

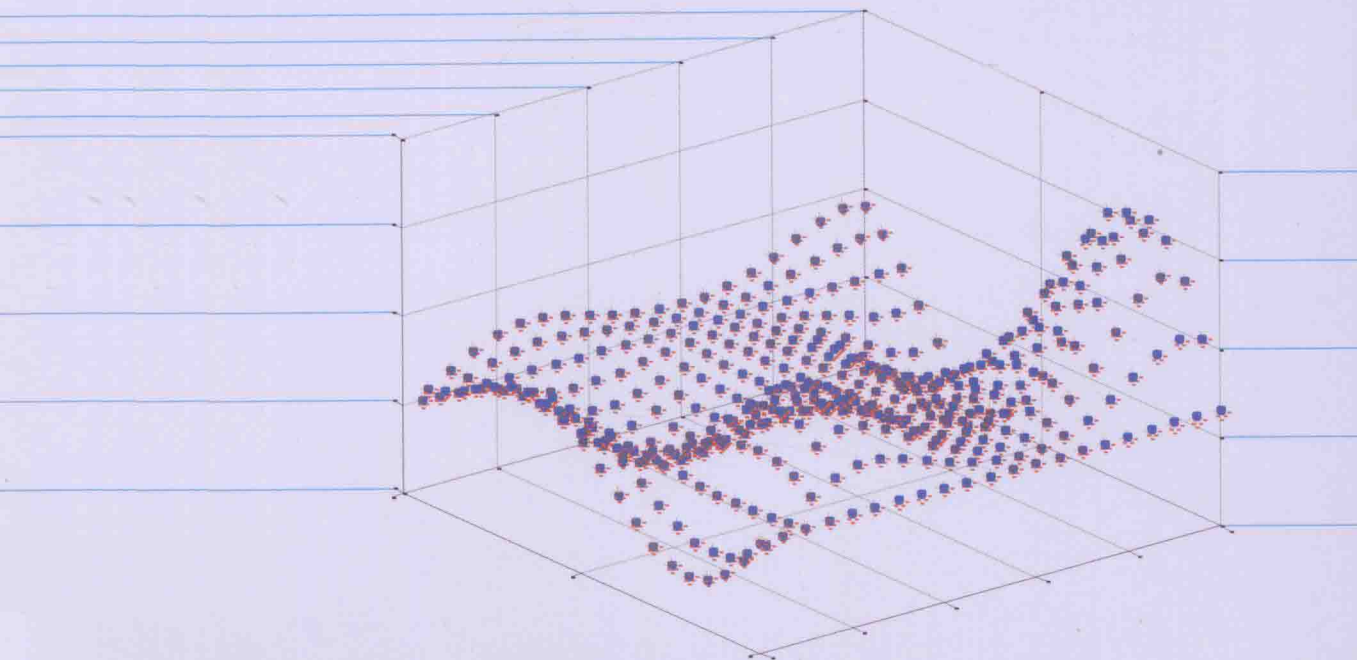


● 张昭昭 乔俊飞 著

模块化神经网络结构分析与设计



Analysis and Design of
Modular Neural Network



辽宁科学技术出版社

LIAONING SCIENCE AND TECHNOLOGY PUBLISHING HOUSE

辽宁省优秀自然科学著作

模块化神经网络结构 分析与设计

张昭昭 乔俊飞 著

辽宁科学技术出版社

沈 阳

© 2014 张昭昭 乔俊飞

图书在版编目 (CIP) 数据

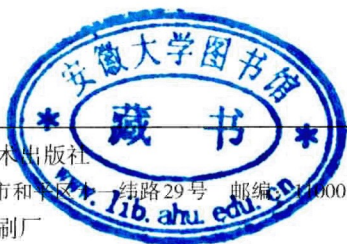
模块化神经网络结构分析与设计 / 张昭昭, 乔俊飞
著. —沈阳: 辽宁科学技术出版社, 2014.4

(辽宁省优秀自然科学著作)

ISBN 978-7-5381-8545-4

I. ①模… II. ①张… ②乔… III. ①模块化—神经网络—结构分析 ②模块化—神经网络—结构设计 IV. ①
TP183

中国版本图书馆CIP数据核字(2014)第074948号



出版发行: 辽宁科学技术出版社

(地址: 沈阳市和平区十一纬路29号 邮编: 110003)

