

机器学习 第9周

【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>

- 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。
- 我们课程里关于人工神经网络的部分是为了将来深度学习的有关课程作知识准备
- 炼数成金深度学习专题：<http://f.dataguru.cn/topic-deep-learning.html>

- 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来说已经变的非常重要，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。

- 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

Geoffrey Hinton

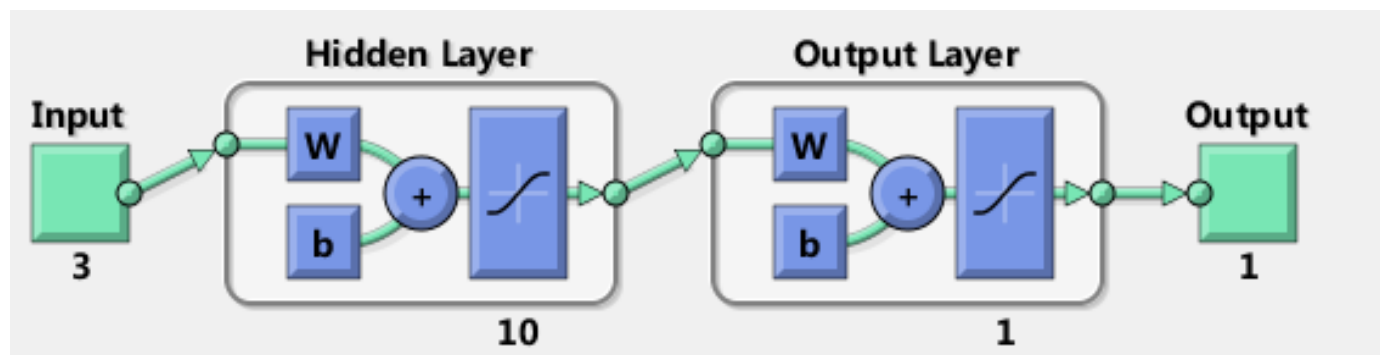
- <http://www.cs.toronto.edu/~hinton/>
- 多伦多大学计算机科学系教授，机器学习与人工智能顶级专家，诸多神经网络原创性研究的发起者和关键技术推手
- 目前领导Google Brain项目



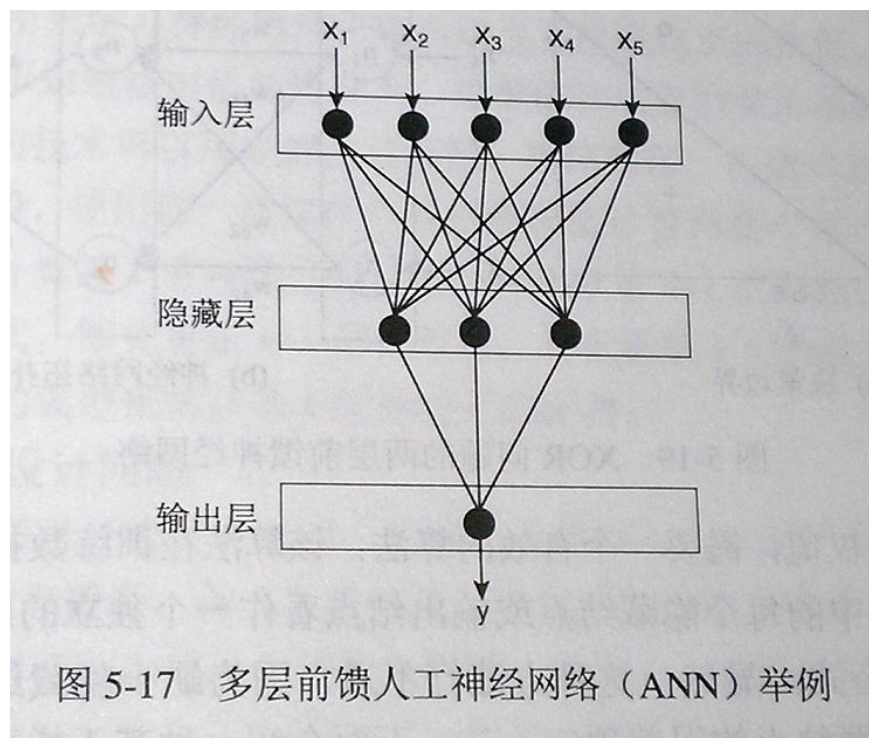
- Yann LeCun是人工智能尤其是深度学习领域最知名的学者之一，1960年生于法国巴黎。在居里大学读博士期间，提出过后向传播学习算法。
- 在多伦多大学随Hinton读博士后。1988-2002年加盟贝尔实验室，期间研发了卷积神经网络（Convolutional Neural Networks）与曾广泛用于手写识别和OCR的图变换网络方法。当然，图形压缩技术DjVu可能更为大家所熟知。
- 2003年加入纽约大学，从事广度与深度兼具的各类研究，涉及机器学习、计算机视觉、移动机器人和计算神经学。他甚至还开发了一种开源的面向对象编程语言Lush，比Matlab功能还要强，当然他也是一位Lisp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。

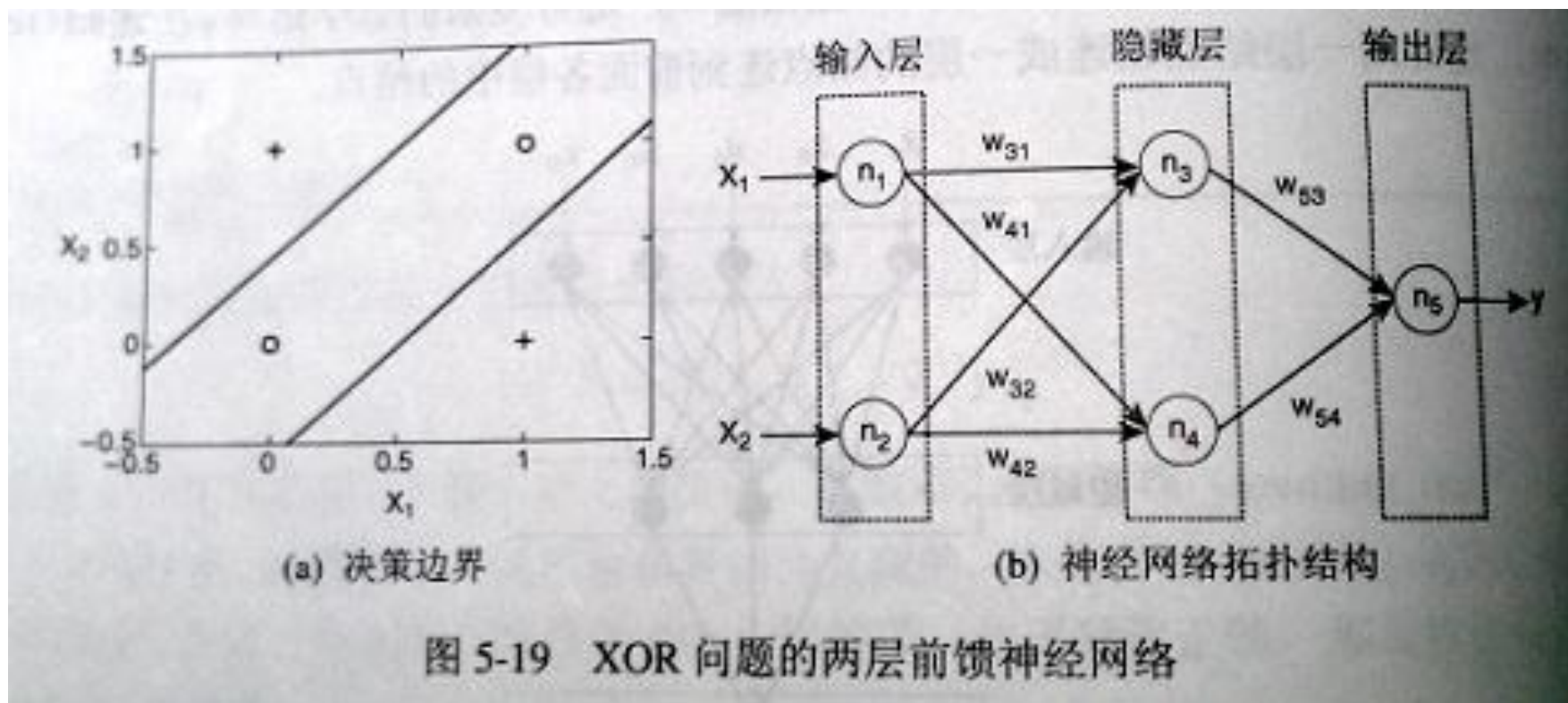


- error Back Propagation , 基于误差反向传播算法
- 1986年由Rumelhart和McClland为首的科学家小组提出
- 多层前馈网络
- 目前应用最广泛的神经网络模型之一
- 多层, 可以调整的权值很多, 可以使用多种激活函数
- 能学习和存贮大量的输入-输出模式映射关系, 而无需事前揭示描述这种映射关系的数学方程
- 学习规则使用最速下降法, 通过反向传播来不断调整网络的权值和阈值, 使网络的误差平方和最小
- BP神经网络模型拓扑结构包括输入层 (input)、隐层(hidden layer)和输出层(output layer)



- 隐藏层与隐藏节点
- 前馈 —— 每一层的节点仅和下一层节点相连





学习规则：最速下降BP法

- 很早就有多层神经网络的想法，但苦于没有合适的学习算法
- 误差反向传播算法催熟了人工神经网络
- 基本思想：学习过程由信号的正向传播和误差的反向传播两个过程组成
- 《人工神经网络理论、设计及应用》第47页

学习规则：最速下降BP法

- 很早就有多层神经网络的想法，但苦于没有合适的学习算法
- 误差反向传播算法催熟了人工神经网络
- 基本思想：学习过程由信号的正向传播和误差的反向传播两个过程组成
- 《人工神经网络理论、设计及应用》第47页

BP神经网络实现图像压缩。依赖BP网络的非线性映射能力进行数据压缩

- JPEG压缩
- 基于小波变换的图像压缩算法
- 分形压缩编码
- 矢量量化压缩编码

采用BP神经网络对灰度图像进行压缩，在保证较好峰值信噪比（PSNR）的情况下，达到了较高的压缩比。

常见的文件压缩软件如WinZip、WinRAR等采用的是无损压缩，能够完全恢复原文件内容。多媒体信息具有信息量大、冗余信息多的特点，往往采用有损压缩技术。

1.根据大面积着色原理，图像必须在一定面积内存在相同或相似的颜色，对于人眼的观察来说才有意义，否则看到的只是杂乱无章的雪花。因此，图像中相邻像素间存在相似性，这样就产生了图像的预测编码。

