



“十二五”江苏省高等学校重点教材项目

混合神经网络技术

(第二版)

田雨波 陈 风 张贞凯◎著



科学出版社



“十二五”江苏省高等学校重点教材项目(编号:2014-1-026)

混合神经网络技术

(第二版)

田雨波 陈 风 张贞凯 著

科学出版社

北 京

内 容 简 介

本书在论述神经网络基本概念和基本原理的基础上,重点介绍混合神经网络技术,同时,介绍各种混合神经网络技术在电磁建模和优化问题中的应用。全书共 11 章,内容主要包括神经网络的基本概念、基础知识、BP 神经网络、RBF 神经网络、Hopfield 神经网络、粒子群神经网络、模糊神经网络、混沌神经网络、小波神经网络、知识神经网络和神经网络集成。

本书可供从事神经网络理论与技术、计算电磁学、电磁场工程等领域研究和开发工作的科技人员和高校教师参考阅读,也可作为高等院校相关专业的高年级本科生和研究生的教学用书。

图书在版编目(CIP)数据

混合神经网络技术/田雨波,陈风,张贞凯著.—2 版.—北京:科学出版社,2015.10

ISBN 978-7-03-045932-9

I. ①混… II. ①田…②陈…③张… III. ①人工神经—神经网络
IV. ①TP18

中国版本图书馆 CIP 数据核字(2015)第 240203 号

责任编辑:孙 芳 张艳芬 / 责任校对:桂伟利

责任印制:张 倩 / 封面设计:蓝 正

科学出版社出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

新科印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2009 年 6 月第 一 版 开本:720×1000 1/16

2015 年 10 月第 二 版 印张:18 1/2

2015 年 10 月第二次印刷 字数:360 000

定价:98.00 元

(如有印装质量问题,我社负责调换)

第二版前言

本书第一版于 2009 年出版,2014 年被评为“十二五”江苏省高等学校重点教材。

本次修订吸收了混合神经网络技术发展的最新成果,同时融入了作者近年的科研成果。本次修订在保持第一版基本框架的基础上做了以下调整:

(1) 增加了广义回归神经网络一节,并给出了广义回归神经网络在宽带信号到达角方向估计上的应用。作为 RBF 神经网络的一个重要分支,广义回归神经网络具有训练效果好、非线性映射能力强、收敛速度快、计算量小等优点。

(2) 增加了基于图形处理单元的粒子群神经网络并行实现一节,主要考虑混合神经网络训练时间较长,通过并行计算可以大幅减少网络的训练时间,提高计算效率;同时,基于图形处理单元的并行计算也是目前并行计算的发展方向之一。

(3) 增加了知识神经网络一章。知识神经网络是专家知识和神经网络相互相合的一种神经网络模型,该模型在保留传统神经网络模型优点的同时,大大简化了传统神经网络的结构,并且不受训练数据采样范围的限制,因而其推广性大大优于传统神经网络,已经得到了广泛应用。

(4) 对神经网络集成一章进行了完善。

本书共 11 章,分为两部分:第 1 章~第 5 章为基本神经网络部分,主要包括神经网络的基本概念、基础知识、BP 神经网络、RBF 神经网络、Hopfield 神经网络;第 6 章~第 11 章为混合神经网络部分,主要包括粒子群神经网络、模糊神经网络、混沌神经网络、小波神经网络、知识神经网络和神经网络集成。

本书在修订过程中参阅了大量的国内外文献,在此对相关作者深表感谢。

本次修订工作是在“十二五”江苏省高等学校重点教材项目(2014-1-026)、江苏省“青蓝工程”中青年学术带头人项目、国家自然科学基金项目(61401179)、江苏科技大学研究生院出版基金的资助下完成的,在此对上述资助单位表示诚挚谢意。同时,对在本书修订过程中给予作者大力支持的南京理工大学陈如山教授、南京大学伍瑞新教授、东北大学李鸿儒教授、南京邮电大学徐荣青教授和江苏科技大学朱志宇教授等表示衷心的感谢。