印刷者: 沈阳新华印刷厂

经销者: 各地新华书店

幅面尺寸: 185mm × 260mm

印 张: 13.5

字 数: 285千字

印 数: 1~1000

出版时间: 2014年4月第1版

印刷时间: 2014年4月第1次印刷

责任编辑: 郑 红 李伟民

特邀编辑: 王奉安

封面设计: 嵘 嵘

责任校对: 李淑敏

书 号: ISBN 978-7-5381-8545-4

定 价: 40.00元

联系电话: 024-23284360

邮购电话: 024-23284502

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

《辽宁省优秀自然科学著作》评审委员会

主任：

康捷 辽宁省科学技术协会党组书记、副主席

执行副主任：

黄其励 东北电网有限公司名誉总工程师

中国工程院院士

辽宁省科学技术协会副主席

副主任：

金太元 辽宁省科学技术协会副主席

宋纯智 辽宁科学技术出版社社长兼总编辑 编审

委员：

郭永新 辽宁大学副校长

陈宝智 东北大学安全工程研究所所长

刘文民 大连船舶重工集团有限公司副总工程师

李天来 沈阳农业大学副校长

刘明国 沈阳农业大学林学院院长

邢兆凯 辽宁省林业科学研究院院长

辽宁省科学技术协会委员

吴春福 沈阳药科大学校长

辽宁省科学技术协会常委

张兰 辽宁中医药大学附属医院副院长

王恩华 中国医科大学基础医学院副院长

李伟民 辽宁科学技术出版社总编室主任 编审

前 言

在人类几千年的文明发展史中，人们始终探索人类自身高级智能的奥秘。人们从认知科学、生物学、生物物理学、生物化学、医学、数学、信息与计算科学等领域进行广泛的研究和探索。在这个过程中逐渐形成了一门具有广泛学科交叉特点的学科——人工神经网络（Artificial Neural Network, ANN）。人们力图构建“人造”的生物神经细胞（人工神经元）和神经网络，在不同程度和不同层次上实现人脑神经系统的信息处理、学习、记忆、知识的存储和检索方面的功能。随着生产力发展水平的提高和实验手段的进步，人们在这个技术领域的各个方面都取得了巨大的成就。但是，由于人脑结构和运行机理无比的复杂，应该说到目前为止，人们对脑式信息处理的深层次机理和规律的认识还是相当粗浅的。

人工神经网络包括神经网络模型结构与神经网络学习算法，是在细胞的水平上模拟脑结构和脑功能的科学。人工神经网络模型结构与人工神经网络学习算法两者相互联系，人工神经网络模型结构是人工神经网络学习算法的前提，而人工神经网络学习算法是人工神经网络模型结构中神经信息运动演化的过程。人工神经网络的核心目标是在神经细胞的水平上，模拟生物神经网络的结构特征和生物神经信息的演化过程，构造人工神经网络模型结构，并建立能够在人工神经网络模型结构中有效的人工神经网络学习算法。

从20世纪40年代M-P神经元模型的提出开始，人工神经网络经历了艰辛的发展历程。1965年Minsky和Papert的《感知机》使得人工神经网络的发展停滞了10余年，直到80年代初期误差反向传播算法的提出，才使得人工神经网络的研究逐渐进入恢复期。时至今日，人工神经网络系统研究的重要意义已经得到广泛承认，涉及电子科学与技术、信息与通信工程、计算机科学与技术、电气工程、控制科学与技术等诸多学科，人工神经网络的应用

领域主要包括建模、时间序列预测、模式识别与智能控制等，并且在不断地拓展之中。可以说，人工神经网络是目前非线性科学和智能计算研究的主要内容之一，已经成为解决许多实际问题必要的技术手段。

本书系统地论述了模块化神经网络的主要理论、设计基础及应用实例，旨在使读者了解模块化神经网络的发展背景和研究对象，理解和熟悉模块化神经网络的基本原理和主要应用，掌握模块化神经网络结构模型和设计应用的方法，特别是模块化神经网络的结构设计理念、学习方法，为进一步研究和应用打下基础。为了便于读者理解，书中尽量避免繁琐的数学推导，加强了模块化神经网络在各个领域的应用实例，并在内容的选择和编排上注意到读者初次接触新概念的易接受性和思维的逻辑性。在编写的过程中，大部分内容是作者近几年的研究成果，部分内容取材于最近的国内外文献。本书的各章后面都附有较多的参考文献，以便读者进一步查阅。

本书可作为电子科学与技术、信息与通信工程、计算机科学与技术、电气工程、控制科学与技术等专业研究生和高年级本科生的教材，同时对有关专业领域的研究人员和工程技术人员也有重要的参考价值。

本书部分工作得到了国家自然科学基金重点项目(No.61034008)、辽宁省教育厅科学技术一般项目(L2013129)的支持。本书第1章和第5章由乔俊飞著，其余由张昭昭著。北京工业大学智能信息处理研究所李凡军、韩广、陈启丽、李薇等博士研究生为本书的编写做了大量工作；本书部分内容引用了北京工业大学智能信息研究所韩红桂和薄迎春博士的研究成果和领域内其他学者的一些研究成果；还得到辽宁工程技术大学电子与信息工程学院的大力支持。在此一并感谢！

由于作者的学术水平和思想境界有限，另外，人工神经网络的理论与应用还在不断地丰富和发展中，书中不妥之处在所难免，恳请广大读者斧正！

作 者

2013年11月16日

目 录

| | |
|---------------------------------|-----|
| 1 绪 论 | 001 |
| 1.1 引言 | 001 |
| 1.2 神经网络及其发展 | 002 |
| 1.3 人工神经网络结构设计发展现状 | 007 |
| 1.4 本书主要内容 | 014 |
| 2 感知器神经网络 | 018 |
| 2.1 引言 | 018 |
| 2.2 感知器神经网络分析 | 018 |
| 2.3 感知器神经网络学习算法 | 024 |
| 2.4 附录：数学基础 | 037 |
| 2.5 本章小结 | 041 |
| 3 前馈神经网络学习方法研究 | 043 |
| 3.1 引言 | 043 |
| 3.2 神经网络学习 | 043 |
| 3.3 快速下降算法 | 046 |
| 3.4 极速学习方法 | 053 |
| 3.5 实验与分析 | 057 |
| 3.6 附录：数学基础 | 065 |
| 3.7 本章小结 | 073 |
| 4 基于信息熵的前馈网络结构优化设计 | 074 |
| 4.1 引言 | 074 |
| 4.2 熵 | 077 |
| 4.3 基于改进拟熵的神经网络权衰减算法 | 081 |
| 4.4 基于神经网络复杂度的修剪算法 | 090 |

