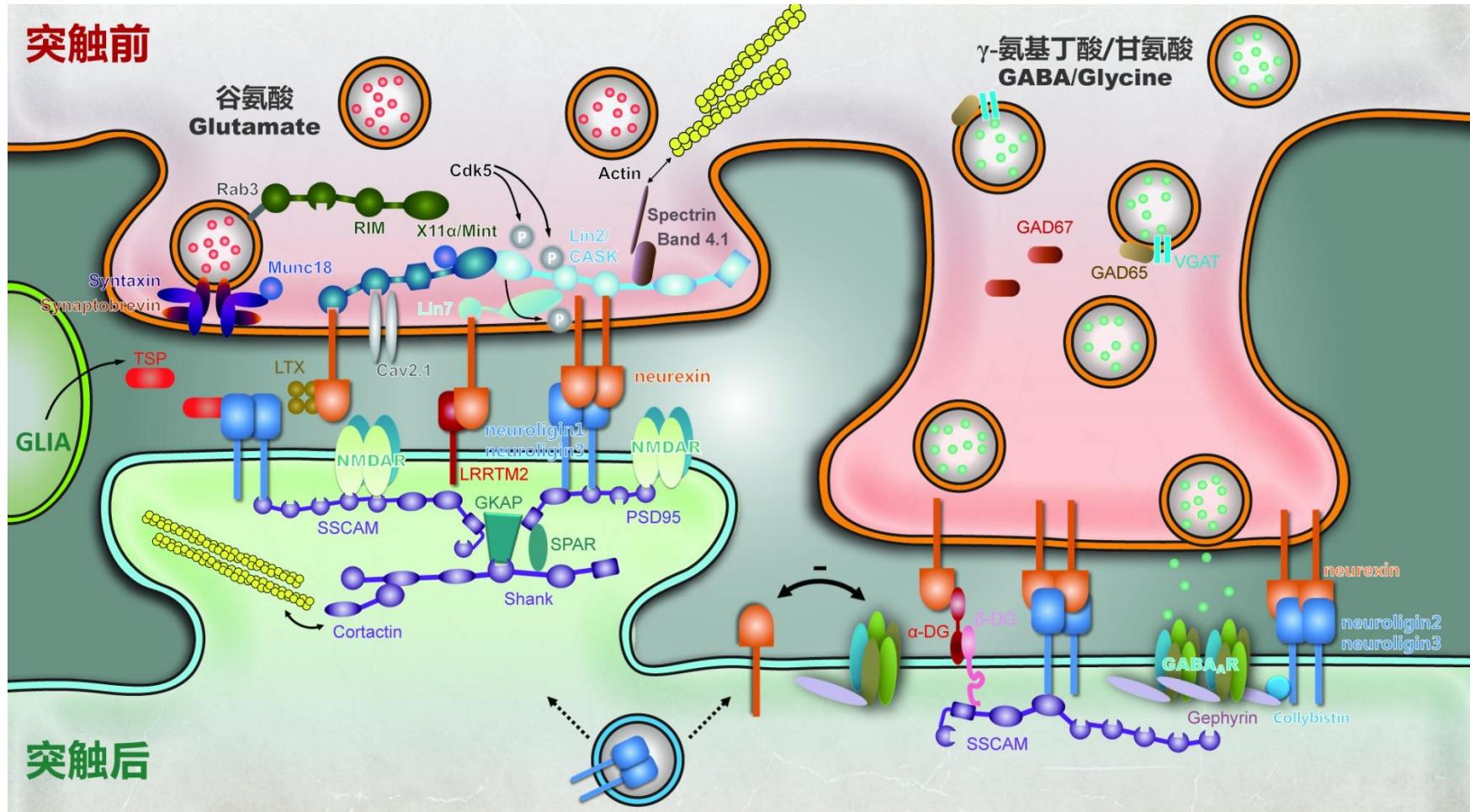


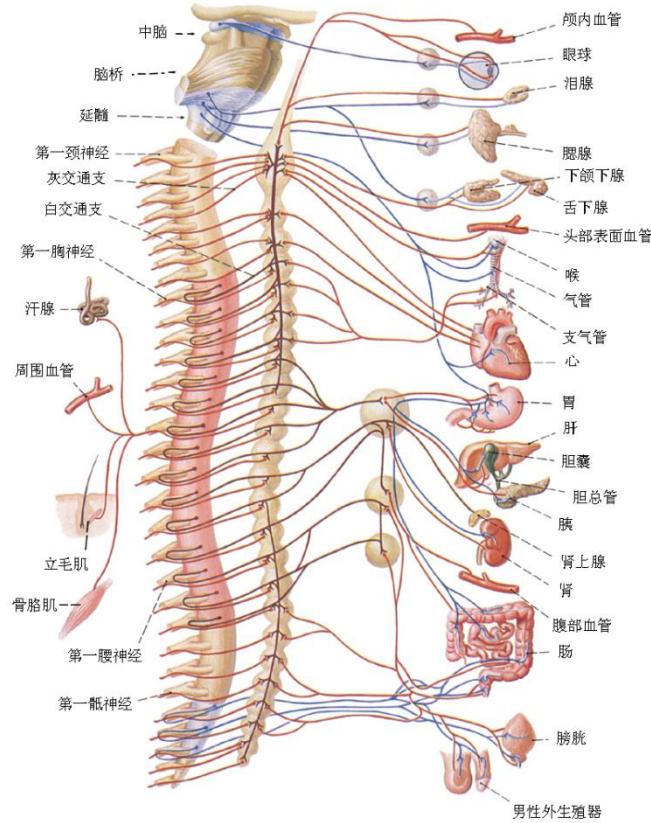
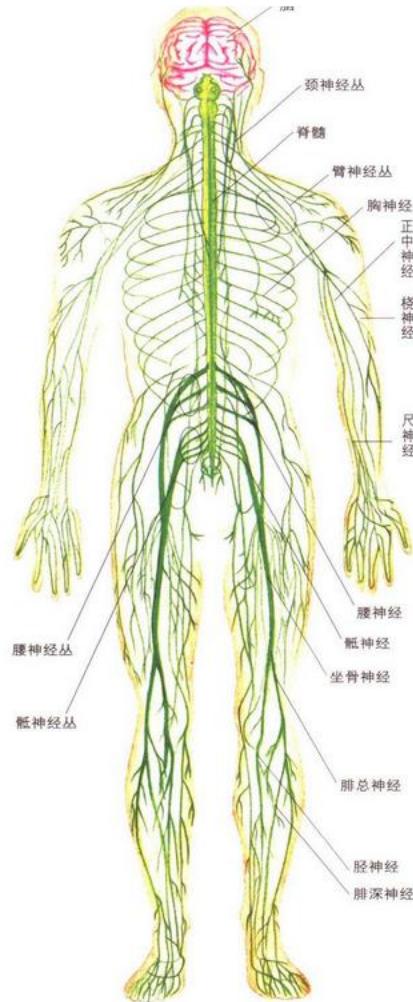
DATAGURU专业数据分析社区

