

清华大学信息科学技术学院教材
——自动化系列
新编《信息、控制与系统》系列教材

Artificial Neural Networks and Evolutionary Computing

人工神经网络与模拟进化计算 (第2版)

阎平凡 张长水 编著
Yan Pingfan Zhang Changshui



TUP
清华大学出版社

 Springer



新编《信息、控制与系统》系列教材

人工神经网络与模拟 进化计算(第2版)

阎平凡 张长水 编著

清华大学出版社
北 京

内 容 简 介

本书较系统全面地讨论了人工神经网络与模拟进化计算的理论和工程应用,特别在学习理论和网络结构选择、动态神经网络、贝叶斯方法的应用以及模拟进化计算中的一些理论问题等方面的论述更为系统深入。讲解中力求讲清物理概念,以便读者深入理解一些主要方法的思路。

第2版加强了关于统计学习理论、核方法与支持向量机、自组织网络的灵活应用,盲信号处理等方面的内容;增加了神经网络在生物信息学和金融方面应用的实例,以及最近的一些参考文献,以便反映这一领域的新进展;为了便于掌握主要内容,对章节顺序也做了调整,模拟进化计算部分增加了分布估计算法一章。

本书适合用作研究生课程教材,或作为希望深入学习神经网络和进化计算的科技工作者的自学参考书。

版权所有,翻印必究。举报电话:010-62782989 13501256678 13801310933

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

本书防伪标签采用特殊防伪技术,用户可通过在图案表面涂抹清水,图案消失,水干后图案复现;或将表面膜揭下,放在白纸上用彩笔涂抹,图案在白纸上再现的方法识别真伪。

图书在版编目(CIP)数据

人工神经网络与模拟进化计算/阎平凡等编著. —2版. —北京:清华大学出版社,2005.9

ISBN 7-302-10663-0

I. 人… II. 阎… III. 人工神经网络—计算 IV. TP183

中国版本图书馆 CIP 数据核字 (2005) 第 020103 号

出 版 者: 清华大学出版社

<http://www.tup.com.cn>

社 总 机: 010-62770175

地 址: 北京清华大学学研大厦

邮 编: 100084

客户服务: 010-62776969

组稿编辑: 王一玲

文稿编辑: 赵从棉

印 刷 者: 北京市密云胶印厂

装 订 者: 三河市化甲屯小学装订二厂

发 行 者: 新华书店总店北京发行所

开 本: 175×245 印张: 41.75 字数: 977 千字

版 次: 2005 年 9 月第 2 版 2005 年 9 月第 1 次印刷

书 号: ISBN 7-302-10663-0/TP·7222

印 数: 1~3000

定 价: 49.00 元

新编《信息、控制与系统》系列教材

出版说明

信息、控制与系统学科是在 20 世纪上半叶形成和发展起来的一门新兴技术科学。在人类探索自然和实现现代化的进程中,信息、控制与系统学科的理论、方法和技术始终起着重要的和基础的作用。基于信息、控制与系统科学的自动化的发展和应用水平在一定程度上是一个国家和社会的现代化程度的重要标志之一。本系列教材是关于信息、控制与系统学科所属各个领域的基本理论和前沿技术的一套高等学校系列教材。

本系列教材所涉及的范围包括信号和信息处理、模式识别、知识工程、控制理论、智能控制、过程和运动控制、传感技术、系统工程、机器人控制、计算机控制和仿真、网络化系统、电子技术等方面。主要读者对象为自动控制、工业自动化、计算机科学和技术、电气工程、机械工程、化工工程和热能工程等系科有关的高年级大学生和研究生,以及工作于相应领域和部门的科学工作者和工程技术人员。

10 多年前,清华大学出版社会同清华大学自动化系,曾经组编出版过一套《信息、控制与系统》系列教材,产生了较大的社会影响,其中多数著作获得过包括国家级教学成果奖和部委优秀教材奖在内的各种奖励,至今仍为国内众多院校所采用,并被广大相关领域科技人员作为进修和自学读物。我们现在组编的这套新编《信息、控制与系统》系列教材,从一定意义上说,就是先前那套教材的延伸和发展,以反映近些年来学科的发展和在科学研究与教学实践上的新成果和新进展,以适应当前科技发展和教学改革的新形势和新需要。列入这套新编系列教材中的著作,大多是清华大学自动化系开设的课程中经过较长教学实践而形成的,既有多年教学经验和教学改革基础上的新编著的教材,也有部分原系列教材的更新和修订版本。这套新编系列教材总体上仍将保持原系列教材求新与求实的风格,力求反映所属领域的基本理论和新近进展,力求做到学科先进性和教学适用性统一。需要说明的是,此前我们曾以《信息技术丛书》为名组编这套教材,并已出版了若干种著作。现为使“书”和“名”更为相符,这些已出版的著作将在重印或再版时列入这套新编系列教材。