| | | |
|----------|----------------------------|------------|
| 4.5 | 自适应前馈神经网络结构设计 | 096 |
| 4.6 | 本章小结 | 103 |
| 5 | 模块化神经网络结构设计 | 105 |
| 5.1 | 引言 | 105 |
| 5.2 | 模块化神经网络 | 105 |
| 5.3 | 模块化神经网络的构架与实现 | 110 |
| 5.4 | 模块化神经性能分析 | 112 |
| 5.5 | 本章小结 | 117 |
| 6 | 多层协同模块化神经网络 | 119 |
| 6.1 | 引言 | 119 |
| 6.2 | 视觉神经系统 | 119 |
| 6.3 | 多层协同模块化神经网络 | 121 |
| 6.4 | 实验与分析 | 126 |
| 6.5 | 本章小结 | 131 |
| 7 | 局部互连BP神经网络 | 132 |
| 7.1 | 引言 | 132 |
| 7.2 | 局部互连BP神经网络结构 | 132 |
| 7.3 | RBF神经元参数的确定 | 135 |
| 7.3 | 局部互连BP神经网络的学习及网络结构设计 | 138 |
| 7.4 | 实验与分析 | 139 |
| 7.5 | 本章小结 | 145 |
| 8 | 动态自适应模块化神经网络结构设计 | 146 |
| 8.1 | 引言 | 146 |
| 8.2 | 动态自适应模块化神经网络结构 | 147 |
| 8.3 | 动态自适应模块化神经网络结构设计 | 150 |
| 8.4 | 动态自适应模块化神经网络学习算法 | 159 |
| 8.5 | 实验与分析 | 163 |
| 8.6 | 本章小结 | 178 |
| 9 | 具有群集属性的回声状态网络结构设计 | 180 |
| 9.1 | 引言 | 180 |
| 9.2 | 回声状态网络 | 181 |
| 9.3 | 具有Small World特性的回声状态网络结构设计 | 183 |
| 9.4 | 实验与分析 | 186 |
| 9.5 | 本章小结 | 195 |
| | 参考文献 | 196 |

1 绪论

1.1 引言

高性能、低成本、普适计算和智能化等是当前信息科学发展的主要方向，寻求新的计算与信息处理模型和物理实现是未来信息技术领域面临的重大挑战。神经网络理论的研究正是在与传统计算方法和理论挑战的过程中得以发展壮大，目前已经成为人工智能领域中最为活跃的研究方向之一。

神经网络理论是巨量信息并行处理和大规模并行计算的基础，神经网络不仅是高度非线性动力学系统，而且也是自适应自组织系统，可用来描述认知、决策和控制的智能行为。神经网络理论研究的核心问题是智能的认知和模拟，以人工方法模拟人脑的某些高级功能或者信息处理方式，毫无疑问将有助于加深对思维和智能的认识。20世纪80年代初崛起的神经网络理论，已经对认知和智力本质的基础研究乃至计算机产业都产生了空前的刺激和极大的推动作用。

人工神经网络(ANN)就是借鉴生物神经网络(BNN)结构和生物神经元工作机理，在一定程度上模拟人脑某些功能的信息处理系统。神经生理学研究结果表明，人脑是有许多神经元、神经元集群或者多个脑区相互连接而成的庞杂复杂网络，这个高度复杂的庞大网络是大脑进行信息处理和认知表达的生理基础。人脑在信息处理方面有着突出的优点，如学习、感知、模式识别以及在复杂环境中作出决策的能力等都是现有机器学习系统所无法比拟的，因此，研究智能理论与技术必须考虑脑的结构与功能。近年来，随着脑科学、分子学和细胞生物学的发展，使得大规模开展脑研究成为可能，而在人工智能领域，探求大脑的组织结构和运行机制，从模仿大脑智能的角度出发，寻求新的信息处理方法也成为该领域的一个重要研究方向。国家中长期科学和技术发展规划纲要中也明确指出：脑功能的细胞和分子机理，学习记忆和思维等脑高级认知功能的过程及其神经基础，脑信息表达与脑式信息处理系统，人脑与计算机对话等研究方向属于当前的科学前沿问题。

近几十年来，脑神经科学家为了阐明脑神经系统不同层次的活动规律和特征，做了大量的实验性研究。Duncan 等对信息在脑内的处理方式的实验结果不支持大脑中只有一个信息处理中心。倘若大脑中只有一个信息处理中心，那么当大脑处理多个信息流时，可能会因争夺信息处理中心资源而发生相互干扰的现象。实验结果更加支持平行分布加工(PDP)模型，即大脑中存在着许多信息处理单元，每个信息处理单元处理一些