第一版前言

人工神经网络是在对人脑认识的基础上,以数学和物理方法及从信息处理的角度对人脑生物神经网络进行抽象并建立起来的某种简化模型,它是计算智能和机器学习研究的最活跃的分支之一。近年来,神经网络在理论研究、实现技术和应用研究等方面取得了引人注目的成果,为此,国内外已经出版了有关神经网络方面的著作、教材、论文集等。信息科学与包括生命科学在内的其他智能技术的相互交叉、相互渗透和相互促进是现代科学技术发展的一个显著特点。神经网络与各种智能信息处理方法有机结合具有很大发展前景,如与模拟退火算法、遗传算法、粒子群算法、模糊理论、混沌理论和小波分析等相结合,即形成所谓的“混合神经网络技术”,目前已经成为一大研究热点。人们希望通过这些理论和算法与神经网络相互混合,获得具有柔性信息处理功能的系统,但关于这方面的书籍并不是很多,这正是本书的主要特点之一。

计算电磁学是在 20 世纪 60 年代随着电子计算机技术的发展而诞生的,它是在电磁学、计算数学和计算机科学的基础上产生的边缘交叉学科。计算电磁学实质上是以电磁场理论为基础,以高性能计算技术为手段,运用计算数学提供的各种方法解决复杂电磁场理论和工程问题的应用科学。经过几十年的发展,计算电磁学内容已经非常丰富,影响非常深广,以致所有与电磁场相关的领域都因其发展而受益,其中,不少领域由于运用了计算电磁学的方法而使其面貌完全改观。然而,对于复杂的电磁系统,对其进行严格的电磁仿真耗时而费力,在保证计算精度的情况下对其进行快速而精确的建模和优化必将成为计算电磁学的发展趋势,而这在一定程度上又是建立在神经网络技术基础之上的。本书在讲述混合神经网络技术的同时,重点给出各种混合神经网络在电磁学方面的应用,这也是本书的另外一个特点。

本书是作者从事神经网络理论与技术和计算电磁学建模及优化的教学和科研工作的系统总结,并从国内外相关文献资料中提取最主要的理论及成果,力图反映最新的研究动态,清楚阐述混合神经网络技术及这些技术在电磁问题的数值仿真、高效建模和优化设计中的具体应用。全书共 12 章。第 1 章介绍神经网络的基本概念、基本功能、基本性质及性能指标、研究内容、发展趋势等,同时,综述神经网络在电磁方面的应用。第 2 章介绍神经网络的基础知识,包括神经网络的基本模型、训练与学习、泛化能力等。第 3 章介绍 BP 神经网络,它是最具代表性的前馈神经网络模型之一,其中,包括 BP 神经网络的网络结构、学习算法、应用要

点、不足及改进等,同时,应用 BP 神经网络对微带天线进行结构设计。第 4 章介绍 RBF 神经网络,它是另外一种具有代表性的前馈神经网络模型,包括 RBF 神经网络的网络结构、学习算法、网络特点等,并与 BP 神经网络作了对比说明。第 5 章介绍 Hopfield 神经网络,它是最常见的反馈神经网络,包括神经动力学和 Lyapunov 定理、连续 Hopfield 神经网络、离散 Hopfield 神经网络等。第 6 章介绍随机神经网络,包括 Boltzmann 机、神经网络的随机训练、模拟退火算法等。第 7 章介绍遗传神经网络,包括遗传算法基本原理及遗传神经网络的实现等,同时,应用遗传 RBF 神经网络解决了自适应波束形成问题。第 8 章介绍粒子群神经网络,包括粒子群优化算法基本原理及粒子群神经网络的实现等。第 9 章介绍模糊神经网络,包括模糊理论基本知识及模糊神经网络的实现等,同时,应用模糊神经网络进行了波导匹配负载设计和微带天线谐振频率计算。第 10 章介绍混沌神经网络,包括混沌理论基本知识及混沌神经网络的实现等,同时,应用混沌神经网络进行了移动通信系统信道分配和自适应雷达目标信号处理。第 11 章介绍小波神经网络,包括小波分析基本知识及小波神经网络的实现等,同时,应用小波神经网络解决了飞机图像识别和微带不连续问题。第 12 章介绍神经网络集成,包括神经网络集成的基本概念、实现方法、理论分析等,同时,应用神经网络集成进行了股市预测、肺癌诊断和谐振频率计算。书末附录给出本书中应用的大部分程序,方便读者理解及使用。另外,本书在编著过程中参阅了大量的国内外文献,在此对作者深表感谢。