我们希望,这套新编系列教材,既能为在校大学生和研究生的学习提供内容先进、论述系统和教学适用的教材或参考书,也能为广大科学工作者与工程技术人员的知识更新与继续学习提供适合的和有价值的进修或自学读物。我们同时要感谢使用本系列教材的广大教师、学生和科技工作者的热情支持,并热忱欢迎提出批评和意见。

新编《信息、控制与系统》系列教材编委会

2002 年 6 月

新编《信息、控制与系统》系列教材编委会

顾问	李衍达	吴澄	边肇祺	王桂增	
主编	郑大钟				
编委	徐文立	王雄	萧德云	杨士元	肖田元
	张贤达	周东华	钟宜生	张长水	王书宁
	范玉顺	蔡鸿程			
责任编辑	王一玲				

前 言

自本书(第1版)出版以来,神经网络和模拟进化计算又有了很大发展。特别是统计学习理论和核方法的引入,使得对学习问题的研究更加深入和系统化。同时,应用领域也不断扩展。在第2版中主要做了如下修改:

1. 较系统地介绍了统计学习理论、核方法与支持向量机。除必要的数学推导外,力求讲清楚物理概念及与其他方法的联系。加强了对自组织映射的分析和独立成分分析的讨论。

2. 增加了一章“分布估计算法”,该算法是对遗传算法的发展,也是进化计算的一个重要组成部分。该章讲述中强调方法的系统性以及各种方法之间的联系。

3. 根据目前各种模型的发展及应用情况,对章节顺序作了一些调整,并增加了一些在生物信息学以及金融领域中的应用实例。

本书各章安排如下。

第1~3章介绍基本概念和前馈网络;第4,5章讨论学习问题、核方法和支持向量机;第6~8章讨论自组网络、主成分分析、自组织映射以及独立成分分析;第9章讨论动态信号和系统的处理;第10章介绍多网络集成及有关算法;第11,12章讨论单层全连接的反馈网络及其在联想存储和优化计算中的应用;第13章讨论了反馈网络中的动力学问题并简单介绍了混沌神经网络;第14~16章进一步讨论了误差函数、参数优化、贝叶斯方法以及神经网络在信号处理中的应用。

第 17~22 章的内容是模拟进化计算。第 17 章讲述基本概念;第 18~21 章讨论遗传算法的理论、算法设计和实现以及在神经网络和作业调度中的应用;第 22 章较系统地讨论了分布估计算法。

这一领域发展很快,新的理论和应用还会不断出现。限于编者水平,遗漏和不当之处在所难免,欢迎读者给予批评指正。

作 者

2005 年 2 月

第一版前言

现在国内有关神经网络和模拟进化计算的书不少,其中也不乏优秀之作,但作为研究生教材,就编者多年的教学经验看,在广度和深度方面尚不能完全满足要求。研究生课程学时较少,来自各专业的研究生参加了各种各样的科研任务,所提出的问题非常广泛和深入,希望教材有一定深度和广度,以便他们进一步扩展和深入,本书向这个方向作了一些努力。

著名学者 D. Marr 认为,对一个复杂的信息处理系统的研究可分为 3 个层次,即计算理论层、算法层和实现层。计算理论层解决处理的目的,用什么理论,并说明所用理论为何能达此目的;算法层解决为实现计算理论所采用的算法;实现部分应给出可执行的或硬件实现的具体算法。编者认为,对于神经网络来说,计算理论层主要是学习理论,而现在的学习理论是基于统计决策理论的;算法部分是非线性优化方法(所有的学习算法都可归结为一种优化算法);实现层背后的理论牵涉到非线性动力学系统的稳定性、可靠性理论等。从这一观点出发,本书加强并扩展了以下几方面的内容。