具体简单的任务，当一个信息处理单元激活时，它可以沿着一个网络激活或抑制其他信息单元的加工过程。大量的实验研究均支持以诺贝尔奖获得者 Edelman 为首的学派提出的动态核心假设 (DCH)，即信息处理时不同的信息对应不同的神经集群；信息处理在空间维度上涉及大脑多区域的协调参与，在时间维度上神经集群的大小时刻都在变化；不同皮层区域的同步神经激活是信息整合的机制。当前神经网络结构设计领域研究较多的是单一的神经网络，而信息处理的丰富性在很大程度上依赖于较高级神经网络的大小和组织性。从仿生学来说，当分子、细胞、神经元回路分别组成新的结构后（小的神经网络组成新的更大的网络），其意义绝不是生物学上简单的叠加，因此，继分子生物学之后，强调整体和整合的观点来研究脑正逐渐成为一个新的趋势。

人类行为的极端复杂性，源于脑的结构和功能在自然界中的长期进化，因此人脑具有今天的形态，是进化的结果，是自组织的结果。从本质上说，神经网络处理信息的过程就是不确定性减少的过程，换句话说，也就是神经网络系统自组织性增强的过程，而 Shannon 信息论告诉我们：衡量一个系统不确定性最本质的参数就是系统的熵。因此，以信息熵为准则，从系统宏观的角度深入研究类脑神经网络结构设计的基础理论，既可以推动神经动力学的发展，也可以为新一代类脑计算机的产生提供理论支持。

人工神经网络作为智能科学的重要组成部分，已经成为脑科学、认知科学、计算机科学、数学和物理学等学科共同关注的焦点。其应用研究已经渗透到工业、农业、国防、航空等领域，并且在信号处理、智能控制、模式识别、图像处理、非线性优化、知识处理等方面取得了令人鼓舞的进展。神经网络实际应用的需求驱动了其理论研究的发展。

1.2 神经网络及其发展

1.2.1 神经网络的定义

人工神经网络简称神经网络，是由大量的简单处理单元——人工神经元 (AN) 相互连接而形成的一个具有高度非线性和并行的自适应信息处理系统。神经网络的非线性动力学行为，打破了传统串行处理计算机的局限，能够并行分布式存储和处理信息。尽管单个神经元的结构和功能非常简单，但由大量简单神经元组成的神经网络系统却具有非常强大的处理问题的能力。神经网络实质上是对人脑功能的一种模拟和简化，具有学习、记忆、联想、类比、计算的能力，是现代神经科学研究与工程技术相结合的产物。神经网络理论的开创与发展，对智能科学和信息处理技术的发展产生了重大的影响和积极的推动作用。

1943年，美国神经生理学家 McCulloch 和 Pitts 提出的第一个神经网络模型 M-P 模型，开创了微观人工智能的研究工作，奠定了人工神经网络发展的基础。人工神经网络经过几十年的发展，无论是在理论研究还是在工程应用方面都取得了丰富的研究成果。

1.2.2 神经网络的功能

神经网络是通过对人类大脑结构和功能的模拟建立起来的一个非线性、自适应的高级信息处理系统。它是现代神经科学研究与工程技术应用相结构的产物。神经网络具有强大的计算能力,是由其本身大规模的并行分布式结构和较好的学习能力以及由此延伸而来的泛化能力决定的。神经网络具有非线性、分布处理、容错性、自适应性等显著特点。

1.2.2.1 非线性

神经网络的单个处理单元——人工神经元可以是线性或非线性的,但由大量神经元互连而成的神经网络系统本身却是非线性的。另外,非线性是一种分布于整个网络的特殊本质。因此,神经网络具有很强的非线性映射能力,并且理论研究已经证明,一个三层的神经网络能够以任意精度逼近非线性系统。

1.2.2.2 分布处理

神经网络是为模拟大脑的结构和功能而建立的一种数学模型,大量的人工神经元相互连接成一个高度并行的非线性动力学系统。尽管单个人工神经元的功能非常简单,但大量神经元的并行行为使得整个网络呈现出很强的信息处理能力。神经网络中的信息的存储体现在神经元之间互相连接的并行分布结构上,进而使得信息的处理必然采用大规模的并行分布处理方式进行,即神经网络中信息的存储和处理是在整个网络中同时进行的,信息不是存储在神经网络中的某个局部,而是分布在网络的所有单元之中。一个神经网络可以存储多种信息,而神经元连接权值中只存储多种信息的一部分。神经网络的内在结构,决定了其信息的存储和处理在时间和空间分布上都是并行的。神经网络中的数据及处理是全局的而不是局部的。

1.2.2.3 容错性

神经网络善于联想、概括、类比和推广,加之神经网络信息储存和处理的并行特性,使得神经网络在以下两个方面表现出较好的容错性:一方面,由于神经网络的信息采用分布式存储,信息存储在各个神经元的连接权值中,当神经网络中某一个神经元或者连接权值出现问题时,局部的改变不会影响神经网络的整体非线性映射。这一点与人的大脑中每时每刻都有神经细胞的正常死亡和分裂,但不会影响大脑的整体智能类似;另一方面,当神经网络的输入信息模糊、残缺或不完整时,神经网络能够通过联想、记忆等方式实现对输入信息的正确识别。