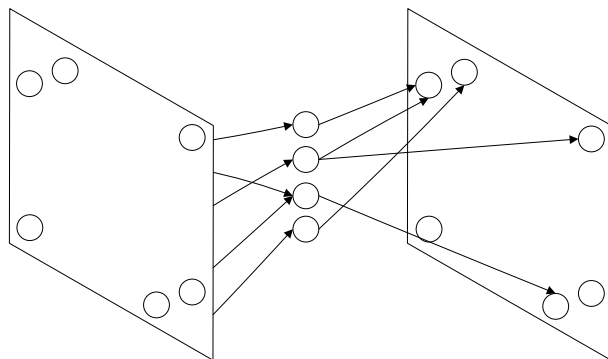
2.由于存在视觉的掩盖效应，因此人眼对于颜色细节往往并不敏感。图像信息上的微小损失往往是无法感知或可以接受的，这样就提供了广阔的压缩空间

3.数据都存在统计上的冗余，如在某一幅描绘海洋的图像中，蓝颜色出现的频率可能远高于红颜色，通过去除统计上的冗余同样可以实现压缩。

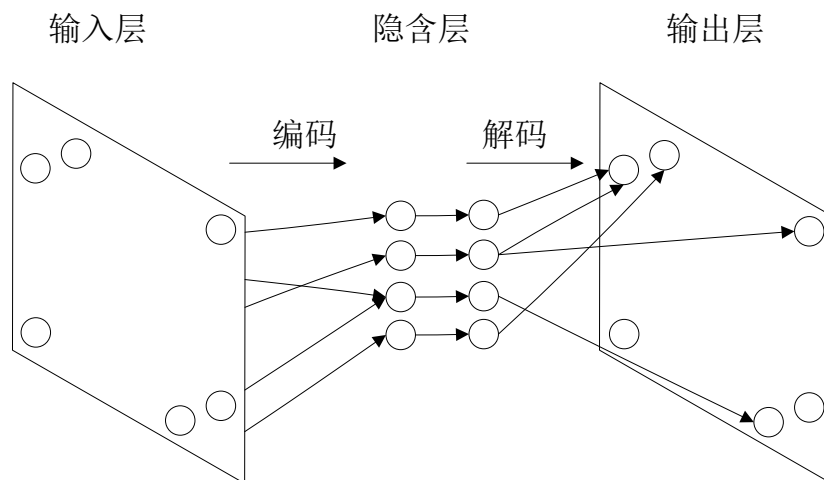
BP神经网络用于压缩：

只采用一个隐含层，因此整体构成了一个三层的网络。

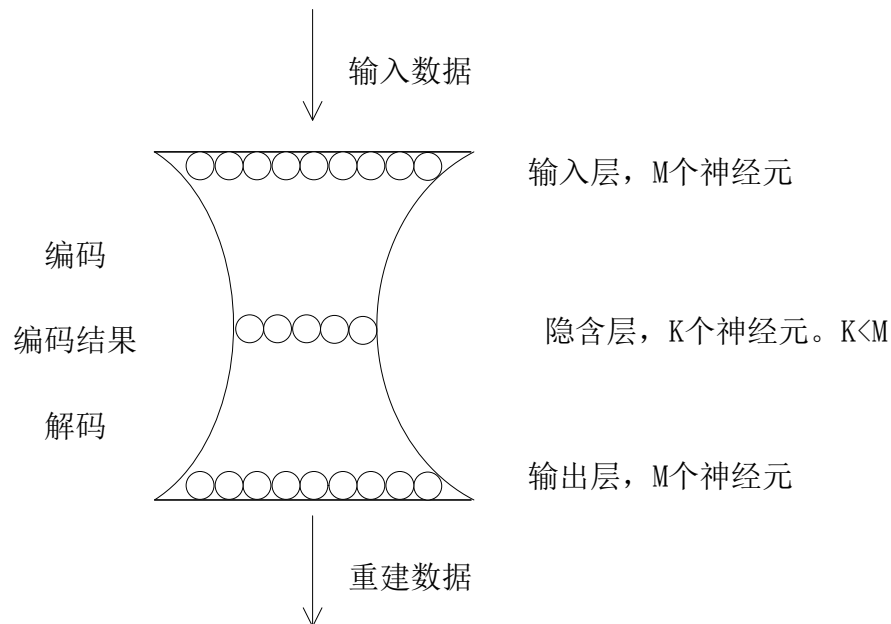
把一组输入模式通过少量的隐含层单元映射到一组输出模式，并使输出模式尽可能等于输入模式。因此，隐含层神经元的值和相应的权值向量可以输出一个与原输入模式相同的向量。当隐含层的神经元个数较少时，就意味着隐含层能用更少的数来表现输入模式，而这，实际上就是压缩。



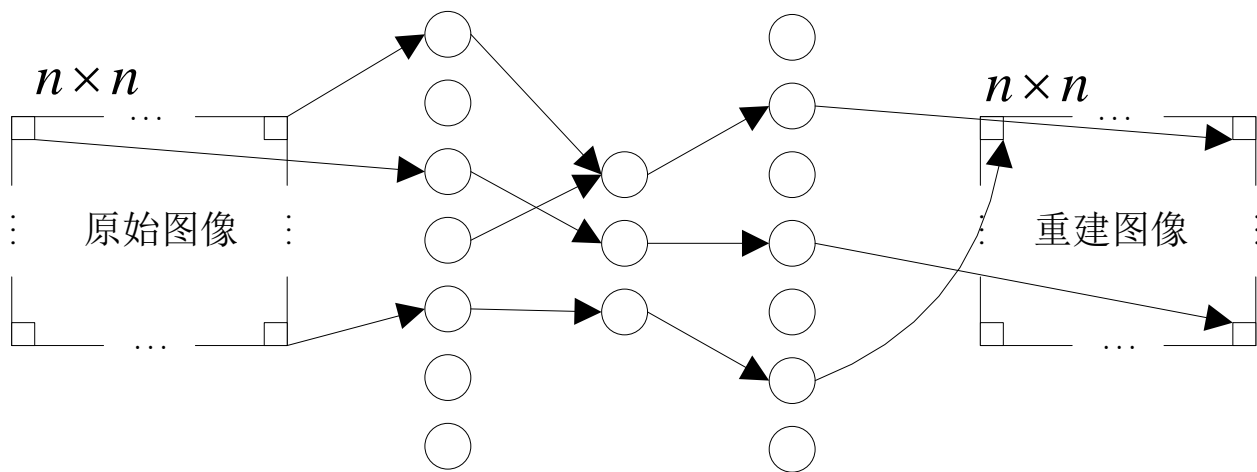
第一层为输入层，中间层为隐含层，网络的映射功能依赖隐含层实现。输入层到隐含层的变换相当于压缩的编码过程；而从隐含层到输出层的变换则相当于解码过程



输入层和输出层均由M个神经元组成，隐含层包含K个神经元。网络对数据进行压缩编码，保存在隐含层，再由隐含层输出到下一层

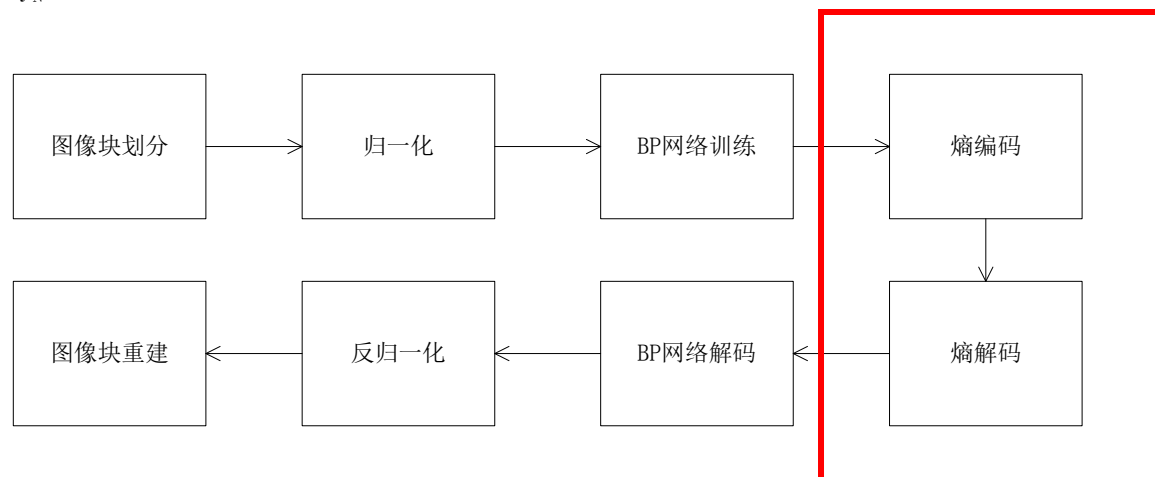
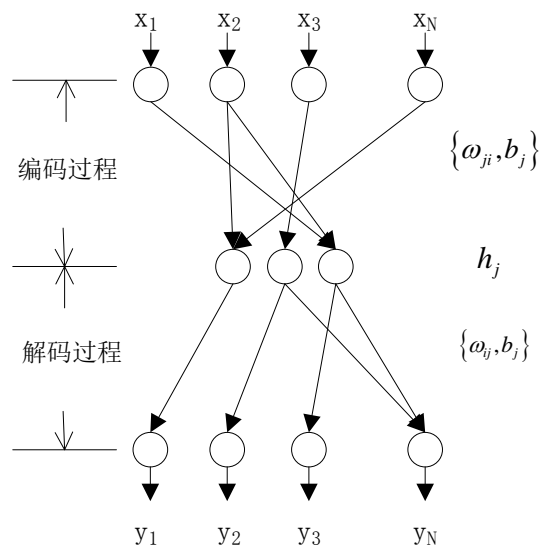