# 受体与离子通道



# 化学过程



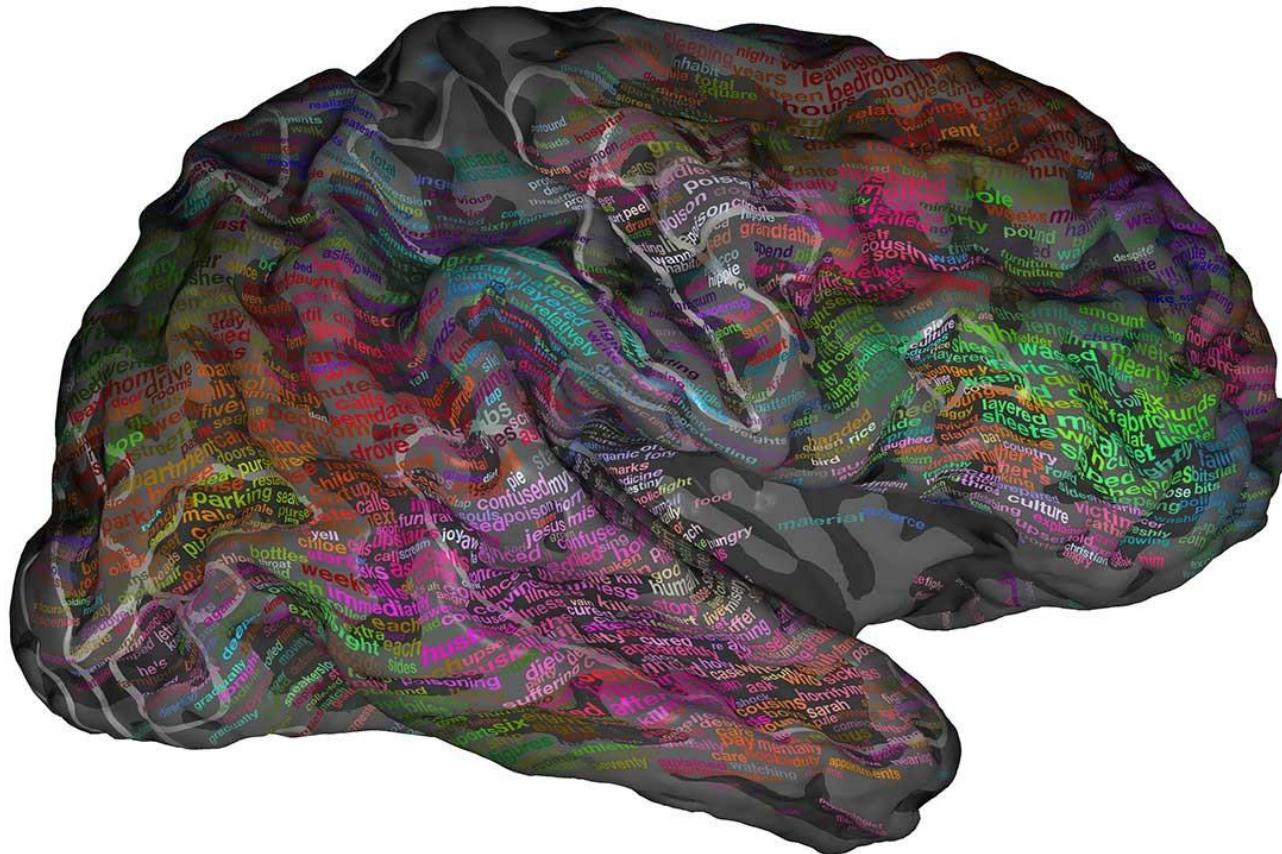


自主神经系统概观

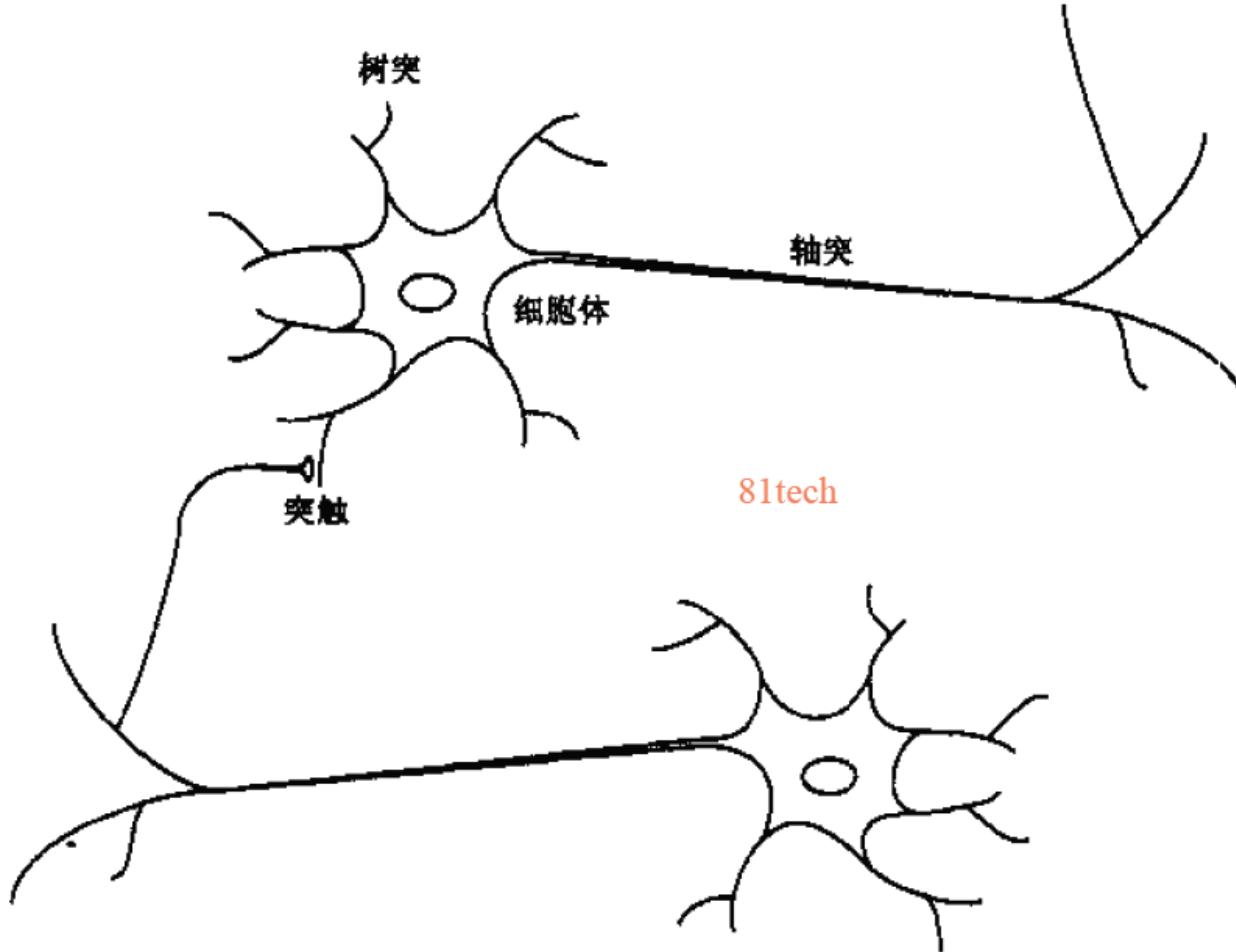
# 大脑地图



- <http://www.51hei.com/bbs/dpj-49269-1.html>

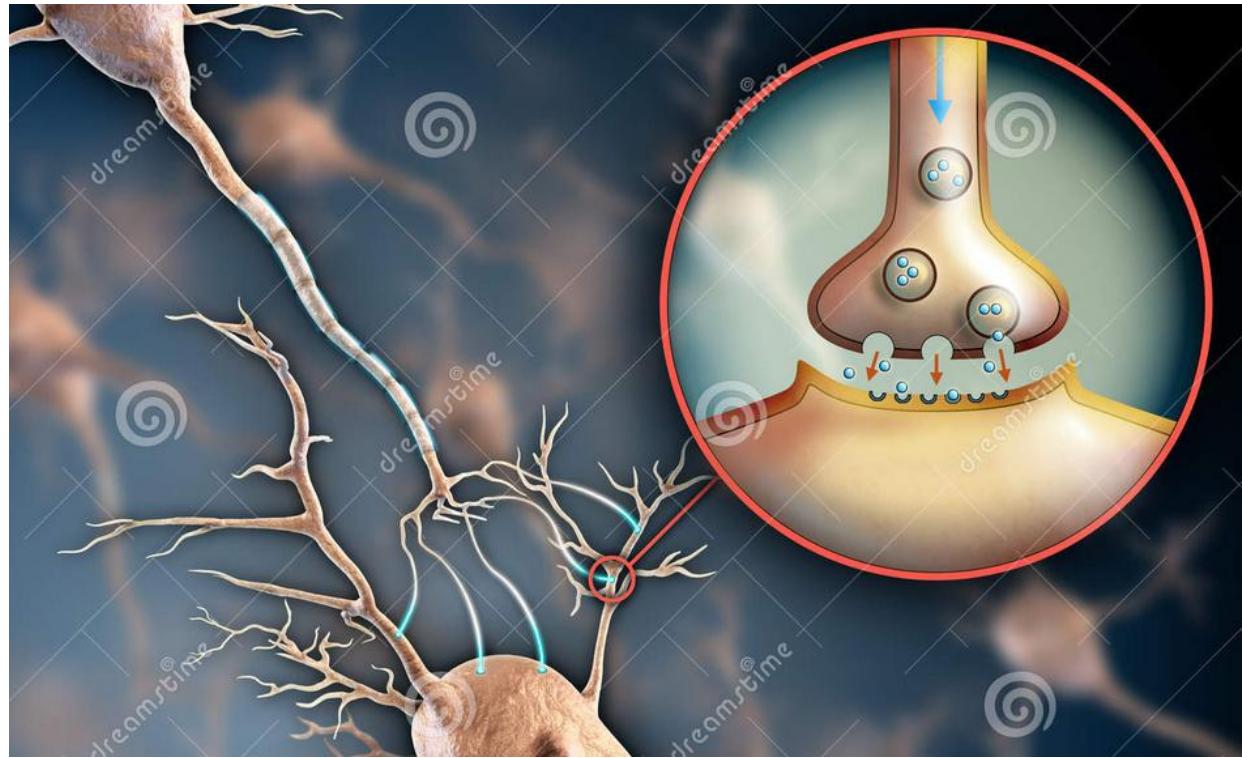


# 神经元简化原理图



# 学习与记忆

- 部分结构是与生俱来（非条件反射？）
- 大部分是经过后天学习建立连接（也可能消除连接）
- 记忆的本质是改变突触强度



# 人工神经网络（ANN）

- 这是迄今为止几乎最为成功的仿生学数学模型，是机器学习领域的热点，符合智能化机器的时代潮流
- 有统一的模型框架，很多算法问题可以划归为神经网络系统学习问题加以解决（例如 SVM）
- 容易硬件化，元器件化，高集成化，并行化，性能优异
- 神经网络设计具有较高技巧。同时它也是一个灰箱系统，容易掩盖某些业务背景细节
- 容易产生过度拟合

# 用途

## ■ 《神经网络设计》第4页

### 汽车

汽车自动导航系统，担保行为分析器

### 银行

支票和其他公文阅读器，信贷申请的评估器

### 国防

武器操纵，目标跟踪，目标辨识，面部识别、新型的传感器，声纳、雷达和图像信号处理(包括数据压缩、特征提取、噪声抑制、信号/图像的识别)

### 电子

代码序列预测，集成电路芯片布局，过程控制，芯片故障分析，机器视觉，语音综合，非线性建模

### 娱乐

动画，特技，市场预测

### 金融

不动产评估，借贷咨询，抵押审查，公司证券分级，投资交易程序，公司财务分析，通货价格预测

### 保险

政策应用评估，产品优化

### 制造

生产流程控制，产品设计和分析，81tech过程和机器诊断，实时微粒识别，可视质量监督系统，啤酒检测，焊接质量分析，纸张质量预测，计算机芯片质量分析，磨床运转分析，化工产品设计分析，机器性能分析，项目投标，计划和管理，化工流程系统动态建模

# 神经网络芯片



**ZISC78**是IBM公司和SILICon公司联合生产的具有自学习功能的径向基函数神经网络芯片。  
33MHz主频、CMOS型的VLSI芯片

**ZISC78**的特点：

- 内含78个神经元；
- 采用并行结构,运行速度与神经元数量无关；
- 支持RBF/KNN算法；
- 内部可分为若干独立子网络；
- 采用菊花链连接,扩展不受限制；
- 具有64字节宽度向量；
- L1或LSUP范数可用于距离计算；
- 具有同步/异步工作模式.