1.2.2.4 自组织性

神经网络既是高度非线性动力学系统,又是自适应组织系统。神经网络的自组织性包括学习自组织能力和结构自组织能力,是指神经网络系统能够通过改变系统本身的某些性能以适应外界环境的变化。自组织性是神经网络的一个重要特性。神经网络的学习自组织能力表现在:当外界的环境发生改变时,即神经网络的输入有变化时,神经网络通过一段时间的学习和训练,能够自动调整网络的结构和参数,从而得到期望的输出。

也就是说，神经网络可以在学习过程中不断地完善自身，具有创新的特点。神经网络的结构自组织特性表现在：神经网络在接收外部激励后可以根据一定的规则通过对神经网络权值的调整以及神经元的增减来重新构建新的神经网络。神经网络不但可以处理各种变化的信息，而且在学习阶段能够根据流过神经网络的外部 and 内部信息对自身的结构进行调整，从而改变神经网络本身的非线性动力学特性，以适应外界环境的变化。

1.2.3 神经网络的发展

神经网络是一门活跃的边缘性交叉学科，是基于人脑的结构和功能而构建的一种智能信息处理系统，它的中心问题是智能的认知和模拟。从解剖学和生理学来看，人脑对信息的处理不同于传统的冯·洛伊曼式计算机，它具有“认知”、“意识”和“感情”等高级脑功能。因此，神经网络目前已成为脑科学、神经科学、认知科学、心理学、计算机科学、数学和物理学等学科关注的焦点，且在过去半个多世纪中一直在智能信息处理的程序化设计中占主导地位。神经网络的应用和发展不但推动了神经动力学本身的发展，而且为智能信息处理提供了新的方法，有可能为信息科学带来革命性的变化。

虽然目前神经网络的理论和应用研究处在前所未有的热潮中，但它的发展却并不是一帆风顺的。从研究时间递推的角度来看，神经网络的研究主要经历了兴起与高潮、萧条、稳步发展的较为曲折的道路，如图 1-1 所示。

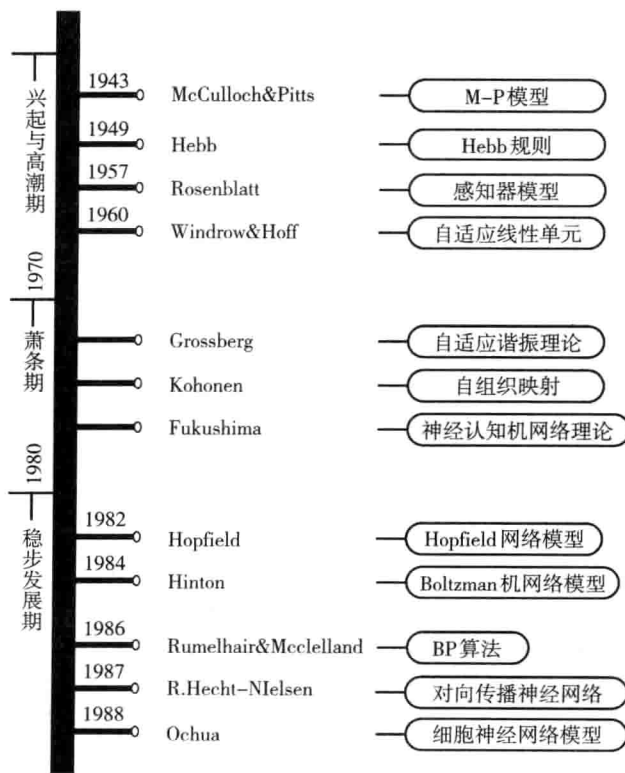


图 1-1 人工神经网络的发展

1.2.3.1 兴起与高潮期（1940—1970年）

1943年，心理学家McCulloch和数理逻辑学家Pitts提出了M-P模型，这是第一个用数理语言描述脑信息处理过程的模型，虽然神经元的功能比较弱，但它为以后的研究工作奠定了基础。1949年，心理学家Hebb提出突触连接可变的假设，根据此假设提出的学习规则为神经网络的学习算法提供了依据。1957年，计算机科学家Rosenblatt提出著名的感知模型，该模型包含了现代计算机的一些原理，是第一个完整的人工神经网络模型，首次把神经网络研究引入工程实现。1960年，Windrow提出了自适应线性单元，主要用于自适应系统。这些简单的神经网络中体现出了许多特性，如并行处理、分布式存储、连续计算、可学习性等，因此引起许多研究者的兴趣。从此，神经网络的研究进入第一个高潮期。

1.2.3.2 萧条期（1970—1980年）

人工智能的创始人之一，美国麻省理工学院的Minsky教授潜心研究数年，对以感知器为代表的人工神经网络系统的记忆功能局限性从数学的角度进行了深入的研究，并于1969年出版了颇具影响的《感知器：计算几何引论》一书。该书中对感知器分析的结论是悲观的，甚至认为其毫无科学价值可言。Minsky在学术界的地位和影响力使得其后若干年内，神经网络领域的研究一直处于低潮期。造成该结局的另一个重要原因是，传统的冯·洛伊曼型数字计算机当时正处于发展的全盛时期，人工智能得到迅速的发展并取得显著的成就，整个学术界陶醉于计算机的成功之中，从而掩盖了发展新型计算模型和人工智能技术的必要性和迫切性。由于上述原因，当时相当多的研究者也都认为神经网络的研究前途渺茫，因而放弃了在该领域继续探索的努力。