1. 学习理论与网络结构选择 第 4 章较系统地介绍了学习理论方面最近的研究成果,从工程应用的角度简要介绍了 V. Vapnik 的统计学习理论和经验风险最小化方法;特别强调了有限样本下的学习和网络结构选择问题。贝叶斯理论是研究归纳学习的有力工具,第 14 章从贝叶斯学习的观点进一步讨论了参数学习和网络结构选择问题。

2. 优化方法 第 13 章专门讨论了误差函数和参数优化方法,介绍了一些不同的误差函数,重点讨论了各种基于梯度的优化方法,而进化计算部分则主要讨论全局优化方法。

3. 动态系统和信号的处理 神经网络在系统辨识和随时间变化的信号处理(如时间序列的建模和预测)等方面的应用越来越多,第10章专门讨论了这方面的问题。第15章讨论神经网络在信号处理中的应用,可看作是第10章的扩展。

模拟进化计算的具体算法及其改进很多,本书不去求全,只结合最常用的遗传算法和进化策略从理论上作了较深入的分析。

在编写过程中,参考了 S. Haykin 编写的“Neural Networks, A Comprehensive Fundation”(1994 年版)以及 C. M. Bishop 编写的“Neural Networks for Pattern Recognition”(1995 年版)两本书。有许多内容取材于最近的国内外文献。各章后面都附有较多的参考文献,以便读者查阅。

本书各章安排如下:第1~15章为第一部分。第1章是绪论,介绍了人工神经元模型、网络结构和工作方式以及一些重要的学习方法。第2~4章讲述前馈网络。其中第2章讨论了前馈网络的一般问题,反向传播学习算法,为便于理解前馈网络作用的机理,2.6节对前馈网络各层的作用从物理概念上作了分析;第3章讲述径向基函数网络和与之有关的概率神经网络等;第4章主要讨论学习和网络结构选择问题,同时也介绍了新近出现的一些神经元模型。第5,6章讲述单层全反馈网络。第5章讨论反馈网络的一般问题和联想存储器,重点讨论了联想存储器的设计,最后介绍了玻耳兹曼机;第6章讲述反馈网络用于优化计算中的一些问题。第7~9章讲述自组织网络。第7章讨论基于 Hebb 学习的自组织网络,讲述了主成分分析网络和独立成分分析问题;第8章讨论基于竞争学习的自组织网络,重点讨论常用的自组织特征映射算法和向量量化算法,对自适应谐振理论也做了介绍;第9章从信息理论角度研究自组织过程。第10章专门讨论用于处理动态系统和信号的神经网络,介绍了带外加延时的前馈网络和部分反馈网络这两种最常用的模型;讨论了动态网络的一些主要学习算法;最后介绍了再励学习算法。第11章讨论反馈网络的动力学问题,由于反馈网络是一个非线性动力学系统,从数学上严格分析很困难,只能从工程应用的角度,对一些常用的网络模型的稳定性作了定性分析,最后,对混沌神经网络作了介绍。第12~14章的内容主要针对前馈网络。第12章讨论了用多个子网络组成复合网络来解决一些较复杂问题的方法;第13章针对学习过程中误差函数的作用和重要的参数优化方法作了介绍;第14章讨论了基于贝叶斯理论的学习方法和网络结构选择问题。第15章主要讨论神经网络在信号处理中的应用,它是第10章内容的深入和扩充。

第16~20章属于本书的第二部分。第16章对进化计算的一般概念作了介绍,系统讲述了进化策略的基本原理和算法;第17章讨论遗传算法,介绍了标准的遗传算法,对遗传算法的理论和收敛问题做了较深入的讨论;第18章进一步从实现的角度讨论遗传算法的设计和实现,包括编码方法、适应度函数的选择、遗传算子和参数选择问题;第19,20章讲述遗传算法的应用,分别讨论了在神经网络结构选择和权值学习中的应用以及在作业