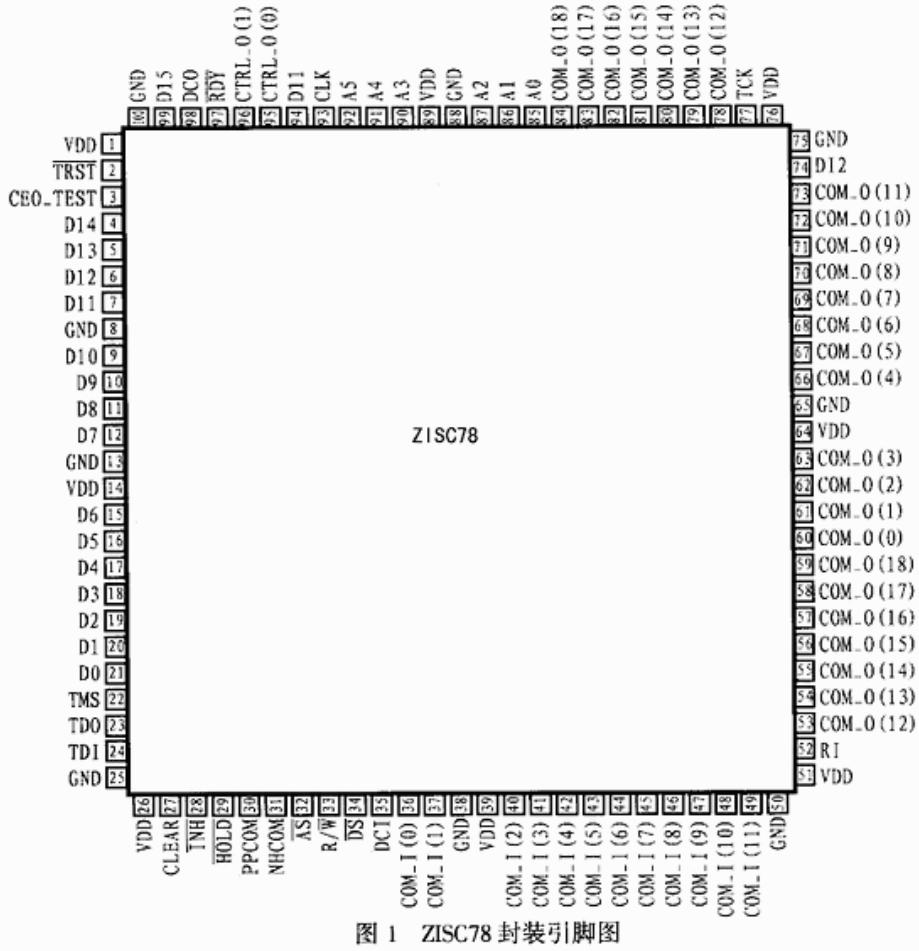
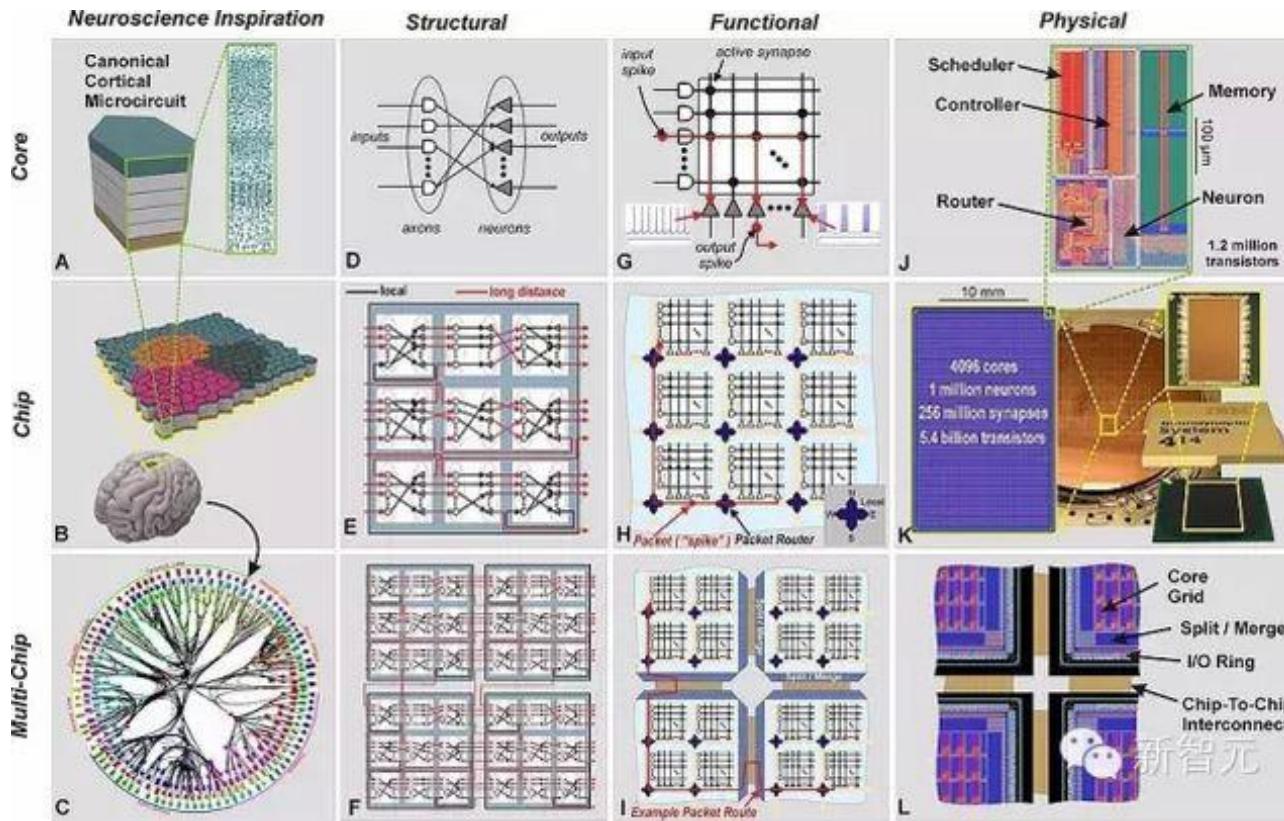


图 1 ZISC78 封装引脚图

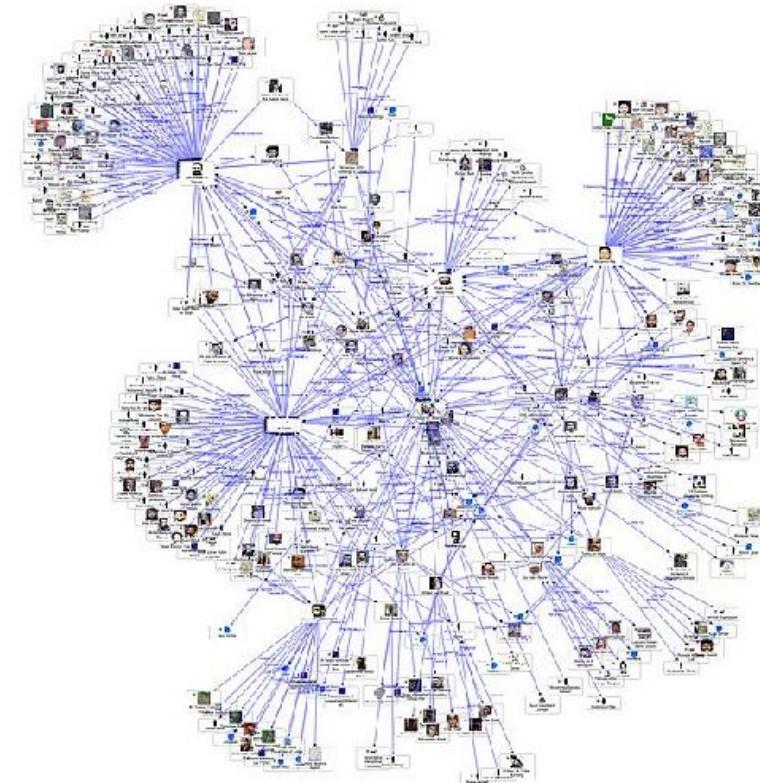
# 类脑超级计算机

- <http://it.dataguru.cn/article-9126-1.html>
- <http://aiera.baijia.baidu.com/article/415992>

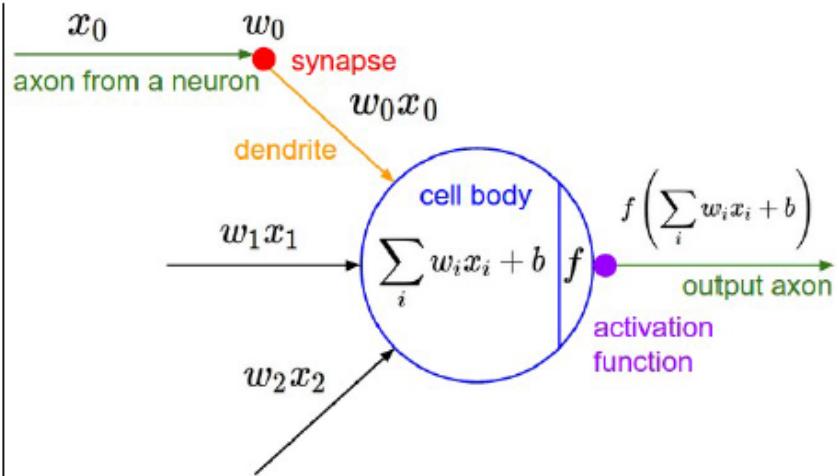
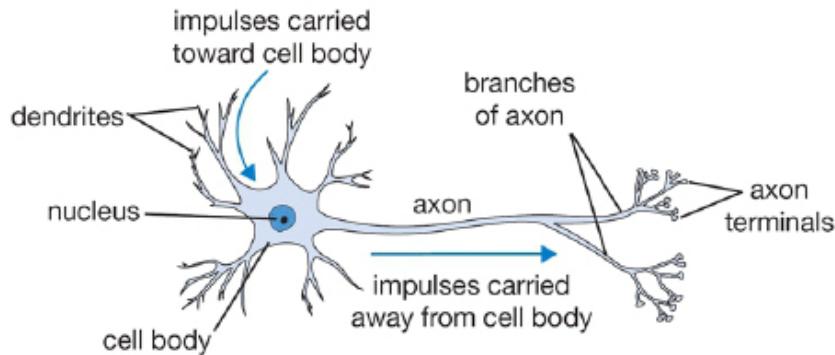


# 冯诺依曼型计算机 vs 神经网络

- 冯诺依曼型计算机 vs 神经网络
- 神经元的多寡以及信号传递的效率决定智力
- 语言，文字，社会化智力



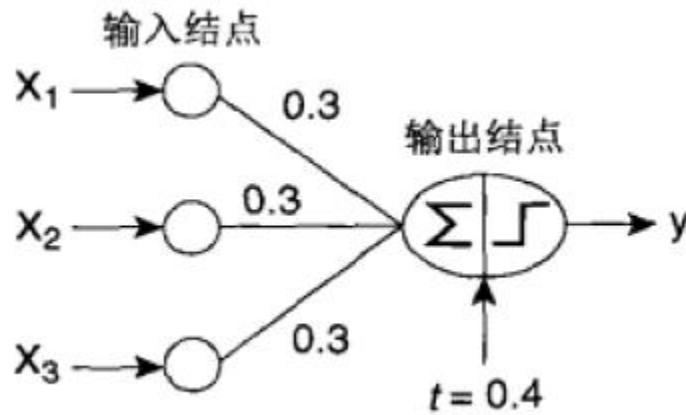
# 神经元仿生：单层感知器



A cartoon drawing of a biological neuron (left) and its mathematical model (right).