虽然在此后的十几年中，神经网络的研究陷入了暗淡的境地，大多数相关的研究人员把注意力转向人工智能。然而，在此期间，仍有少数学者继续致力于神经网络的研究。如Carpebter和Grossberg提出了自适应谐振理论，Kohonen提出了自组织映射，Fukushima提出了神经认知机网络理论，Amari则致力于神经网络有关数学理论的研究，Anderson等提出了BSB模型等。这些工作都为神经网络研究的发展奠定了理论基础。

1.2.3.3 稳步发展期

美国生物物理学家Hopfield于1982年、1984年在美国科学研究院刊发表了两篇文章有力地推动了神经网络的研究，引起了研究神经网络的又一次热潮。1982年，他提出了一个新的神经网络模型——Hopfield网络模型。在该网络模型中，首次引入了网络能量函数的概念，并给出网络收敛性的判定依据。1984年，他又提出了网络模型实现的电子电路，为神经网络的工程实现指明了方向。Hopfield的研究成果开拓了神经网络用于联想记忆优化计算的途径，并为神经计算机的研究奠定了基础。1984年Ackley等将模拟退火算法引入到神经网络中，提出了Boltzmann机网络模型，Boltzmann网络算法为神经网络优化计算提供了一个有效的方法。1986年，McClelland等提出了误差反向传播算法（Error Back Propagation, BP），成为迄今为止影响最大的一种神经网络学习算

法。1987年美国神经计算机专家Robert提出了对向传播神经网络,该网络具有分类灵活,算法简练的优点,可用于模式分类、函数拟合、统计分析和数据压缩等领域。1988年Reid等提出了细胞神经网络模型,它在视觉初级加工上得到了广泛应用。上述这些研究成果为神经网络的研究与发展起到了推波助澜的作用,人工神经网络开始步入了稳步发展的时期。在经历了艰辛的历程后,人工神经网络取得了很大的成就,然而,人工神经网络还有很长的路要走。

人脑是一个功能十分强大、结构非常复杂的信息处理系统,随着信息论、控制论、生命科学、计算机科学的发展,人们越来越惊异于大脑的奇妙,人类大脑信息处理机制对人类自身来说,仍然是一个谜,要揭示人脑信息处理的奥秘需要神经科学家、计算机科学家、微电子科学家、数学家等专家共同的努力,对人类智能行为不断探索研究,为人工神经网络发展提供丰富的理论源泉。另外,通过哲学思想和多种自然科学的深层结合,逐步孕育出探索人类思维本质和规律的新方法。纵观神经网络的发展历程,没有相关学科的贡献和不同学科专家的竞争与协同,神经网络就没有今天。当然,人工神经网络在各个学科领域应用的研究反过来也推动了其他学科自身的完善与发展。总之,人工神经网络研究是人类智能研究的一个重要组成部分,对人类智能的发展,特别是计算智能的发展起到了重要的推动作用。

1.2.4 神经网络的应用

神经网络属于连接主义,是一种并行分布式的信息处理系统,采用不同于传统的人工智能和信息处理技术,突破了传统的串行处理的数字电子计算机的局限,可用于解决知识表达、联想记忆、推理学习以及一些复杂系统问题。神经网络的实用性在于其可以通过观测值推导出对应的数学模型,尤其是当数据或任务的复杂性使得采用手动推导不切合实际时,采用神经网络建模推导函数关系十分有效。由于神经网络优点众多,神经网络已经广泛应用于众多领域,在工业、农业、商业、国防和其他科学技术领域都得到了广泛而成功的应用。神经网络的应用领域主要包含系统辨识和控制(车辆控制、过程控制、工厂自动化机械控制等),量子化,博弈和决策制定(国际象棋),模式识别(雷达系统、人脸识别、语音识别、目标识别、边缘检测、视觉索引引擎等),序列识别(姿态、语音、手写体辨识),语音合成(文本朗读——NETtalk等),医学诊断、数据挖掘、数据压缩(语音信号、图像等),可视化以及垃圾邮件过滤等,如图1-2所示。

根据神经网络应用性质的不同,其应用领域主要体现在以下两个方面。

1.2.4.1 理论分析

理论分析主要是应用神经网络的函数逼近能力,包含时间序列预测、适应度逼近和建模。理论研究表明,给定一个合适结构的人工神经网络,能够以任意精度逼近任意的非线性映射。这一点为人工神经网络的应用提供了理论依据和保障。在人工神经网络的应用上,很多方面就是应用神经网络的这一特性进行的。最基本的是对非线性系统的黑箱建模,神经网络可以根据系统的输入—输出数据对,经过学习对数据之间的非线性关