假设输入图像为像素 $N \times N$ 大小，被细分为多个 $n \times n$ 的图像块。如果将图像块中每一个像素点与一个输入或输出神经元相对应

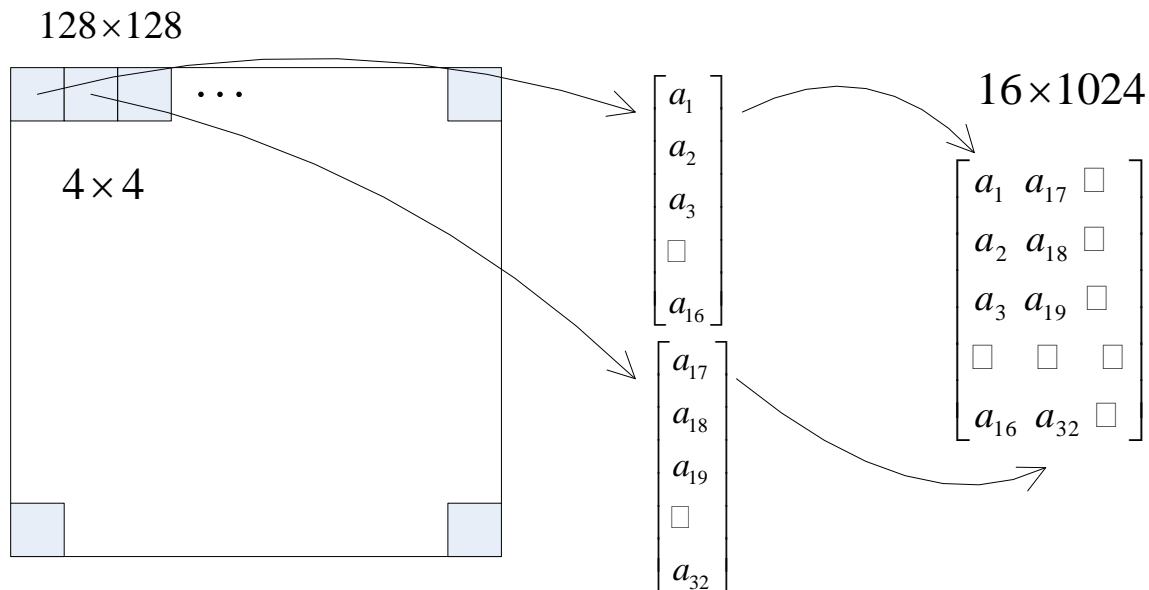


网络随机地抽取图像中各 $n \times n$ 图像块作为学习模式，使用反向传播算法进行学习，通过调整网络中神经元之间的连接权值，使训练集图像的重建误差 $E=f-g$ 的均值达到最小。

BP神经网络实现图像压缩



图像块的划分



归一化。神经网络的输入样本一般都需要进行归一化处理，这样更能保证性能的稳定性。归一化可以使用`mapminmax`函数进行，考虑到图像数据的特殊性，像素点灰度值为整数，且处于0-255之间，因此归一化处理统一将数据除以255即可。

采用MATLAB神经网络工具箱的**feedforwardnet**函数创建BP网络，并指定训练算法。为了达到较好效果，采用LM训练法。

保存结果。训练完成后，压缩的结果是每个输入模式对应的隐含层神经元向量的值，以及网络的权值和阈值。 使用**save**命令保存为MAT文件

BP神经网络实现图像压缩

解压。

使用load命令加载数据文件

数据反归一化

重建

图像反归一化

图像块恢复

运行bp_imageCompress.m进行压缩

bp_imageRecon进行解压



■ ftp://ftp.ics.uci.edu/pub/machine-learning-databases



← → ↻ ftp://ftp.ics.uci.edu/pub/machine-learning-databases		
README	35.7 kB	07-6-25 上午12:00:00
SUMMARY-TABLE	3.9 kB	98-9-1 上午12:00:00
abalone/		96-12-3 上午12:00:00
abscisic-acid/		08-6-17 上午12:00:00
access-lists/		96-12-3 上午12:00:00
acute/		09-9-14 上午12:00:00
adult/		01-1-31 上午12:00:00
annealing/		96-12-3 上午12:00:00
anonymous/		98-11-30 上午12:00:00
arcene/		08-3-25 上午12:00:00
arrhythmia/		98-3-20 上午12:00:00
artificial-characters/		97-12-9 上午12:00:00
audiology/		96-12-3 上午12:00:00
auto-mpg/		96-12-3 上午12:00:00
autos/		96-12-3 上午12:00:00
badges/		96-12-3 上午12:00:00
bag-of-words/		09-2-2 上午12:00:00
balance-scale/		96-12-3 上午12:00:00
balloons/		96-12-3 上午12:00:00
blood-transfusion/		08-11-30 上午12:00:00
breast-cancer/		00-1-6 上午12:00:00
breast-cancer-wisconsin/		96-12-3 上午12:00:00
bridges/		96-12-3 上午12:00:00
car/		97-6-19 上午12:00:00
census-income/		96-12-3 上午12:00:00
character-trajectories/		08-11-30 上午12:00:00
chess/		96-12-3 上午12:00:00
chorales/		96-12-3 上午12:00:00
cmc/		99-3-4 上午12:00:00
communities/		09-9-14 上午12:00:00
concrete/		09-9-14 上午12:00:00
connect-4/		96-12-3 上午12:00:00
contacts/		96-12-3 上午12:00:00
covtype/		99-12-14 上午12:00:00

← → ↺ <ftp://ftp.ics.uci.edu/pub/machine-learning-databases/statlog/german/german.data>

```
A11 6 A34 A43 1169 A65 A75 4 A93 A101 4 A121 67 A143 A152 2 A173 1 A192 A201 1
A12 48 A32 A43 5951 A61 A73 2 A92 A101 2 A121 22 A143 A152 1 A173 1 A191 A201 2
A14 12 A34 A46 2096 A61 A74 2 A93 A101 3 A121 49 A143 A152 1 A172 2 A191 A201 1
A11 42 A32 A42 7882 A61 A74 2 A93 A103 4 A122 45 A143 A153 1 A173 2 A191 A201 1
A11 24 A33 A40 4870 A61 A73 3 A93 A101 4 A124 53 A143 A153 2 A173 2 A191 A201 2
A14 36 A32 A46 9055 A65 A73 2 A93 A101 4 A124 35 A143 A153 1 A172 2 A192 A201 1
A14 24 A32 A42 2835 A63 A75 3 A93 A101 4 A122 53 A143 A152 1 A173 1 A191 A201 1
A12 36 A32 A41 6948 A61 A73 2 A93 A101 2 A123 35 A143 A151 1 A174 1 A192 A201 1
A14 12 A32 A43 3059 A64 A74 2 A91 A101 4 A121 61 A143 A152 1 A172 1 A191 A201 1
A12 30 A34 A40 5234 A61 A71 4 A94 A101 2 A123 28 A143 A152 2 A174 1 A191 A201 2
A12 12 A32 A40 1295 A61 A72 3 A92 A101 1 A123 25 A143 A151 1 A173 1 A191 A201 2
A11 48 A32 A49 4308 A61 A72 3 A92 A101 4 A122 24 A143 A151 1 A173 1 A191 A201 2
A12 12 A32 A43 1567 A61 A73 1 A92 A101 1 A123 22 A143 A152 1 A173 1 A192 A201 1
A11 24 A34 A40 1199 A61 A75 4 A93 A101 4 A123 60 A143 A152 2 A172 1 A191 A201 2
A11 15 A32 A40 1403 A61 A73 2 A92 A101 4 A123 28 A143 A151 1 A173 1 A191 A201 1
A11 24 A32 A43 1282 A62 A73 4 A92 A101 2 A123 32 A143 A152 1 A172 1 A191 A201 2
A14 24 A34 A43 2424 A65 A75 4 A93 A101 4 A122 53 A143 A152 2 A173 1 A191 A201 1
A11 30 A30 A49 8072 A65 A72 2 A93 A101 3 A123 25 A141 A152 3 A173 1 A191 A201 1
A12 24 A32 A41 12579 A61 A75 4 A92 A101 2 A124 44 A143 A153 1 A174 1 A192 A201 2
A14 24 A32 A43 3430 A63 A75 3 A93 A101 2 A123 31 A143 A152 1 A173 2 A192 A201 1
A14 9 A34 A40 2134 A61 A73 4 A93 A101 4 A123 48 A143 A152 3 A173 1 A192 A201 1
A11 6 A32 A43 2647 A63 A73 2 A93 A101 3 A121 44 A143 A151 1 A173 2 A191 A201 1
A11 10 A34 A40 2241 A61 A72 1 A93 A101 3 A121 48 A143 A151 2 A172 2 A191 A202 1
A12 12 A34 A41 1804 A62 A72 3 A93 A101 4 A122 44 A143 A152 1 A173 1 A191 A201 1
A14 10 A34 A42 2069 A65 A73 2 A94 A101 1 A123 26 A143 A152 2 A173 1 A191 A202 1
A11 6 A32 A42 1374 A61 A73 1 A93 A101 2 A121 36 A141 A152 1 A172 1 A192 A201 1
A14 6 A30 A43 426 A61 A75 4 A94 A101 4 A123 39 A143 A152 1 A172 1 A191 A201 1
A13 12 A31 A43 409 A64 A73 3 A92 A101 3 A121 42 A143 A151 2 A173 1 A191 A201 1
A12 7 A32 A43 2415 A61 A73 3 A93 A103 2 A121 34 A143 A152 1 A173 1 A191 A201 1
A11 60 A33 A49 6836 A61 A75 3 A93 A101 4 A124 63 A143 A152 2 A173 1 A192 A201 2
A12 18 A32 A49 1913 A64 A72 3 A94 A101 3 A121 36 A141 A152 1 A173 1 A192 A201 1
A11 24 A32 A42 4020 A61 A73 2 A93 A101 2 A123 27 A142 A152 1 A173 1 A191 A201 1
A12 18 A32 A40 5866 A62 A73 2 A93 A101 2 A123 30 A143 A152 2 A173 1 A192 A201 1
A14 12 A34 A49 1264 A65 A75 4 A93 A101 4 A124 57 A143 A151 1 A172 1 A191 A201 1
A13 12 A32 A42 1474 A61 A72 4 A92 A101 1 A122 33 A141 A152 1 A174 1 A192 A201 1
A12 45 A34 A43 4746 A61 A72 4 A93 A101 2 A122 25 A143 A152 2 A172 1 A191 A201 2
A14 48 A34 A46 6110 A61 A73 1 A93 A101 3 A124 31 A141 A153 1 A173 1 A192 A201 1
A13 18 A32 A43 2100 A61 A73 4 A93 A102 2 A121 37 A142 A152 1 A173 1 A191 A201 2
A13 10 A32 A44 1225 A61 A73 2 A93 A101 2 A123 37 A143 A152 1 A173 1 A192 A201 1
A12 9 A32 A43 458 A61 A73 4 A93 A101 3 A121 24 A143 A152 1 A173 1 A191 A201 1
A14 30 A32 A43 2333 A63 A75 4 A93 A101 2 A123 30 A141 A152 1 A174 1 A191 A201 1
A12 12 A32 A43 1158 A63 A73 3 A91 A101 1 A123 26 A143 A152 1 A173 1 A192 A201 1
A12 18 A33 A45 6204 A61 A73 2 A93 A101 4 A121 44 A143 A152 1 A172 2 A192 A201 1
A11 30 A34 A41 6187 A62 A74 1 A94 A101 4 A123 24 A143 A151 2 A173 1 A191 A201 1
```