# 单层感知器

- 输入节点
- 输出节点
- 权向量
- 偏置因子
- 激活函数
- 学习率
- 单层感知器学习算法
- 单层感知器的局限



# 感知器学习规则

1958年，美国学者 Frank Rosenblatt 首次定义了一个具有单层计算单元的神经网络结构，称为感知器（Perceptron）。感知器的学习规则规定，学习信号等于神经元期望输出（教师信号）与实际输出之差

$$r = d_j - o_j \quad (2.15)$$

式中  $d_j$  为期望的输出， $o_j = f(\mathbf{W}_j^T \mathbf{X})$ 。感知器采用了与阈值转移函数类似的符号转移函数，其表达为

$$f(\mathbf{W}_j^T \mathbf{X}) = \text{sgn}(\mathbf{W}_j^T \mathbf{X}) = \begin{cases} 1, & \mathbf{W}_j^T \mathbf{X} \geq 0 \\ -1, & \mathbf{W}_j^T \mathbf{X} < 0 \end{cases} \quad (2.16)$$

因此，权值调整公式应为

$$\Delta \mathbf{W}_j = \eta [d_j - \text{sgn}(\mathbf{W}_j^T \mathbf{X})] \mathbf{X} \quad (2.17a)$$

$$\Delta w_{ji} = \eta [d_j - \text{sgn}(\mathbf{W}_j^T \mathbf{X})] x_i \quad i = 0, 1, \dots, n \quad (2.17b)$$

式中，当实际输出与期望值相同时，权值不需要调整；在有误差存在情况下，由于  $d_j, \text{sgn}(\mathbf{W}_j^T \mathbf{X}) \in \{-1, 1\}$ ，权值调整公式可简化为

$$\Delta \mathbf{W}_j = \pm 2 \eta \mathbf{X} \quad (2.17c)$$

感知器学习规则只适用于二进制神经元，初始权值可取任意值。

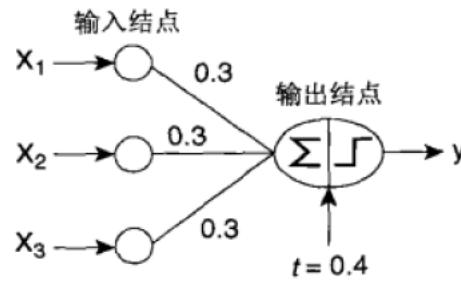
感知器学习规则代表一种有导师学习。由于感知器理论是研究其他神经网络的基础，该规则对于神经网络的有导师学习具有极为重要的意义。

# 例子

考虑图 5-14 中的图表。左边的表显示一个数据集，包含三个布尔变量( $x_1, x_2, x_3$ )和一个输出变量  $y$ ，当三个输入中至少有两个是 0 时， $y$  取 -1，而至少有两个大于 0 时， $y$  取 +1。

| $x_1$ | $x_2$ | $x_3$ | $y$ |
|-------|-------|-------|-----|
| 1     | 0     | 0     | -1  |
| 1     | 0     | 1     | 1   |
| 1     | 1     | 0     | 1   |
| 1     | 1     | 1     | 1   |
| 0     | 0     | 1     | -1  |
| 0     | 1     | 0     | -1  |
| 0     | 1     | 1     | 1   |
| 0     | 0     | 0     | -1  |

(a) 数据集



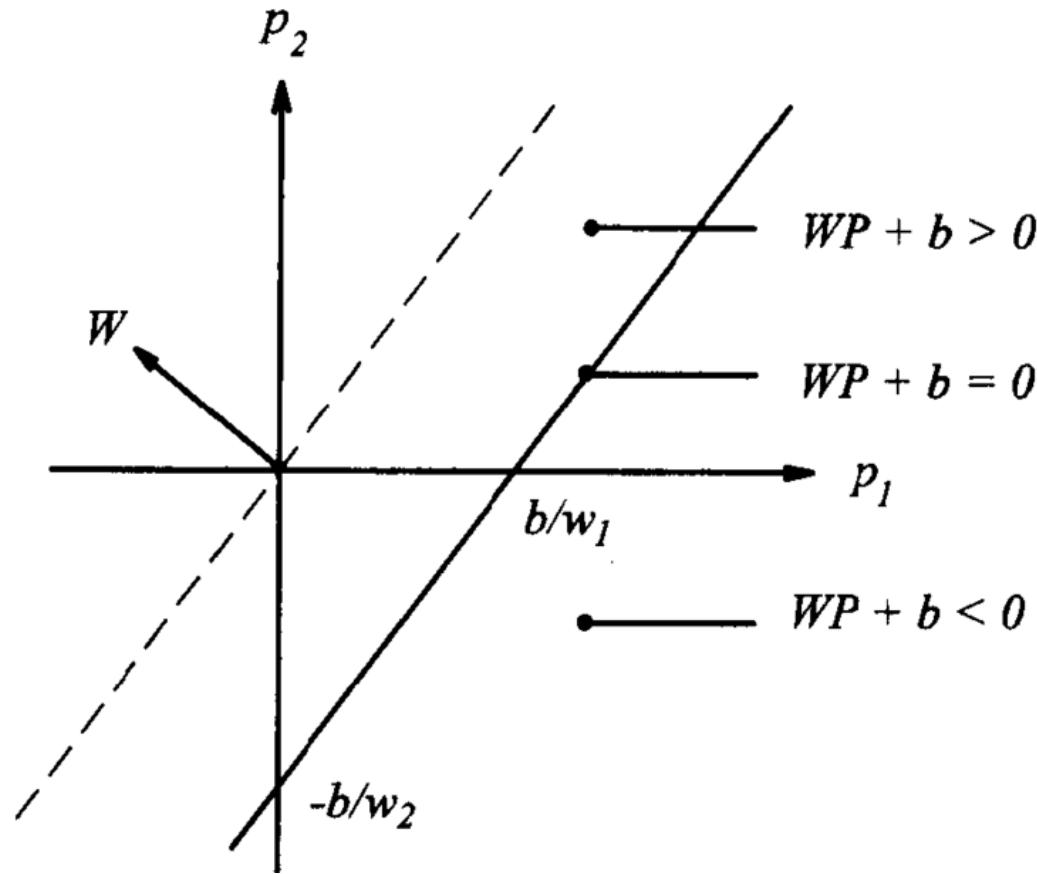
(b) 感知器

图 5-14 使用感知器模拟一个布尔函数

$$\hat{y} = \begin{cases} 1 & \text{如果 } 0.3x_1 + 0.3x_2 + 0.3x_3 - 0.4 > 0 \\ -1 & \text{如果 } 0.3x_1 + 0.3x_2 + 0.3x_3 - 0.4 < 0 \end{cases}$$

# 几何解释

## ■ 分离超平面



# 收敛条件

- 误差（实际输出与期望输出的差的绝对值）小于某个预先设定的较小的值
- 两次迭代之间的权值变化已经很小
- 设定最大迭代次数
- 单层感知器只对线性可分问题收敛

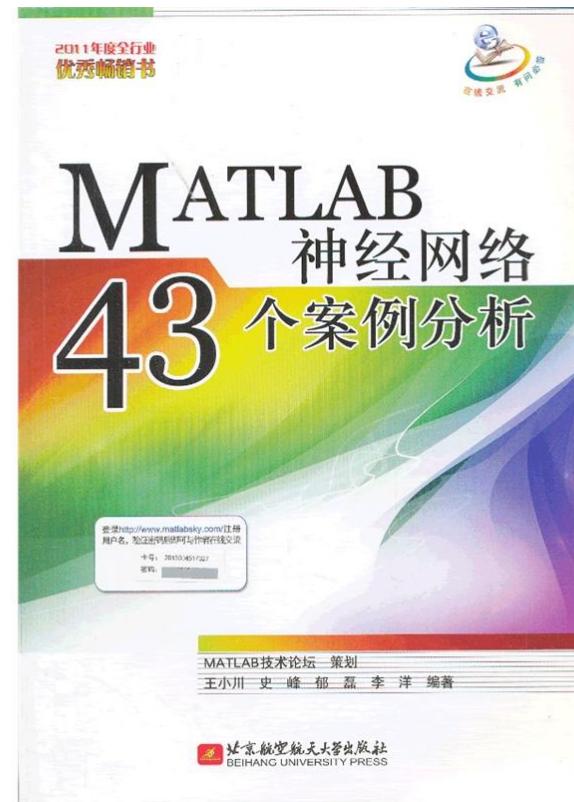
# 学习率的选择

- 学习率太大，容易造成权值计算不稳定
- 学习率太小，不能充分体现对权值的修正，迭代次数很多
- 比喻（显微镜对焦）

# 单层感知器的局限

- 无泛化能力
- 结构简单，激活函数只能是符号函数
- 只对线性可分问题收敛
- 如果存在离群点，则需要花费较多的训练时间

- MATLAB=matrix+laboratory，是由美国mathworks公司发布的主要面对科学计算、可视化以及交互式程序设计的高科技计算环境。
- MATLAB和Mathematica、Maple并称为三大数学软件。它在数学类科技应用软件中在数值计算方面首屈一指。MATLAB可以进行矩阵运算、绘制函数和数据、实现算法、创建用户界面、连接其他编程语言的程序等，主要应用于工程计算、控制设计、信号处理与通讯、图像处理、信号检测、金融建模设计与分析等领域。
- 具有功能完备强大的神经网络包