系进行映射。这一方面广泛引用于时间序列预测、函数的逼近、系统的建模等方面。

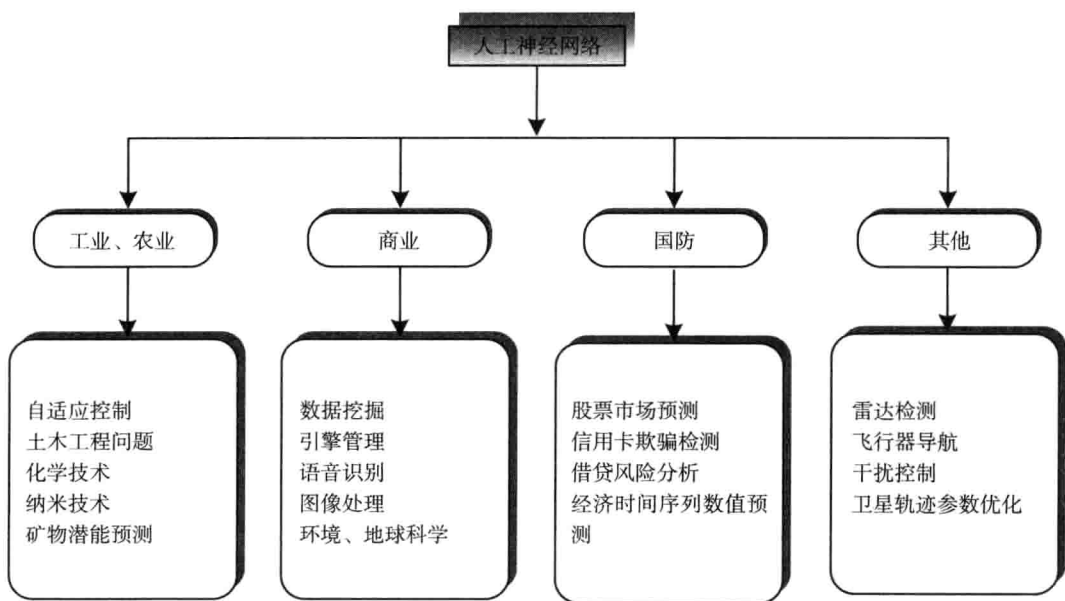


图 1-2 神经网络的应用

1.2.4.2 工程实践

由于神经网络具有良好的模式辨识和数据趋势的能力，能够满足预测的需求，因此神经网络已经广泛地应用于实际的工业应用和商业问题中。例如，销售预测、工业过程控制、视觉导航、轨迹控制、消费者研究、数据验证、风险管理等。此外，对于给定的特殊问题，神经网络业能应用于下列领域，如说话者在交流中的分辨、疾病诊断、三维目标识别、手写体识别等。

1.3 人工神经网络结构设计发展现状

1.3.1 前馈神经网络结构设计研究现状

由于本书的模块化神经网络中的子神经网络拟采用前馈神经网络，所以首先对前馈神经网络结构设计研究现状进行介绍。

前馈神经网络是目前应用最为广泛的神经网络，前馈神经网络得以成功应用的前提是如何针对具体的问题确定合适的神经网络结构。就一个训练样本数一定的具体问题而言，由于训练样本不可避免地包含误差，若神经网络结构冗余，那么在神经网络训练的后期，这些误差会影响到整个神经网络训练的收敛方向，从而造成全局最优点的偏离和泛化能力（generalization ability）或者推广能力的降低，这种现象被称之为过训练或过拟合（overtraining/overfitting）。一般情况下，过拟合是由于神经网络的自由度大于训练数据的信息自由度造成的。另外，规模冗余的神经网络除了容易出现过拟合现象外，还

容易导致训练陷入局部极小。在确定神经网络结构的问题上, 尽管目前存在一些定性的指导原则: 在没有其他先验知识的情况下, 与给定训练样本一致的规模最小的神经网络就是最好的神经网络。从数学的角度看, 该原则相当于在样本点偏差允许范围内, 用最平滑的函数去逼近未知的非线性映射。但对于一特定问题而言, 至今还没有一种在训练之前就能根据训练样本确定最优神经网络结构的方法。事实上, 给定训练样本后, 最佳神经网络结构选择问题是一个NPC问题。理论上讲, 调整神经网络的规模使之刚好与给定的训练样本实现“匹配”(此时的神经网络结构被认为是最优神经网络结构), 是提高神经网络泛化能力的一个有效方法。由于神经网络的泛化能力是学习后的神经网络对测试样本或工作样本(没有参与神经网络训练的样本)作出正确反应的能力, 所以没有泛化能力的神经网络没有任何实用价值。正因为神经网络的结构直接决定了神经网络的泛化性能、动力学行为、神经网络的稳定状态和神经网络状态空间的大尺度结构, 所以前馈神经网络结构设计的理论和方法一直是神经网络研究领域的一个难点和热点。