调度中的应用。

作为教材,第1~6章、第8章和第10章,16.1,17.1~17.3节属于基本内容,其余各章可作为专题内容。

本书第1~15章由阎平凡编写,第16~20章由张长水编写。清华大学出版社为本书的编写出版做了大量工作,在此表示衷心的感谢。

限于作者水平,书中错误遗漏之处在所难免,恳请读者批评指正。

编 者

1999年12月25日

目 录

第 1 章 绪论	1
1.1 神经网络的发展与应用	1
1.2 人工神经元模型	2
1.3 用有向图表示神经网络	4
1.4 网络结构及工作方式	5
1.5 NN 的学习	7
1.5.1 学习方式	7
1.5.2 学习算法	7
1.5.3 学习与自适应	9
习题	9
参考文献	10
第 2 章 前馈网络	11
2.1 线性阈值单元	11
2.1.1 用线性阈值单元实现布尔函数	11
2.1.2 线性可分性	12
2.1.3 n 维欧氏空间中 m 个点上可实现的线性可分函数的个数	13
2.2 多层前馈网络的计算能力及函数逼近	15

2.3	感知器的学习算法	17
2.4	反向传播学习算法	19
2.5	改进反向传播算法收敛速度的措施	26
2.5.1	加入动量项	26
2.5.2	高阶导数的利用	27
2.5.3	共轭梯度法	27
2.5.4	递推最小二乘法	28
2.5.5	神经元空间搜索法	28
2.5.6	一些其他措施	30
2.6	多层前馈网络作用的分析	31
2.6.1	线性网络	32
2.6.2	非线性情况	33
2.7	应用举例	34
	习题	45
	参考文献	48

第3章	径向基函数网络	51
3.1	φ 可分性	51
3.2	函数逼近与内插	52
3.3	正规化理论	53
3.4	RBF 网络的学习	60
3.5	RBF 网络的一些变形	63
3.6	CMAC 网络	64
3.6.1	模型结构	64
3.6.2	工作原理分析	66
3.6.3	学习算法	69
3.7	概率神经网络	70
3.8	小波网络	71
3.9	泛函连接网络	72
3.10	新一代神经元模型及其计算能力的研究	74
3.10.1	布尔函数的计算	75
3.10.2	连续输入的情况	75
3.10.3	脉冲耦合神经网络	76
	参考文献	81

第4章 学习理论与网络结构选择	86
4.1 经验风险最小化与推广能力	88
4.1.1 经验风险最小化原则	89
4.1.2 推广能力、模型复杂度和样本量	90
4.2 统计学习理论简介	92
4.2.1 学习过程一致性的条件	93
4.2.2 统计学习理论发展中的三个重要结论	96
4.2.3 推广能力的界	97
4.2.4 结构风险最小化	102
4.3 学习的复杂性问题	105
4.3.1 PAC 学习模型	105
4.3.2 PAC 学习的例子	106
4.3.3 PAC 学习模型的一些扩展	106
4.3.4 多层前馈网络的样本数问题	107
4.3.5 学习的计算复杂性	108
4.4 学习的动态特性	109
4.4.1 通用学习方程	109
4.4.2 LMS 规则	110
4.4.3 Hebb 规则	111
4.4.4 Oja 学习规则	111
4.5 推广问题	112
4.5.1 定性分析	112
4.5.2 平均推广能力	113
4.5.3 样本量问题	115
4.5.4 推广误差的实验估计	116
4.6 预测学习	117
4.6.1 模型	117
4.6.2 根本困难	118
4.6.3 维数灾难问题	118
4.6.4 方差与偏置折衷	118
4.7 学习的数学基础	120
4.8 网络模型选择	122
4.8.1 定性分析	122
4.8.2 正规化方法	123

4.8.3 修剪与网络构造法	124
4.9 符号学习与神经网络结合	131
习题	133
参考文献	134
第5章 核方法与支持向量机	140
5.1 现有方法概述	141
5.1.1 基函数法	142
5.1.2 核函数	143
5.2 基函数法与核函数法的对偶关系	144
5.3 自适应方法	147
5.4 核函数法	147
5.5 表示定理	151
5.6 最优线性分界面	152
5.7 支持向量机(SVM)	155
5.7.1 基本原理	155
5.7.2 SVM用于多类(C类)问题	161
5.7.3 SVM用于回归	161
5.8 核 PCA	163
5.9 SVM的算法	165
5.10 贝叶斯方法与高斯过程	171
5.10.1 贝叶斯方法	171
5.10.2 高斯过程	173
5.10.3 稀疏表示	175
5.11 讨论	177
参考文献	181
第6章 自组织系统(I)——Hebb学习	184
6.1 引言	184
6.2 自组织特征检测——一个简单的实验	185
6.3 主成分分析	186
6.4 单个神经元抽取最大主分量	189
6.5 单层网络用于抽取一组主分量	189
6.6 有侧向连接的自适应 PCA	191