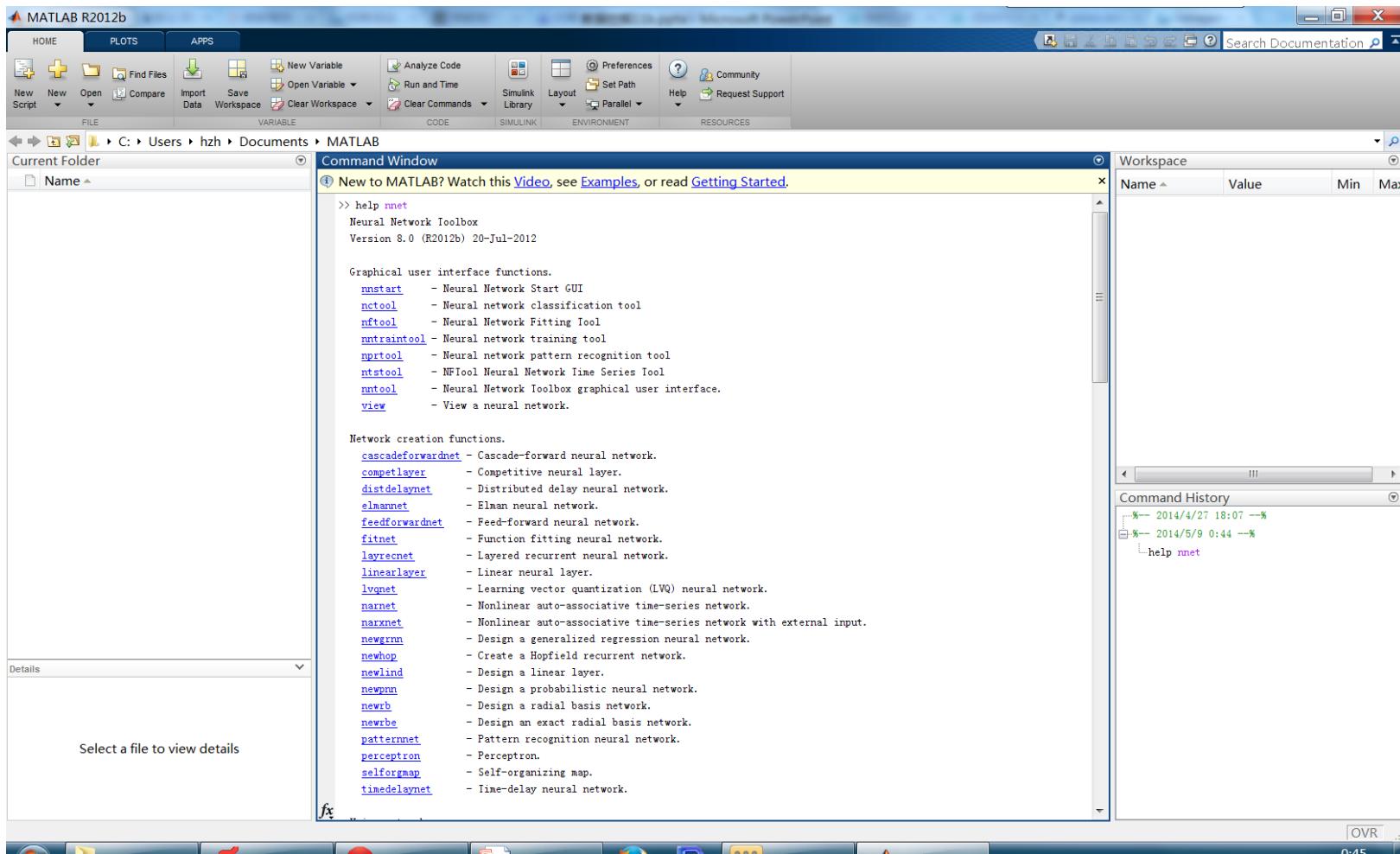
The screenshot shows the MathWorks website for the MATLAB product. The top navigation bar includes links for '中国' (China), '联系我们' (Contact Us), '如何购买' (How to Buy), '创建帐户' (Create Account), and '登录' (Log In). The main menu has categories like '产品和服务' (Products & Services), '解决方案' (Solutions), '教育' (Education), '支持' (Support), '用户中心' (User Center), '活动' (Events), and '公司' (Company). Below the menu, a breadcrumb trail shows '产品和服务 > MATLAB > 新特性'. A large banner for 'MATLAB' features a 3D surface plot. On the left, a floating window displays a portion of the MATLAB command history:

```
clc
plot(f,2*abs(Y(1:NFFT/2+1)))
Fs = 1000;
T = 1/Fs;
t =
t = (0:L-1)*T;
L = 1000;
clc
3x plot(f,2*abs(Y(1:NFFT/2+1)))
fx >> plot(f,2*abs(Y(1:NFFT/2+1)))
```

Below this window, a callout box says '弹出命令历史记录' (Pop-up Command History) and '使用弹出命令历史记录查看, 筛选和搜索最近使用的命令。' (View, filter, and search recently used commands using the pop-up command history.). A blue button labeled '观看视频' (Watch Video) is also visible.

On the right side of the page, there's a sidebar for 'R2014a' with options: '马上下载' (Download Now), '续订软件维护服务' (Renew Software Maintenance Services), '试用软件' (Try Software), '在线购买' (Buy Online), and a link to '» See release highlights for all products'.

# 神经网络工具箱



# 用命令建立数据

Command Window

① New to MATLAB? Watch this [Video](#), see [Examples](#), or read [Getting Started](#).

```
>>
>>
>> nnntool
>> nnntool
>> P=[1, 1, 1, 1, 0, 0, 0, 0; 0, 0, 0, 1, 1, 0, 1, 1, 0; 0, 1, 0, 1, 1, 0, 1, 0; 1, 0, 1, 1, 0, 1, 0, 1];
>> I=[-1, 1, 1, 1, -1, -1, 1, -1];
>> P

P =
1   1   1   1   0   0   0   0
0   0   1   1   0   1   1   0
0   1   0   1   1   0   1   0

>> I

I =
-1   1   1   1   -1   -1   1   -1
```

# 建立单层感应器

```
>> p=[0, 1; 0, 1; 0, 1]

p =
0     1
0     1
0     1

>> t=1

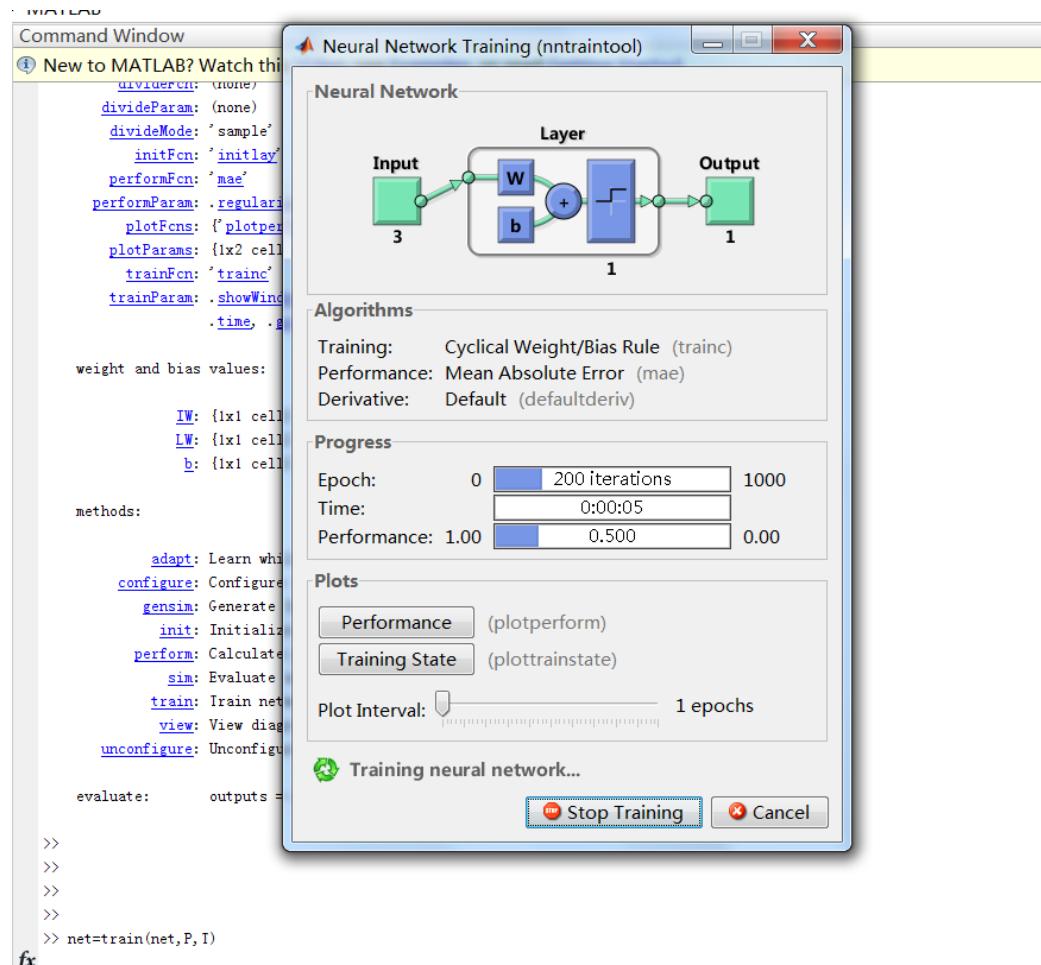
t =
1

>> net=newp(p,t)

net =
Neural Network

    name: 'Custom Neural Network'
    efficiency: .cacheDelayedInputs, .flattenTime,
                  .memoryReduction
    userdata: (your custom info)
```