常用的前馈神经网络结构设计方法有正则化方法、修剪方法、增长型方法和修剪加增长型方法。

在神经网络刚刚兴起的1995年, Bishop指出, 对大多数的应用来说, 采用正则化技术确定神经网络结构是比较好的方法。所谓正则化方法, 就是在神经网络的训练代价函数中增加一个惩罚项。增加了惩罚项的代价函数为

$$E = E_t + \lambda E_c \quad (1-1)$$

其中 E_t 为网络训练误差, E_c 为惩罚项, λ 为一个正的惩罚因子。该惩罚项能起到使神经网络权连接衰减的作用。应用正则化方法的难点在于如何确定一个合适的惩罚因子使得惩罚项能够与神经网络映射的缩放属性相匹配。最优惩罚因子一般情况下可以通过交叉测试(CV)的方法获得, 然而采用交叉测试法计算成本十分昂贵, 特别是当惩罚项中包括多个惩罚属性的时候。采用Bayesian方法可以使惩罚因子在神经网络的训练过程中自适应调整, 从而避免了采用交叉测试法来确定惩罚因子所带来的大量的时间消耗, 然而, Bayesian方法是建立在神经网络权连接的后验概率分布为高斯分布的假设之上, 该假设没有考虑神经网络训练代价函数存在多个局部极小点的事实。

修剪算法是一种自顶向下的神经网络结构设计方法, 神经网络学习时先从一个初始规模较大的神经网络开始, 在网络学习的过程中逐渐删除神经网络中冗余的权连接或隐节点, 最终形成一个既能较好地学习当前任务又具有一定泛化能力的精简神经网络结构。由于修剪过程中神经网络的规模呈现递减趋势, 因此神经网络的学习速度将逐渐变快, 学习任务时所需的存储空间将逐渐变小。修剪算法符合工程上“宁多勿少”的思想, 人们会首选一个神经网络规模相对较大但能很好地学习当前任务且具有一定泛化能力的神经网络。同时, 神经生物学研究表明, 在人类的智力发展阶段, 大脑中确实存在修剪现象, 那些不常用的神经元或神经元之间的连接将会逐渐失去活性而被删除, 因此修剪算法也具有一定的仿生学基础。

经典的修剪型算法如OBD算法, 该算法的最初思想来自Le Cun等根据信息论利用

二次寻优的方法 (second order derivative) 对神经网络的代价函数进行调整。OBD算法中神经网络代价函数泰勒展开式为:

$$\delta E = \delta W^T g + \frac{1}{2} \delta W^T H \delta w + L \quad (1-2)$$

式中, $g = \delta E / \delta W$, 为代价函数 $E(w)$ 的梯度; $H = \delta^2 E / \delta^2 w$, 为代价函数的二阶导数, 称为 Hessian 矩阵。为简化运算, OBD 算法在每次删减神经元权连接时总是假设此时神经网络代价函数已经训练到一个局部最小点, 基于此假设, 此时神经网络代价函数泰勒展开式的一次项为 0, 舍弃所有高次项而只取 2 次项, 则式 (1-1) 简化为:

$$\delta E = \frac{1}{2} H_{jj} w_j^2 \quad (1-3)$$

式中, H_{jj} 为对角阵, $j=1, \dots, j$, j 为此时神经网络隐层节点数。在神经网络训练的过程中, OBD 算法根据性能指标的要求不断调整神经网络参数的同时可以根据式 (1-2) 求得每组神经元权值参数对整个神经网络的影响, 通过删除具有最小特征值的权连接达到精简神经网络结构的目的。其算法计算时间复杂度为 $O(M^2 N)$, M 为神经网络中连接权个数, N 为训练样本数。

在 OBD 算法的基础上, Hassibi 等提出了 OBS 算法。OBS 算法也是利用二次寻优的方法对神经网络的代价函数进行调整。在神经网络的学习过程中 H 不需要假设为对角矩阵, 而是对公式 (1-3) 进行优化处理:

$$e_j^T (\delta w + w) = 0 \quad (1-4)$$

式中, e_j 是向量, 向量中第 j 位为 0, 这样 OBS 的寻优问题就转化为:

$$\min_{1 \leq j \leq J} \{ \min(\frac{1}{2} \delta w^T \frac{\delta^2 E}{\delta^2 w} j e_j^T (\delta w + w) = 0) \} \quad (1-5)$$

与 OBD 算法不同的是, OBS 算法不仅删除影响较小的神经元权值, 而且还提供了一种补偿机制来弥补被删除节点或权连接对神经网络输出的贡献, 从而节约了大量重新训练神经网络的时间。OBS 算法也可以同时删除多个神经元权值, 与 OBD 算法一次只能删除一个神经元权值相比, 其神经网络结构调整速度也有所提高。同时, Hassibi 等人还给出了修剪后神经网络的稳定性证明。

其后, 许多修剪算法都在 OBD 和 OBS 的基础上继续发展并在许多方面超越了 OBD 和 OBS 算法。

李倩等基于进化算法和局部搜索算法两类策略的特点和不足, 提出了混合修剪算法 HAP (Hybrid Algorithm of Pruning, HAP)。HAP 算法首先联合遗传算法 (Genetic Algorithm, GA) 和反向传播算法的不同优势完成神经网络结构和权重进化的初级阶段; 然后应用多重修剪策略进一步简化、确定神经网络结构。HAP 算法在寻优能力、精简神经网络结构、保证稳定性方面均有明显优势, 适合大规模人工神经网络的结构优化问题, 其算法时间复杂度为 $O(MN^2)$, M 为权连接个数, N 为训练样本数。

乔俊飞等在 OBS 的基础上, 通过直接删除冗余的隐层神经元实现神经网络结构优化。该快速修剪算法与常规 OBS 算法相比, 具有更简单的神经网络结构和更快的学习