基于BP网络的个人信贷信用评估

对个人信用的评估方法主要分为定性评估和定量评估两种，前者主要根据信贷人员的主观判断，后者则根据个人客户的资料，利用评分卡和信用评分模型等工具进行分析。

本实例采用**BP**神经网络，以已知用户信息及信用情况为训练样本，学习得出一个抽象模型，然后对新样本进行评估，正确率稳定在**70%**以上

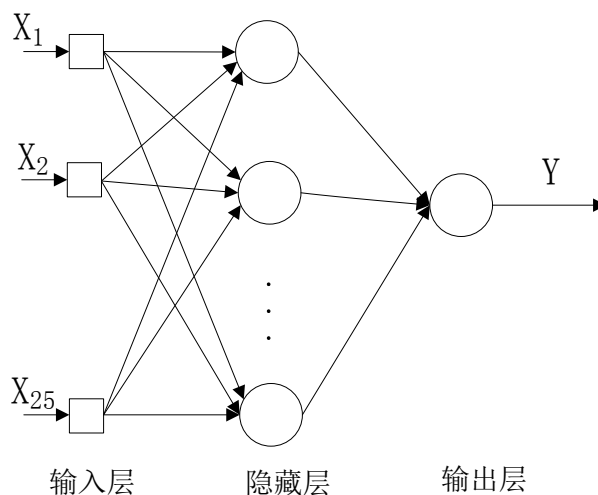
对所有客户做二分类，只区分**好**和**差**两种情况。数据采用德国信用数据库。德国信用数据库由**Hans Hofmann**教授整理，包含**1000**份客户资料，每位客户包含**20**条属性，并给出了信用好或差的标注。

基于BP网络的个人信贷信用评估

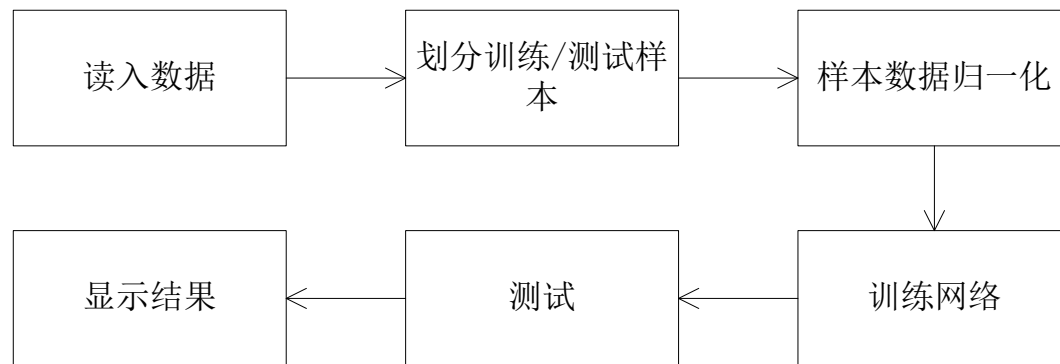
原始数据保存在`german.data`文件中：

经常账户状况、账户持续时间、贷款历史状况、贷款用途、贷款数额。

使用MATLAB实现一个三层的BP神经网络。由于每个个人用户拥有24个属性，因此输入层包含24个神经元节点。该问题为针对信用好/差的二分类问题，因此输出层只包含一个神经元。隐含层的神经元个数与网络性能有关，需要通过实验确定



基于BP网络的个人信贷信用评估



(1) 读入数据。

(2) 划分训练样本与测试样本。在全部1000份样本中，共有700份正例（信誉好），300份负例（信誉差）。划分时取前350份正例和前150份负例作为训练样本，后350份正例和后150份负例作为测试样本。

(3) 样本归一化。使用mapminmax函数对输入样本进行归一化

(4) 创建BP神经网络，并完成训练

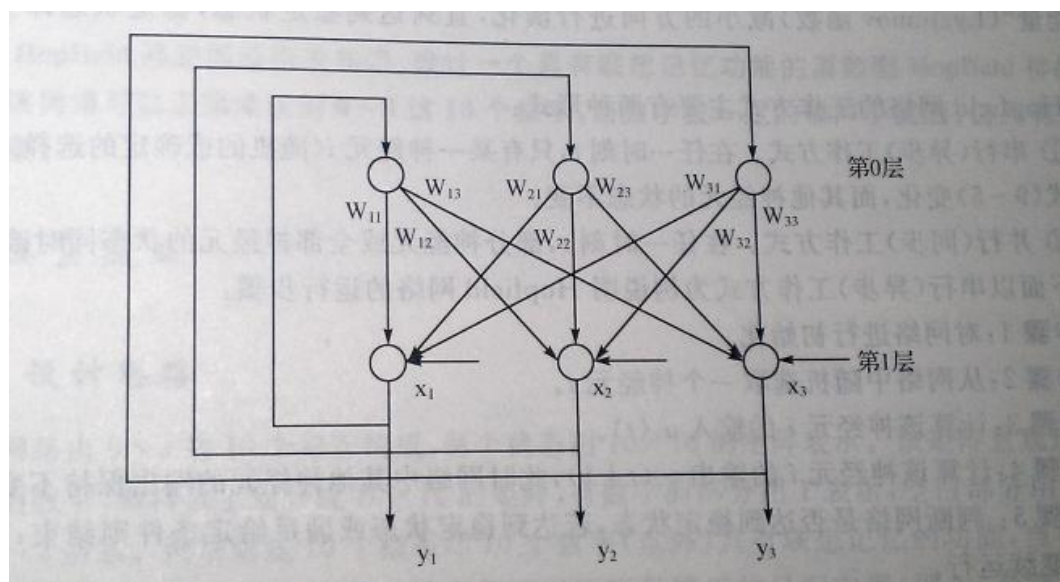
(5) 测试。BP网络输出值并无限定为1或2，而是一个实数，因此还需要将输出转换为整数。取1.5为阈值，小于该阈值的输出判为1（信用好），否则判为2（信用差）

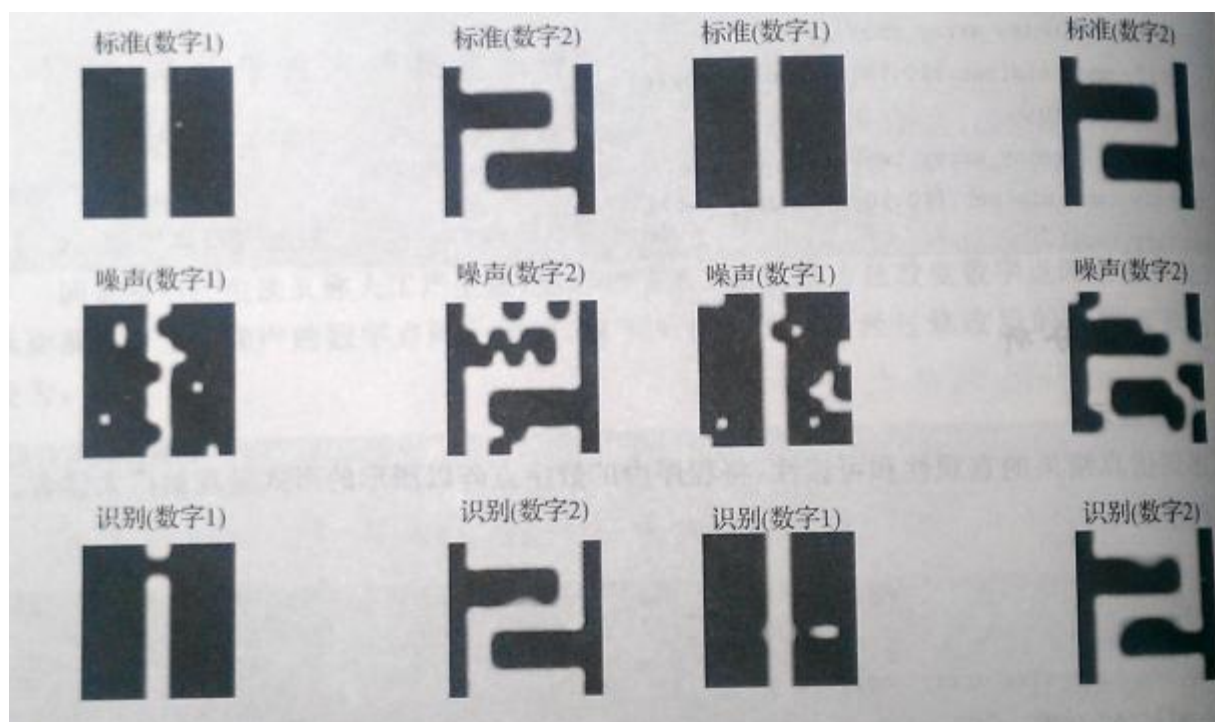
基于BP网络的个人信贷信用评估

运行脚本credit_class.m

测试20次的平均正确率为74.97%，最低正确率为73.4%，迭代次数均为3次

- 人类的联想记忆能力
- Hopfield人工神经网络能模拟联想记忆功能，曾被誉为神经网络方法的以此革命和再生
- Hopfield人工神经网络按动力学方式运行