6.7	最小均方误差重建学习	193
6.8	次分量的提取和应用	195
6.8.1	最优拟合问题	195
6.8.2	用单个神经元实现	197
6.9	PCA 算法的进一步扩展	200
6.9.1	非线性 PCA	200
6.9.2	主曲线与主曲面	200
6.9.3	鲁棒 PCA 算法	202
6.9.4	偏最小二乘方法	202
6.10	用于特征抽取的网络	203
6.10.1	正态分布的数据	204
6.10.2	类内、类间距离的计算	205
6.10.3	Bhattacharya 距离	206
	参考文献	207
第 7 章	自组织系统(Ⅱ)——竞争学习	210
7.1	Hamming 网络与 WTA 网络	210
7.2	自组织特征映射	212
7.3	等效的 SOM 算法	215
7.4	向量量化	220
7.5	广义向量量化	221
7.6	讨论	223
7.7	应用举例	225
7.8	自适应共振理论	235
7.8.1	ART 的基本原理	236
7.8.2	ART 作为分类器时的学习算法	239
	习题	240
	参考文献	241
第 8 章	自组织系统(Ⅲ)——基于信息论的模型	245
8.1	信息论简介	245
8.2	最大信息保持原则	247
8.2.1	单个神经元受噪声干扰	247
8.2.2	输入受加性噪声干扰	248

8.2.3 更复杂些的情况	249
8.3 拓扑有序映射的产生	251
8.4 基于最大熵原则的拓扑映射	253
8.5 盲信号处理	255
8.5.1 一些基本概念	257
8.5.2 ICA 的目标函数	261
8.5.3 盲信号处理的应用	267
8.5.4 非线性独立成分分析(NLICA)	270
参考文献	273
第 9 章 动态信号与系统的处理	276
9.1 延时单元网络	276
9.2 时空神经元模型	280
9.2.1 模型	280
9.2.2 FIR 网络的学习算法	282
9.3 部分反馈网络	286
9.4 有反馈网络的学习算法	289
9.4.1 随时间演化的反向传播算法	289
9.4.2 实时递归学习	290
9.5 应用举例	293
9.6 神经网络在金融领域中的应用	304
9.7 讨论	310
9.8 再励学习及其主要算法	314
9.8.1 时间差分法	315
9.8.2 RL 的主要算法	317
9.9 再励学习在控制中的应用举例	319
习题	327
参考文献	328
第 10 章 多神经网络集成	335
10.1 引言	335
10.2 群体平均法	336
10.3 Boosting 方法	337
10.3.1 Boosting 算法	338

10.3.2	性能分析	339
10.3.3	应用实例——由基因表达谱对肿瘤分类	340
10.4	混合专家网络	342
10.5	分层混合专家网络	344
10.5.1	工作原理	345
10.5.2	EM 算法概述	347
10.5.3	EM 算法用于 HME	349
10.5.4	IRLS 算法	350
10.5.5	EM 算法的步骤	352
10.6	应用举例	353
	参考文献	357
第 11 章	反馈网络与联想存储器	360
11.1	联想存储器	360
11.2	反馈网络	361
11.2.1	离散 Hopfield 网络	362
11.2.2	连续 Hopfield 网络	366
11.3	用反馈网络作联想存储器	368
11.4	相关学习算法	369
11.5	容量分析	373
11.6	伪逆学习算法	377
11.7	基于线性可分性的学习算法	378
11.8	Li 与 Michel 的设计方法(特征结构法)	379
11.9	线性规划方法	379
11.10	多余吸引子问题	381
11.11	双向联想存储器(BAM)	382
11.11.1	能量函数与稳定性	384
11.11.2	BAM 网络的权值设计	384
11.12	玻耳兹曼机	386
11.12.1	随机神经元	386
11.12.2	模拟退火算法	387
11.12.3	玻耳兹曼机	387
11.12.4	玻耳兹曼机的学习	388
11.12.5	平均场学习规则	392