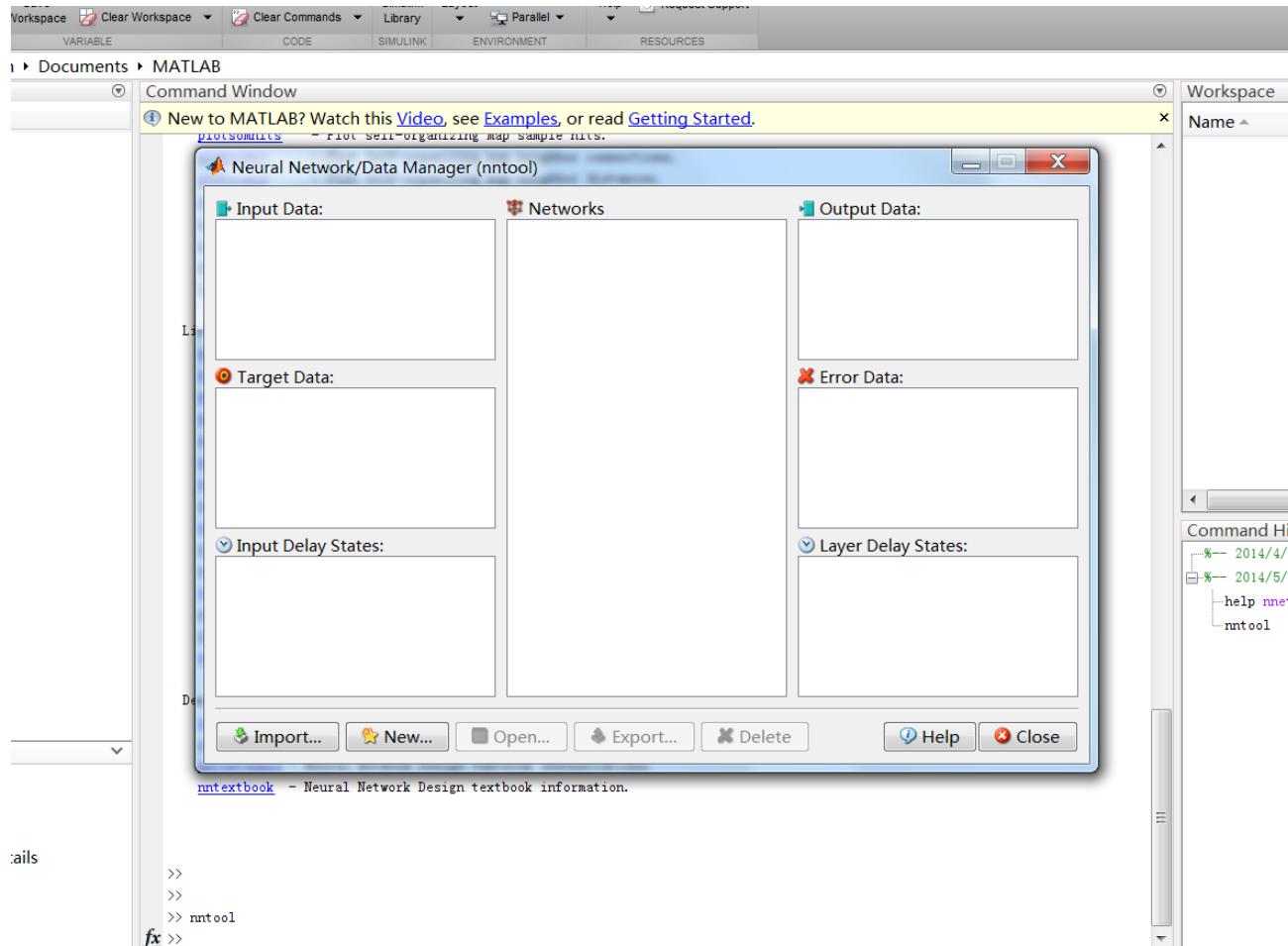
# 训练



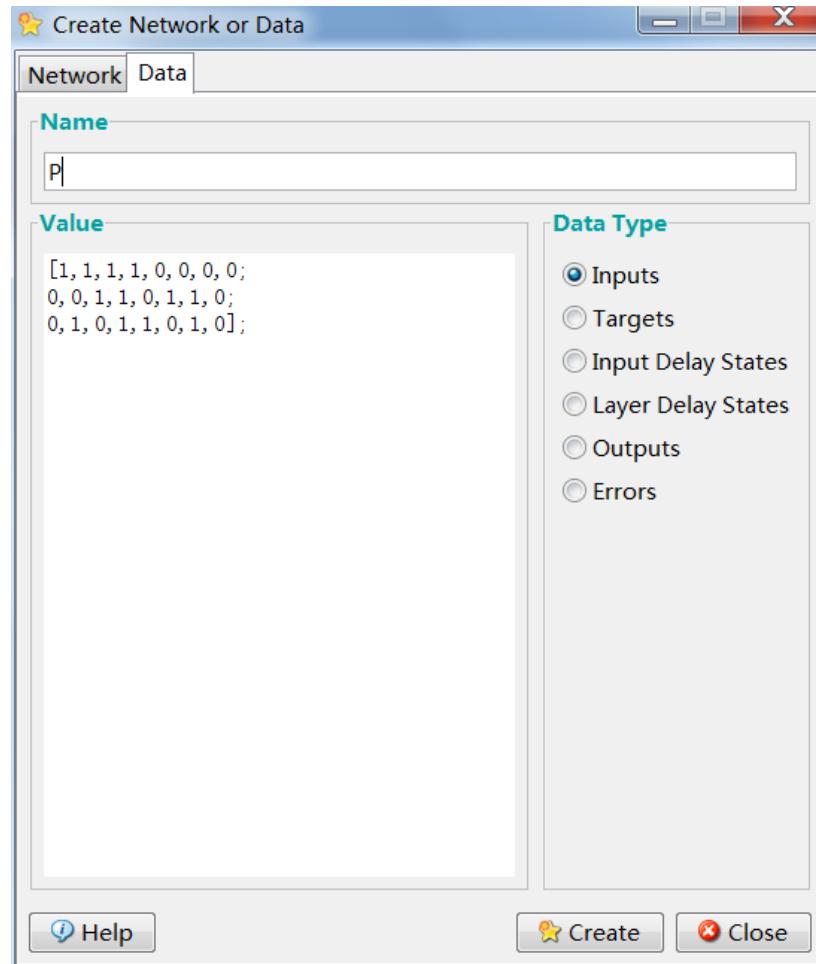
# 测试

```
>> newP=[0, 0, 1]';  
>> newI=sim(net, newP)  
  
newI =  
  
-1
```