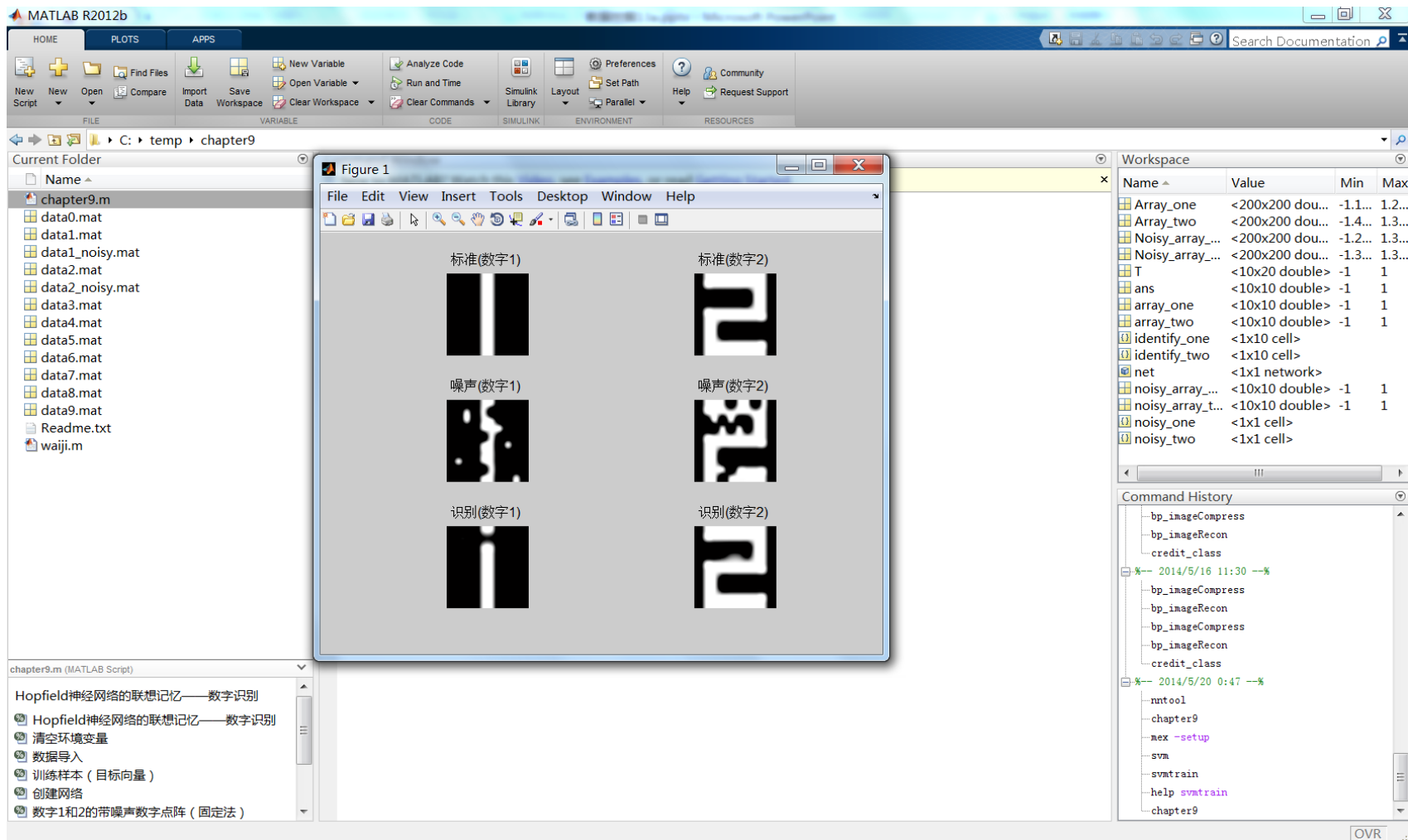




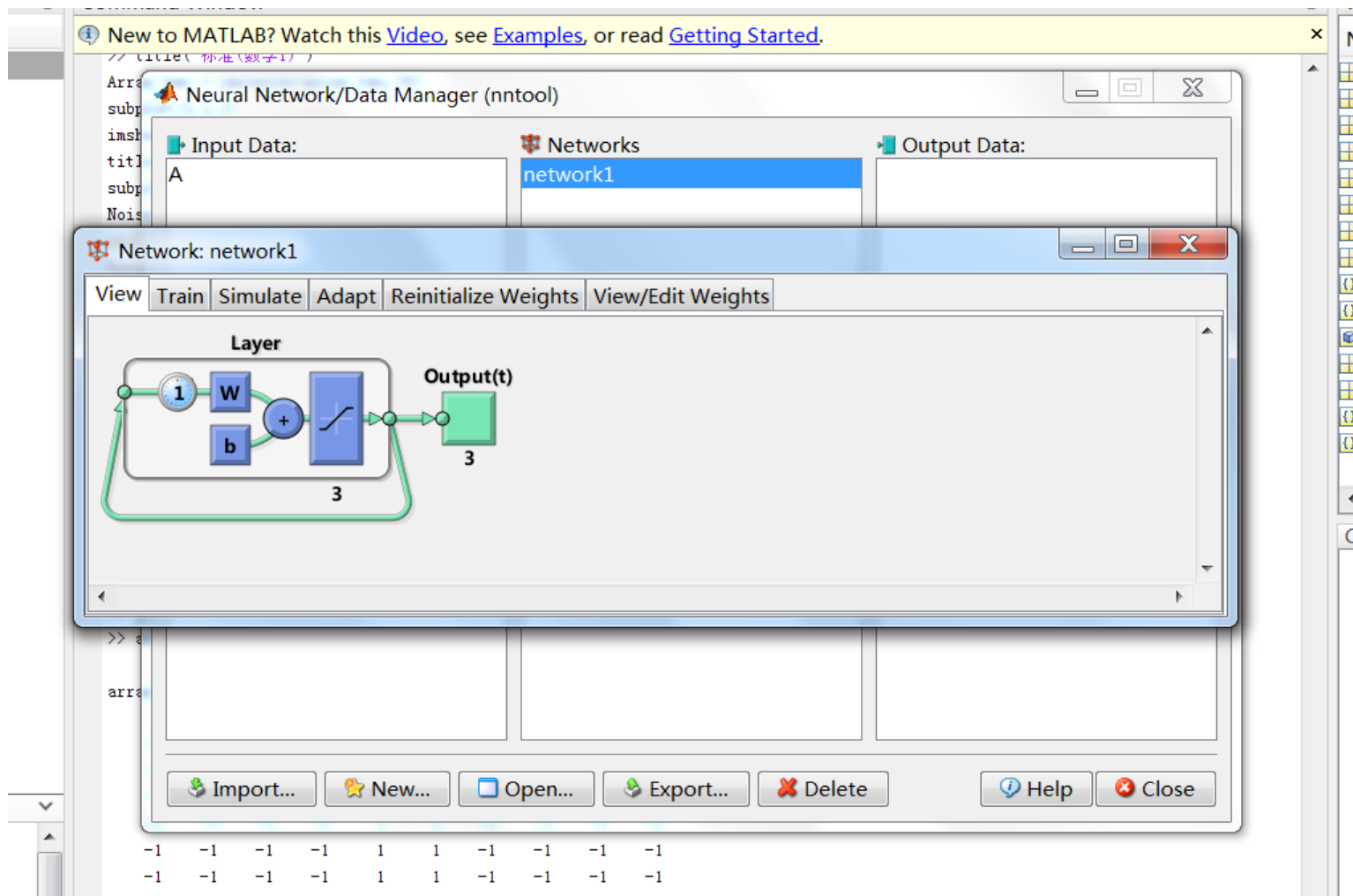


- 把图像信息数字化为1和-1二值矩阵
- 标准图样生成的矩阵作为Hopfield网络的目标向量
- 生成Hopfield网络
- 使用带噪音的矩阵测试
- 输出已经降噪，再和标准目标矩阵（向量）比对，找出最接近者

样例运行结果



创建Hopfield神经网络



- 《Matlab神经网络43个案例分析》第九章

- PCA神经网络（主成分分析）
- SVM神经网络（支持向量机）
- RBF神经网络（基于径向基函数）

设 X 是 p 维随机变量, 并假设 $\mu = E(X)$, $\Sigma = \text{Var}(X)$. 考虑如下线性变换

$$\begin{cases} Z_1 = a_1^T X \\ Z_2 = a_2^T X \\ \vdots \\ Z_p = a_p^T X \end{cases}, \quad (9.1)$$

易见

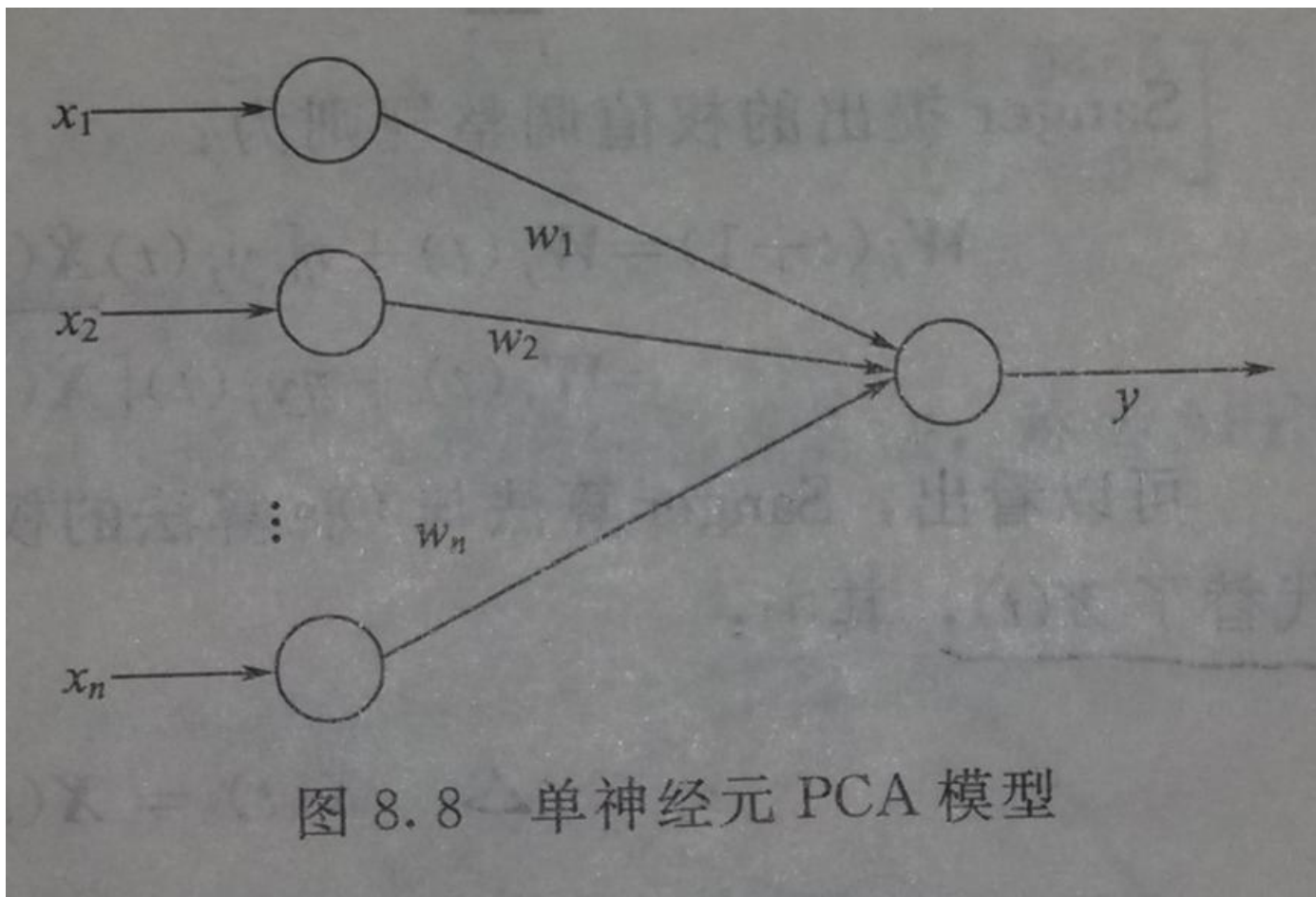
$$\text{Var}(Z_i) = a_i^T \Sigma a_i, \quad i = 1, 2, \dots, p, \quad (9.2)$$

$$\text{Cov}(Z_i, Z_j) = a_i^T \Sigma a_j, \quad i, j = 1, 2, \dots, p, \quad i \neq j. \quad (9.3)$$