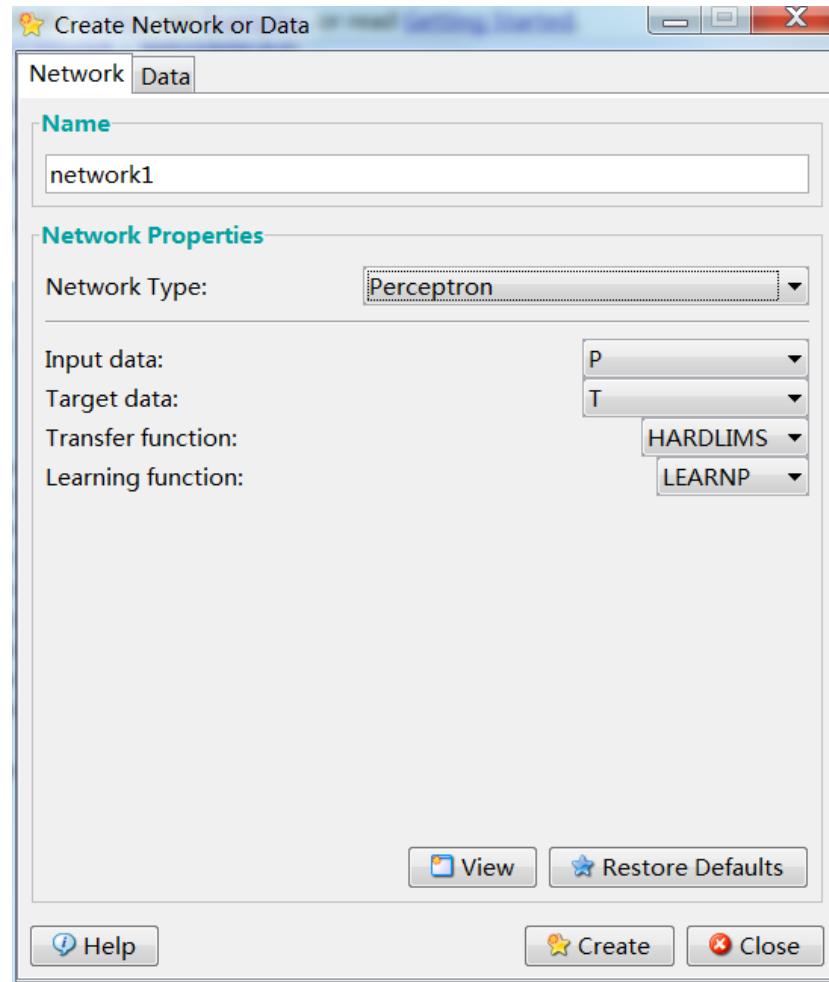
# 图形界面操作 : nntool



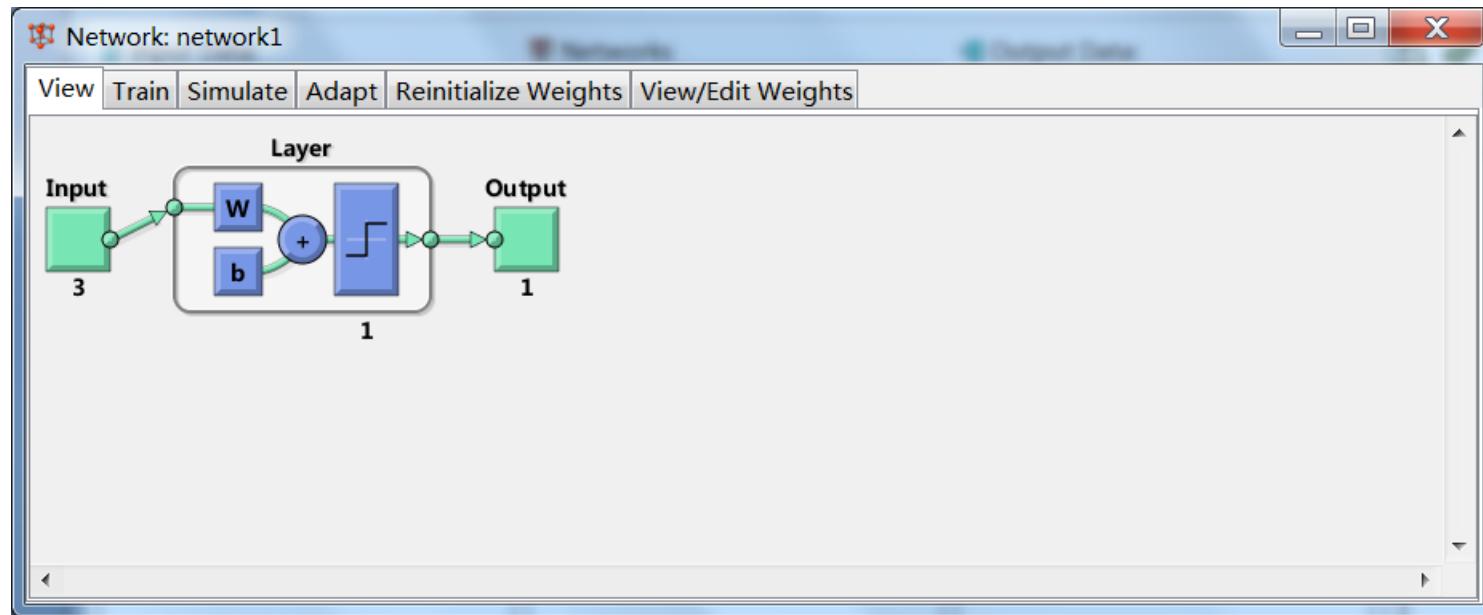
# 输入数据



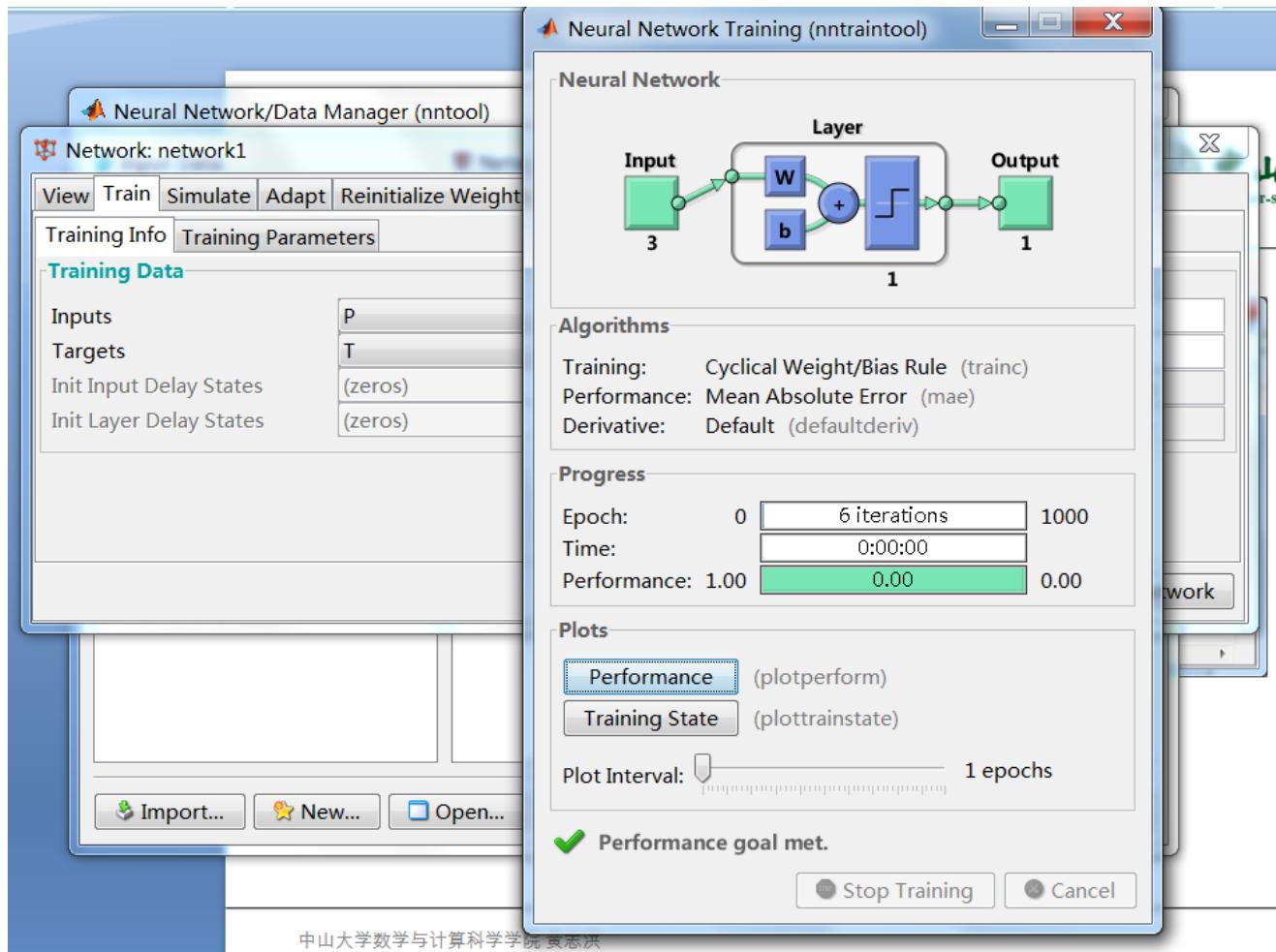
# 创建神经网络



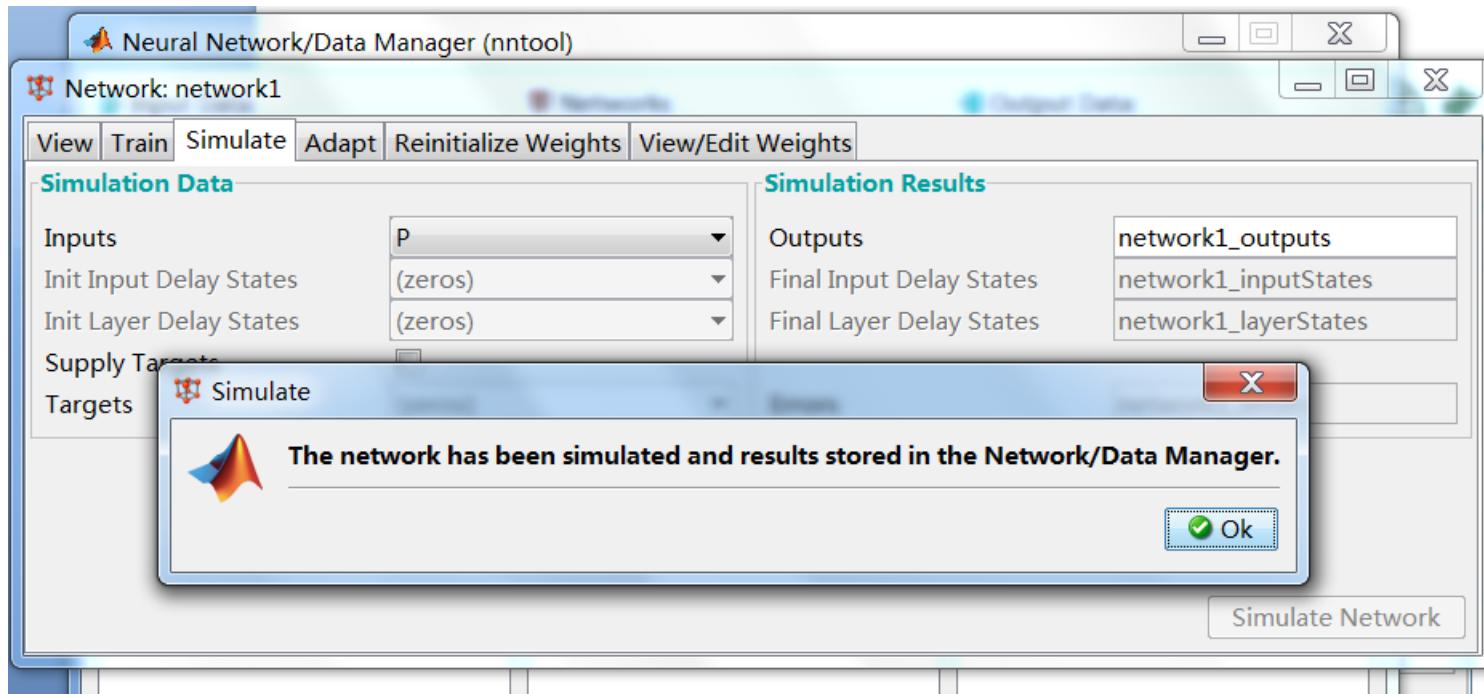
# 打开网络



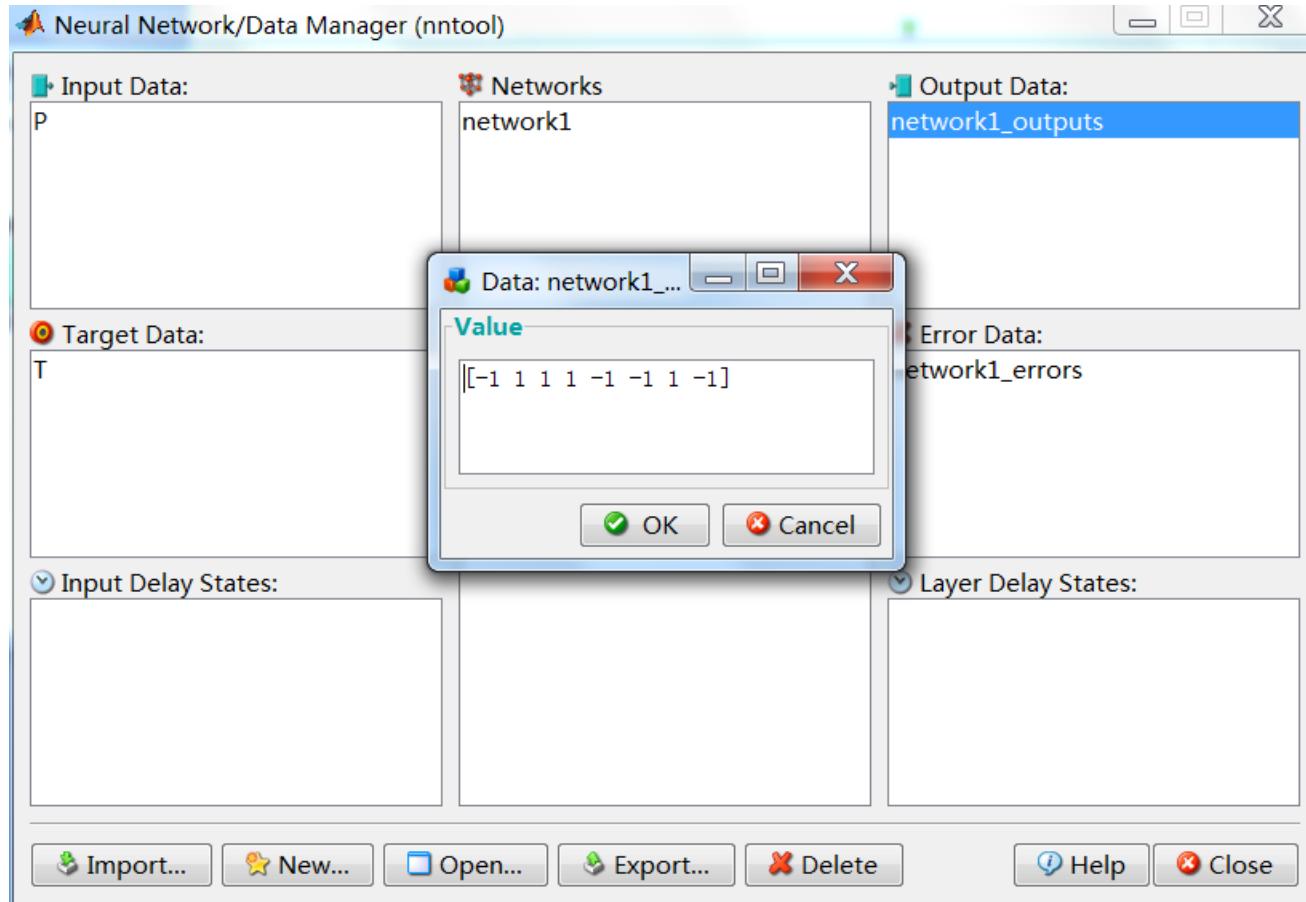
# 训练



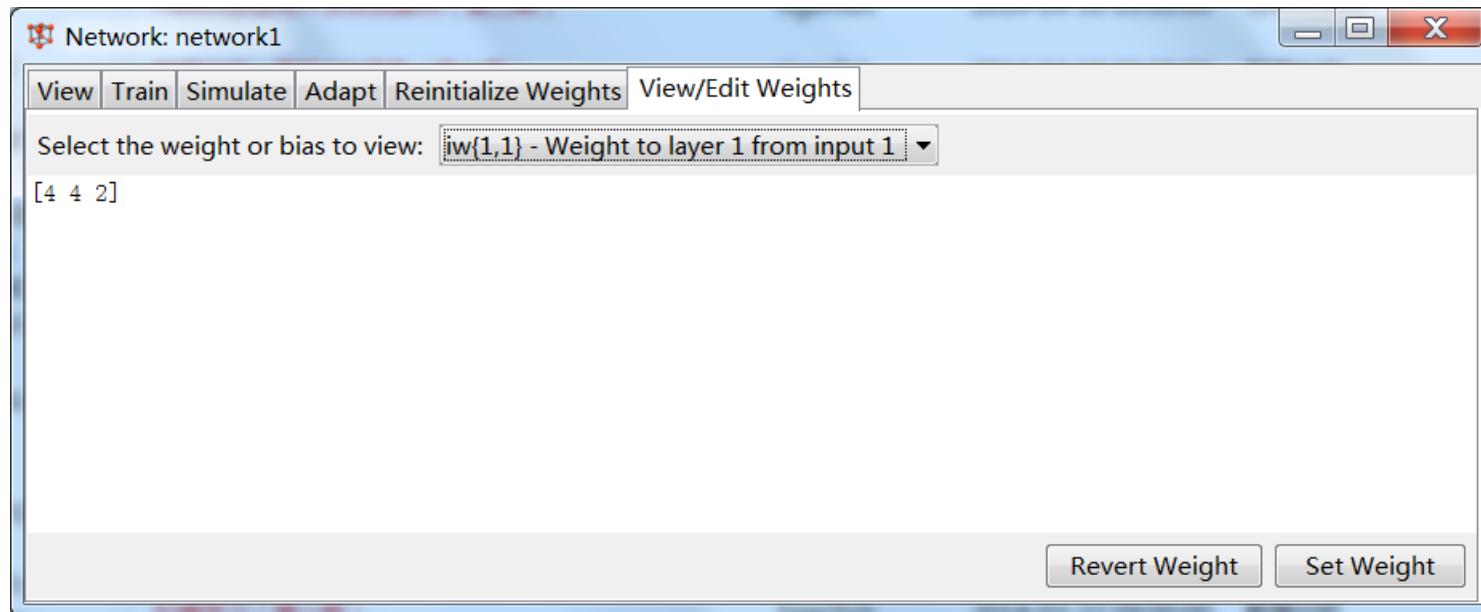
# 测试



# 观看测试结果



# 观察训练出来的权重和偏移



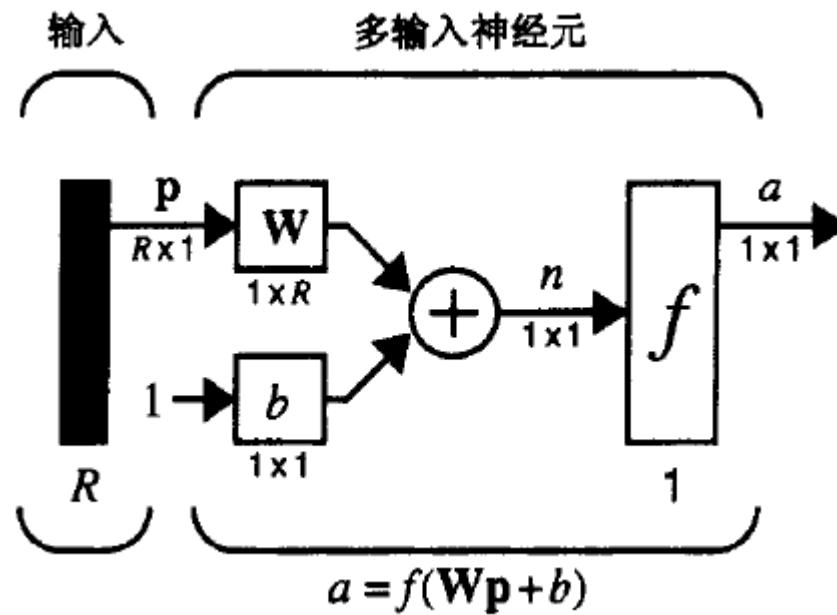
# 一些探讨

- Matlab演示中各个参数的意义
- 齐次化训练偏置值的技巧
- 有哪些激活函数？什么函数可以作为激活函数？
- 单层感知机不需要有样本概率分布的前提假设
- 更换单层感知器的激活函数是否可以获得解决非线性的能力？



# 怎样训练偏置值

- 把偏置值也看成权重之一，对应的变量输入值恒为1



# 性能评估函数

- MAE，平均绝对误差（误差的绝对值的平均值）
- MSE，均方误差（误差平方的平均值）
- SSE，误差平方和

# 各种激活函数

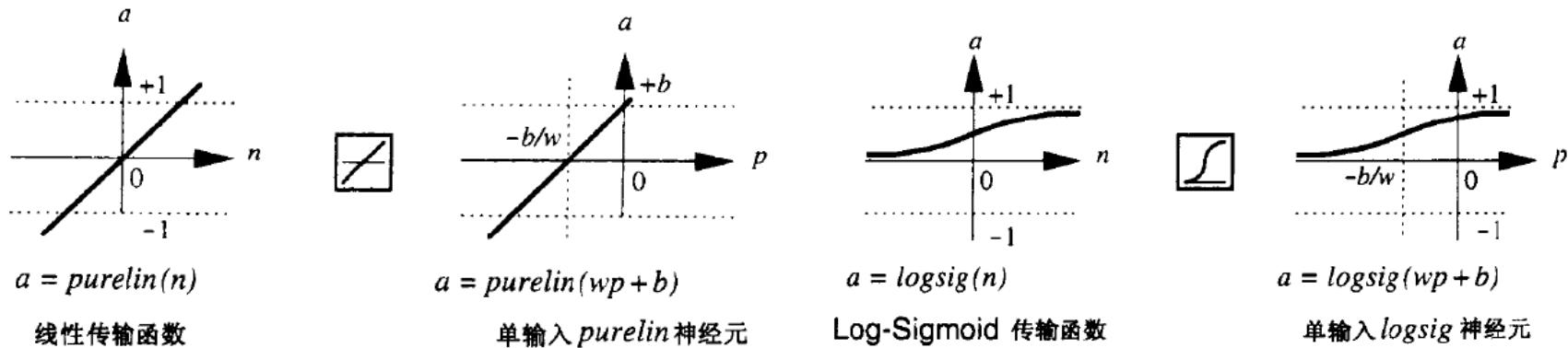


图 2-3 线性传输函数

图 2-4 对数-S 形传输函数

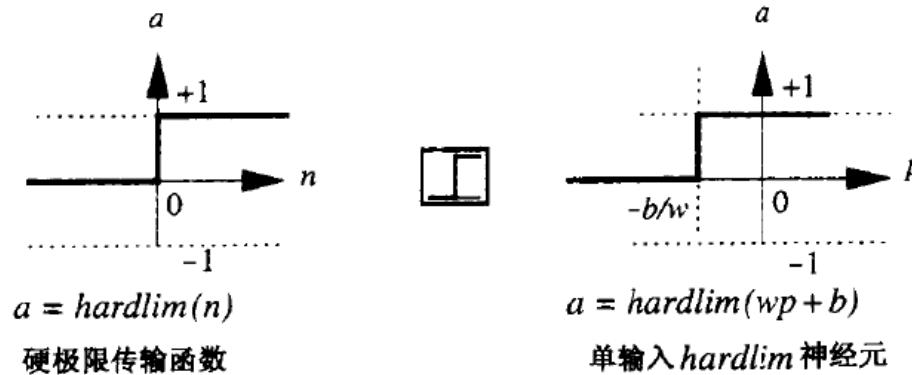


图 2-2 硬极限传输函数

# 各种激活函数

表 2-1 传输函数

| 名称       | 输入/输出关系   | 图标  | MATLAB 函数 |
|----------|---|---|-----------|
| 硬极限函数    | $a = 0, n < 0$<br>$a = 1, n \geq 0$                             |    | hardlim   |
| 对称硬极限函数  | $a = -1, n < 0$<br>$a = +1, n \geq 0$                           |    | hardlims  |
| 线性函数     | $a = n$   |    | purelin   |
| 饱和线性函数   | $a = 0, n < 0$<br>$a = n, 0 \leq n \leq 1$<br>$a = 1, n > 1$    |    | satlin    |
| 对称饱和线性函数 | $a = -1, n < -1$<br>$a = n, -1 \leq n \leq 1$<br>$a = 1, n > 1$ |  | satlins   |

# 各种激活函数

|           |   |  |        |
|-----------|---|--|--------|
| 对称饱和线性函数  | $a = -1, \quad n < -1$<br>$a = n, \quad -1 \leq n \leq 1$<br>$a = 1, \quad n > 1$ |  | satlin |
| 对数-S形函数   | $a = \frac{1}{1 + e^{-n}}$  |  | logsig |
| 双曲正切 S形函数 | $a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$   |  | tansig |
| 正线性函数     | $a = 0, \quad n < 0$<br>$a = n, \quad n \geq 0$                                   |  | poslin |
| 竞争函数      | $a = 1, \text{ 具有最大 } n \text{ 的神经元}$<br>$a = 0, \text{ 所有其他神经元}$                 |  | compet |