我们希望 Z_1 的方差达到最大, 即 a_1 是约束优化问题

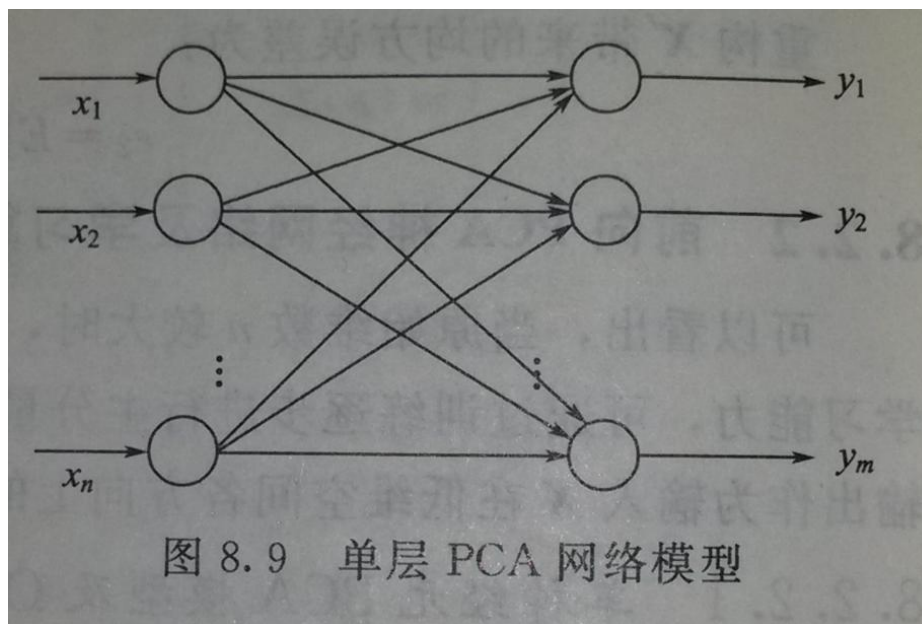
$$\begin{aligned} \max \quad & a^T \Sigma a \\ \text{s.t.} \quad & a^T a = 1 \end{aligned}$$

的解. 因此, a_1 是 Σ 最大特征值 (不妨设为 λ_1) 的特征向量. 此时, 称 $Z_1 = a_1^T X$ 为第一主成分.



- 赫布（1904-1985）是加拿大著名生理心理学家。在1949年出版的《行为的组织》中，赫布提出了其神经心理学理论。赫布认为神经网络的学习过程最终是发生在神经元之间的突触部位，突触的联结强度随着突触前后神经元的活动而变化，变化的量与两个神经元的活性之和成正比。
- Hebb学习规则是一个无监督学习规则，这种学习的结果是使网络能够提取训练集的统计特性，从而把输入信息按照它们的相似性程度划分为若干类。这一点与人类观察和认识世界的过程非常吻合，人类观察和认识世界在相当程度上就是在根据事物的统计特征进行分类。Hebb学习规则只根据神经元连接间的激活水平改变权值，因此这种方法又称为相关学习或并联学习。
- 《人工神经网络理论，设计及应用》第32页
- 把Hebb学习规则应用到PCA神经网络，可以得到方差最大输出，对应于第一主成分。但也有容易发散的问题。

- 《人工神经网络理论，设计及应用》第181页
- Oja学习规则，给Hebb学习规则加上控制项
- Sanger学习规则，可以输出多个主成分



$$L_D = \sum_{i=1}^n \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)$$

使用二次规划技术得到 λ_i 后，就可以通过下面的方程导出参数 \mathbf{w} 和 b ：

$$\mathbf{w} = \sum_i \lambda_i y_i \Phi(\mathbf{x}_i)$$

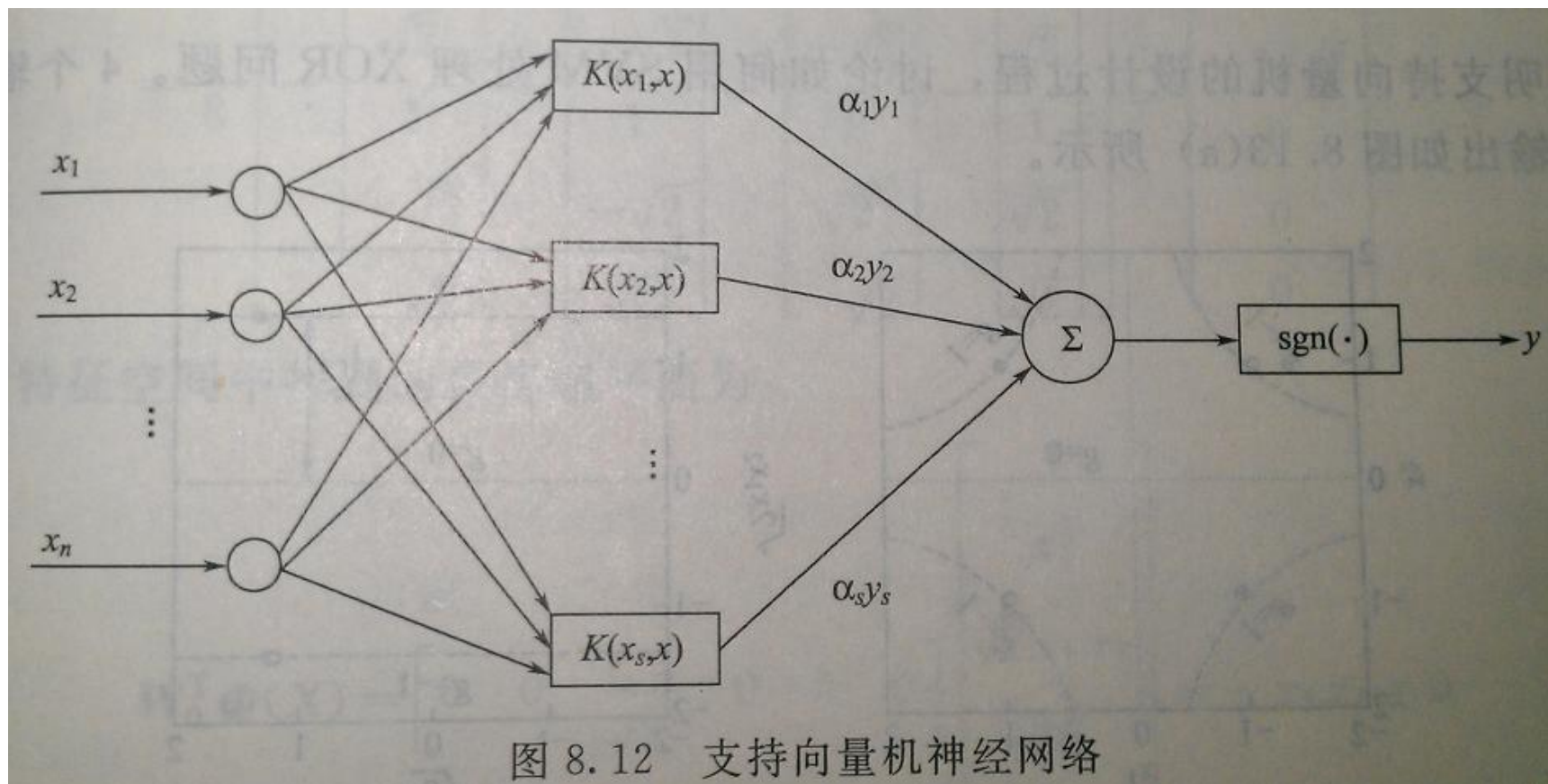
$$\lambda_i \{ y_i \sum_j \lambda_j y_j \Phi(\mathbf{x}_j) \cdot \Phi(\mathbf{x}_i) + b \} = 0$$

这类似于公式 (5-39) 和 (5-40) 的线性 SVM。最后，可以通过下式对检验实例 \mathbf{z} 进行

$$f(\mathbf{z}) = \text{sign}(\mathbf{w} \cdot \Phi(\mathbf{z}) + b) = \text{sign} \left(\sum_{i=1}^n \lambda_i y_i \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{z}) + b \right)$$

$$\max \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j)$$
$$0 \leq \alpha_i \leq C \quad \sum_i \alpha_i y_i = 0$$

$$\begin{aligned} \text{maximize} \quad & L(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j) \\ \text{subject to} \quad & \sum_{i=1}^n \alpha_i y_i = 0 \end{aligned}$$



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.