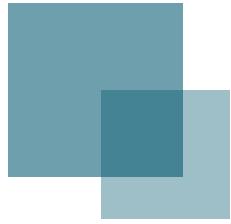
# 较新的激活函数

■ <http://www.zhunniao.com/values/33262.html>

| Name  | Plot  | Equation   | Derivative  |
|---|---|--|---|
| Identity                                      |    | $f(x) = x$   | $f'(x) = 1$   |
| Binary step                                   |    | $f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$               | $f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$             |
| Logistic (a.k.a Soft step)                    |    | $f(x) = \frac{1}{1 + e^{-x}}$  | $f'(x) = f(x)(1 - f(x))$  |
| TanH  |    | $f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1$  | $f'(x) = 1 - f(x)^2$  |
| ArcTan  |    | $f(x) = \tan^{-1}(x)$  | $f'(x) = \frac{1}{x^2 + 1}$   |
| Rectified Linear Unit (ReLU) [2]              |    | $f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$               | $f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$             |
| Parameteric Rectified Linear Unit (PReLU) [2] |   | $f(x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$        | $f'(x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$        |
| Exponential Linear Unit (ELU) [3]             |  | $f(x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$ | $f'(x) = \begin{cases} f(x) + \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$ |
| SoftPlus                                      |  | $f(x) = \log_e(1 + e^x)$   | $f'(x) = \frac{1}{1 + e^{-x}}$  |

# 炼数成金逆向收费式网络课程

- Dataguru ( 炼数成金 ) 是专业数据分析网站 , 提供教育 , 媒体 , 内容 , 社区 , 出版 , 数据分析业务等服务。我们的课程采用新兴的互联网教育形式 , 独创地发展了逆向收费式网络培训课程模式。既继承传统教育重学习氛围 , 重竞争压力的特点 , 同时又发挥互联网的威力打破时空限制 , 把天南地北志同道合的朋友组织在一起交流学习 , 使到原先孤立的学习个体组合成有组织的探索力量。并且把原先动辄成千上万的学习成本 , 直线下降至百元范围 , 造福大众。我们的目标是 : 低成本传播高价值知识 , 构架中国第一的网上知识流转阵地。
- 关于逆向收费式网络的详情 , 请看我们的培训网站 <http://edu.dataguru.cn>



# Thanks

FAQ时间