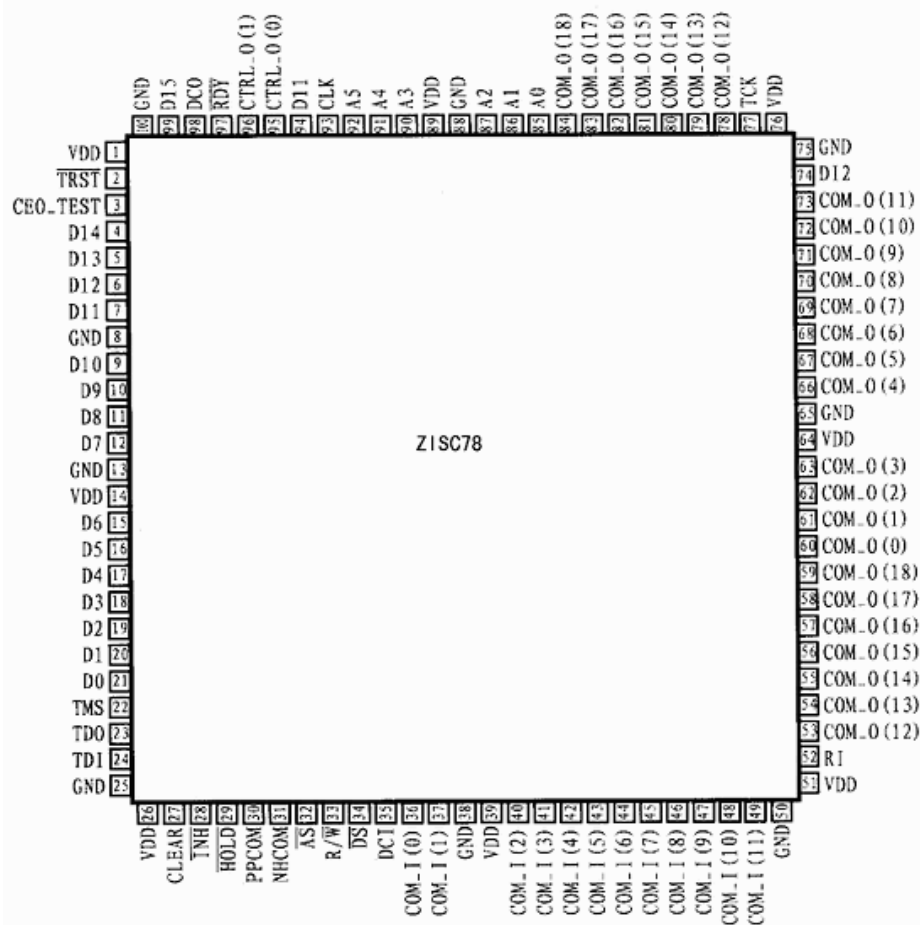


图 1 ZISC78 封装引脚图

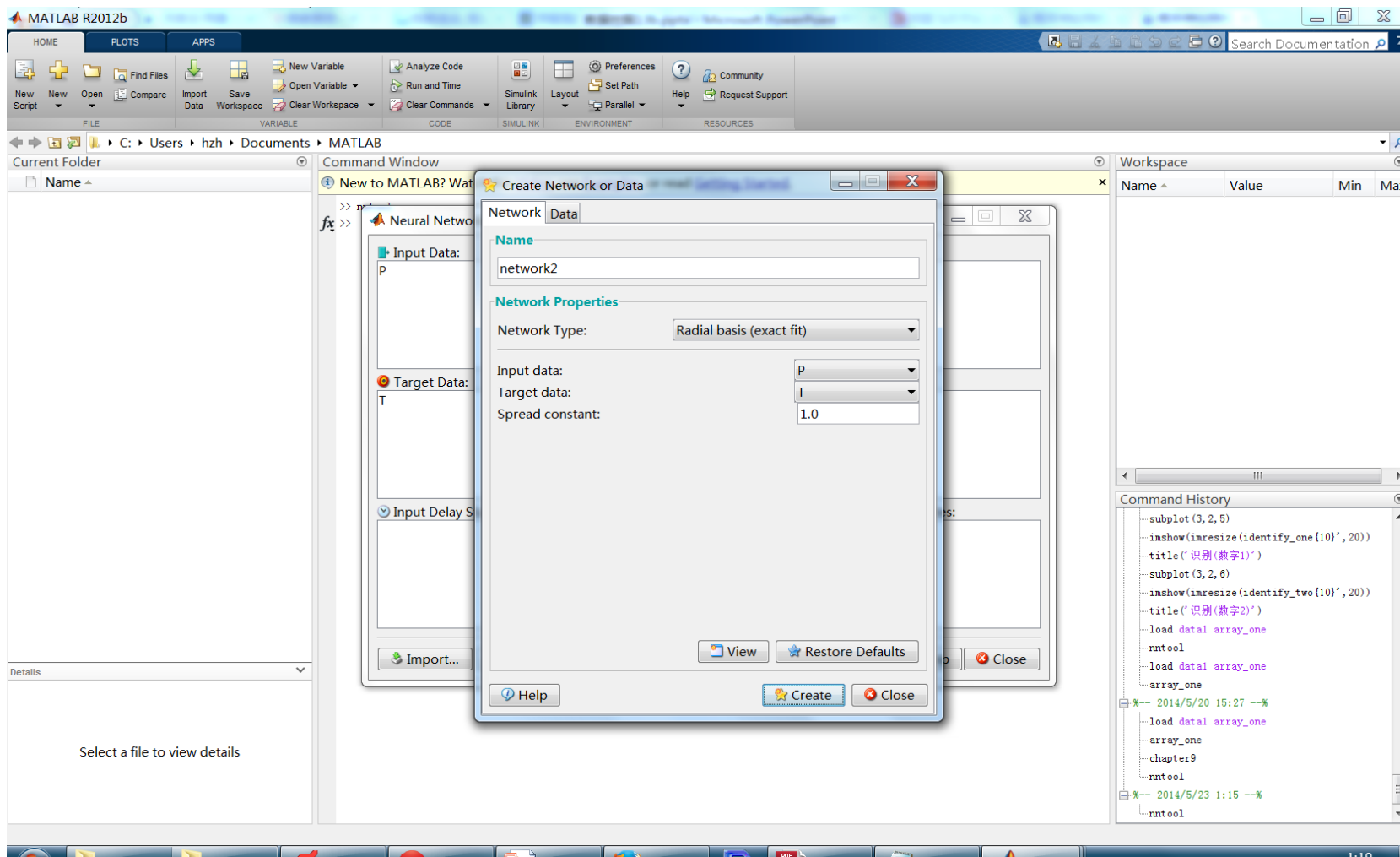
- 范数
- 什么是径向基函数
- 常用径向基函数：高斯函数，反演S型函数，拟多二次函数
- 逼近解释
- 《人工神经网络理论、设计及应用》第165页

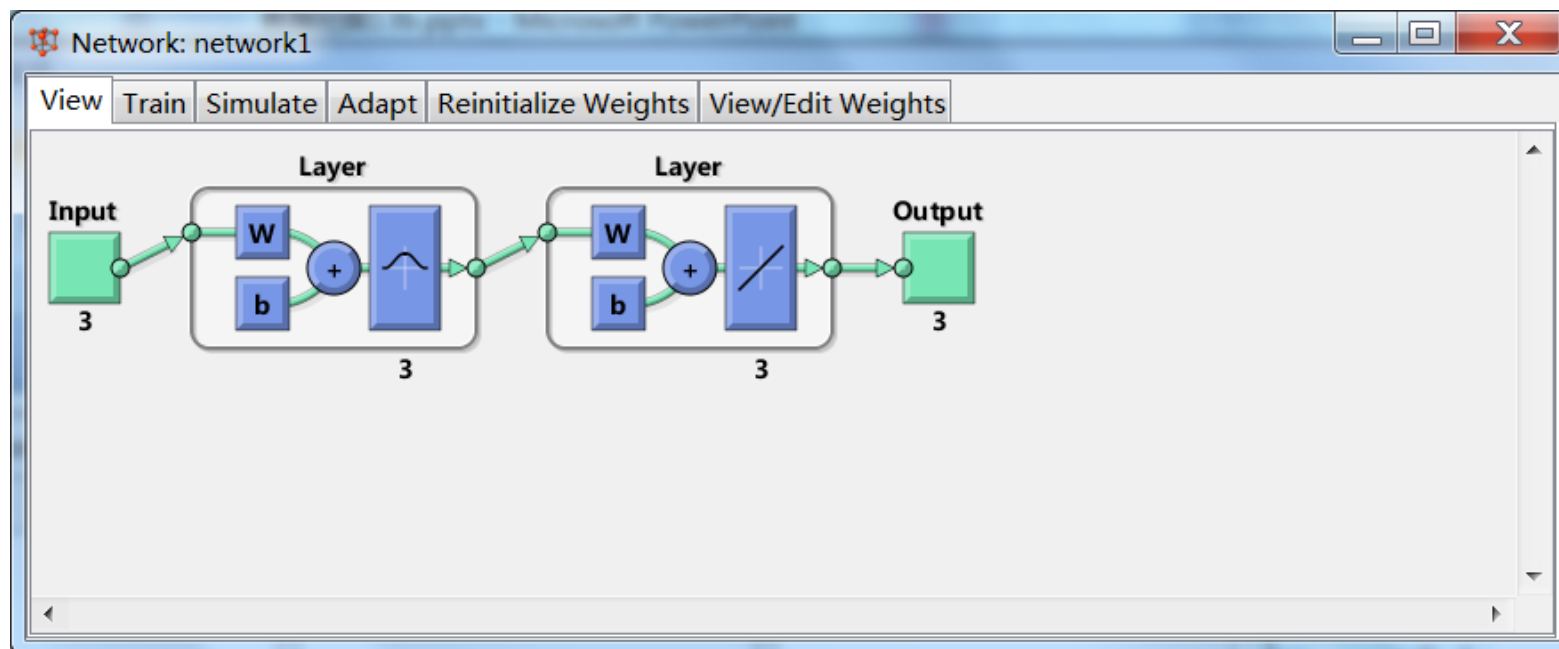
- 隐藏层节点数即为样本数
- 使用高斯函数作为隐藏层激活函数，称为基函数
- 数据中心即样本点，扩展常数可由样本分布（和两两之间距离有关）和样本数决定。
所有节点的基函数取相同的扩展常数
- 输出层使用线性激活函数
- 是一种通用逼近器
- 学习算法，类似LMS
- 《人工神经网络理论、设计及应用》第167页

- 基本思想还是将低维度线性不可分问题（非线性问题）映射到高维空间变成线性可分问题解决（径向基函数用于逼近映射）
- 隐藏层节点数一般远远小于样本数
- 中心不一定在样本点，通过训练算法（可能需要聚类）来选取
- 扩展常数不再统一，由训练算法得出
- 输出函数在线性函数中包括阈值项
- 训练算法（基于k-means聚类）

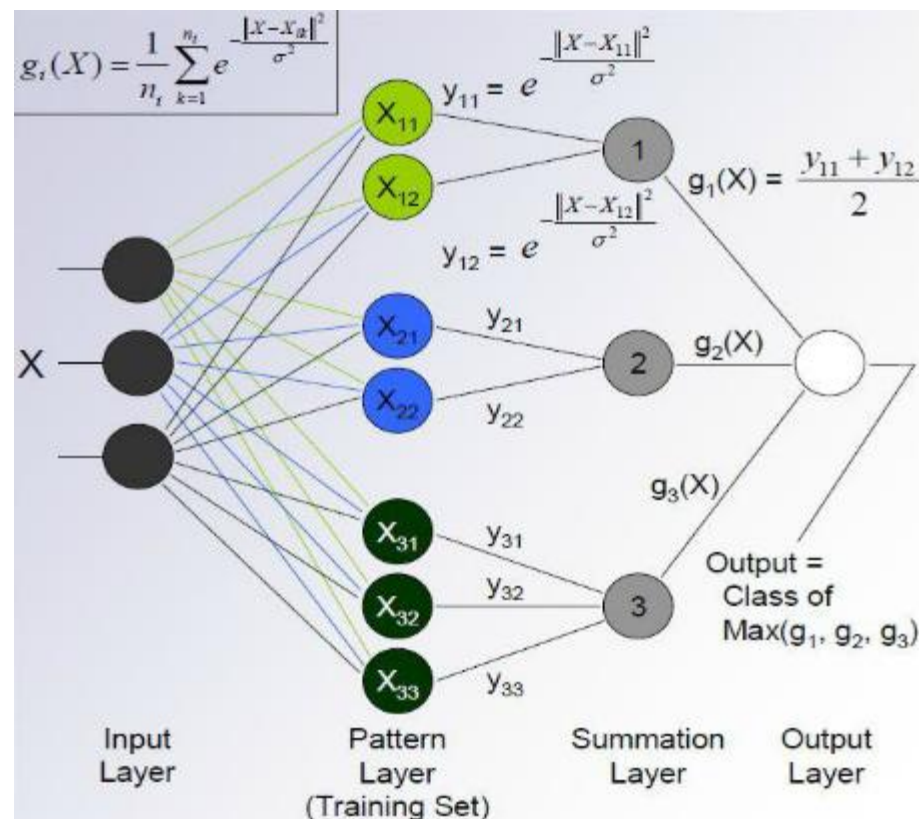
- RBF只有一个隐藏层，BP可以多层
- RBF隐藏层和输出层使用不同的激活函数，BP全相同
- 所使用的激活函数特性不一样，RBF选择局部指数衰减的非线性函数（强调局部），BP则使用输入样本于权向量的内积作为自变量，激活函数采用sigmoid或限幅函数（强调全局）
- BP是对非线性映射的全局逼近，RBF更突出局部特性（回忆SVM的情形）

Matlab实现RBF神经网络





- <http://wenku.baidu.com/view/9cedca1316fc700abb68fcec.html>
- Probabilistic neural networks, 简称PNN
- D. F. Specht博士在1989年首先提出
- RBF神经网络的一种, 基于贝叶斯决策理论, 不同的参数取值, 可以对应于贝叶斯分类器, 最近邻算法等几种传统的分类器算法



- Bayes Belief Network , 简称BBN
- 朴素贝叶斯分类器需要特征之间互相独立的强条件 , 制约了模型的适用
- 用有向无环图表达变量之间的依赖关系 , 变量用节点表示 , 依赖关系用边表示
- 祖先 , 父母和后代节点。贝叶斯网络中的一个节点 , 如果它的父母节点已知 , 则它条件独立于它的所有非后代节点
- 每个节点附带一个条件概率表 (CPT) , 表示该节点和父母节点的联系概率

■ 韩家炜书第256页

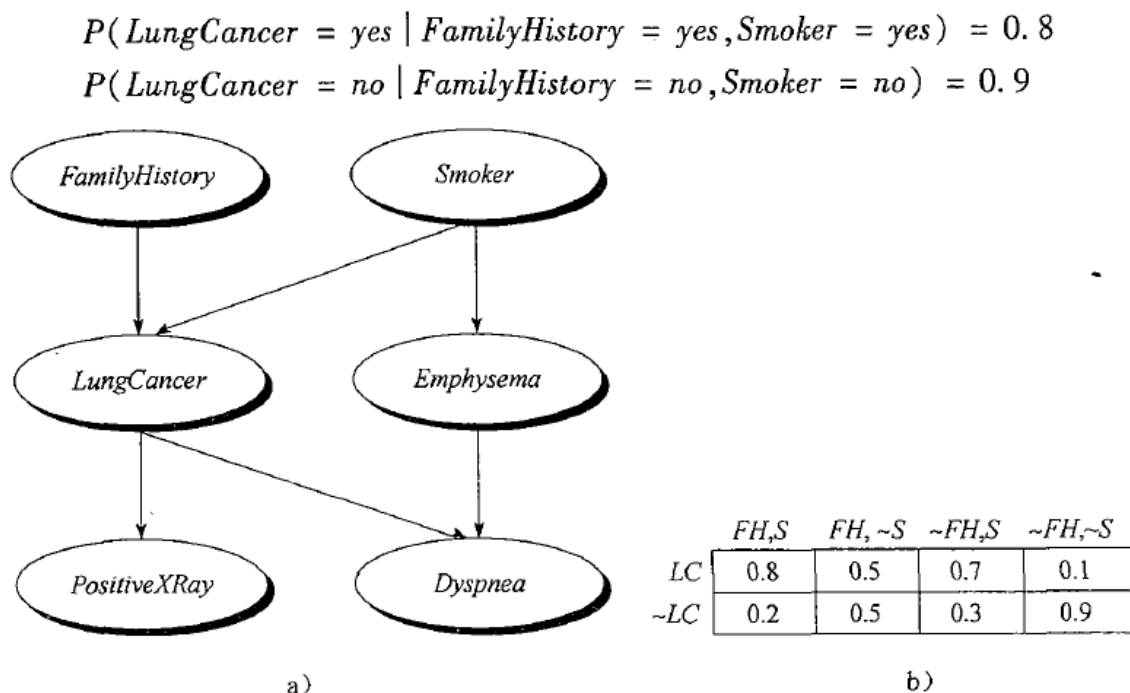


图 9.1 一个简单的贝叶斯信念网络：a) 一个提议的因果模型，用有向无环图表示；b) 变量 *LungCance*(*LC*) 的条件概率表，给出其双亲节点 *FamilyHistory* 和 *Smoke* 的每个可能值组合的条件概率。取自 Russell、Binder、Koller 和 Kanazawa[RBKK95]

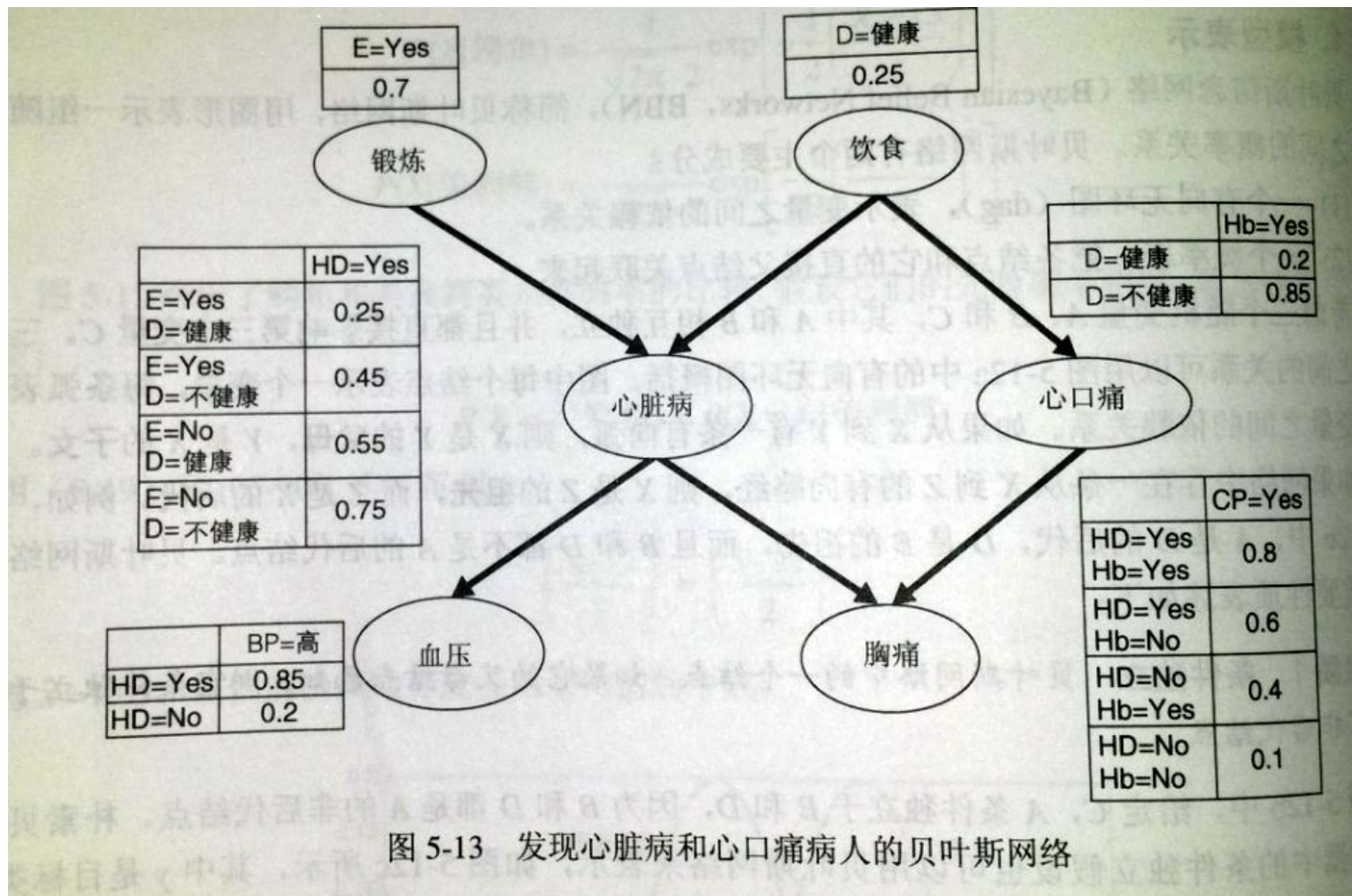


图 5-13 发现心脏病和心口痛病人的贝叶斯网络

- 什么时候要训练？存在隐藏变量或隐藏数据
- 转化为类似神经网络求权值的问题，使用梯度下降法进行训练
- 韩家炜书第257页
- 目标函数及解释

$$P_w(D) = \prod_{d=1}^{|D|} P_w(X_d)$$

$$\frac{\partial \ln P_w(D)}{\partial w_{ijk}} = \sum_{d=1}^{|D|} \frac{P(Y_i = y_{ij}, U_i = u_{ik} | X_d)}{w_{ijk}}$$

$$w_{ijk} \leftarrow w_{ijk} + (l) \frac{\partial \ln P_w(D)}{\partial w_{ijk}}$$

- **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