本书是在江苏省“青蓝工程”优秀青年骨干教师项目、江苏省高校自然科学基金基础研究项目(07KJB510032)、江苏科技大学研究生部出版基金的资助下完成的,在此对上述资助单位表示诚挚谢意。同时,对在本书创作过程中给予作者大力支持的江苏科技大学电子信息学院的领导及各位同仁由衷地表示感谢。

由于作者水平有限,不妥之处在所难免,敬请读者批评指正。

作 者

2009 年 3 月

目 录

第二版前言

第一版前言

第 1 章 绪论	1
1.1 神经网络的概念与分类	1
1.2 神经网络的基本特征和基本功能	3
1.3 神经网络的基本性质、优点及应用	5
1.4 神经网络的性能指标及研究内容	6
1.5 神经网络的发展简史、存在问题及发展趋势	7
1.6 神经网络的电磁应用	11
参考文献	12
第 2 章 基础知识	17
2.1 神经网络模型	17
2.2 神经网络的训练和学习	23
2.3 神经网络的泛化能力	27
2.4 神经网络训练用样本	28
参考文献	33
第 3 章 BP 神经网络	34
3.1 BP 神经网络结构	34
3.2 BP 学习算法	36
3.3 BP 神经网络应用要点	38
3.4 BP 算法的不足及改进	42
3.5 BP 神经网络应用	48
参考文献	57
第 4 章 RBF 神经网络	60
4.1 网络结构和工作原理	60
4.2 网络的生理学基础和数学基础	62
4.3 常用的学习算法	66
4.4 网络的特点及注意事项	71
4.5 RBF 神经网络应用	72

4.6 RBF神经网络与BP神经网络的比较	78
参考文献	79
第5章 Hopfield神经网络	82
5.1 Hopfield神经网络简介	82
5.2 神经动力学	82
5.3 Lyapunov定理	83
5.4 连续Hopfield神经网络	85
5.5 离散Hopfield神经网络	88
5.6 Hopfield神经网络应用	90
5.7 Hopfield神经网络特点	92
参考文献	92
第6章 粒子群神经网络	94
6.1 粒子群优化算法	94
6.2 粒子群神经网络原理及实现	108
6.3 粒子群神经网络应用	113
参考文献	125
第7章 模糊神经网络	128
7.1 模糊理论	128
7.2 模糊神经网络原理及实现	139
7.3 模糊神经网络应用	151
参考文献	169
第8章 混沌神经网络	173
8.1 混沌理论	173
8.2 混沌神经网络原理及实现	190
8.3 混沌神经网络应用	201
参考文献	214
第9章 小波神经网络	217
9.1 小波分析概述	217
9.2 小波神经网络原理及实现	223
9.3 小波神经网络应用	232
参考文献	242
第10章 知识神经网络	244
10.1 知识神经网络基本概念	244
10.2 基于局部保持投影的知识神经网络原理及实现	249

10.3 基于局部保持投影的知识神经网络应用	251
参考文献	255
第 11 章 神经网络集成	256
11.1 神经网络集成基本概念	256
11.2 神经网络集成实现方法	257
11.3 神经网络集成的应用	268
参考文献	284

第 1 章 绪 论

人类具有高度发达的大脑,大脑是思维活动的物质基础,而思维是人类智能的集中体现。长期以来,人们想方设法了解人脑的工作机理和思维本质,向往构造出人工智能系统来模仿人脑的功能,其中的一个重要成果就是人工神经网络(artificial neural networks, ANN)。本章主要讲述人工神经网络的概念、基本性质、基本功能和应用,以及人工神经网络的性能指标、研究内容和发展趋势,最后介绍人工神经网络的电磁应用。

1.1 神经网络的概念与分类

1.1.1 神经网络的概念

人工神经网络又称神经网络(neural network)、人工神经系统(artificial neural systems)、自适应系统(adaptive systems)、自适应网(adaptive networks)、连接模型(connectionism)、神经计算机(neurocomputer)等,它是在对人脑认识的基础上,以数学和物理方法及从信息处理的角度对人脑生物神经网络(biological neural network, BNN)进行抽象并建立起来的某种简化模型。它是对人类大脑系统特性的一种描述,是由多个非常简单的处理单元彼此按某种方式连接而形成的计算机系统,该系统是靠其状态对外部输入信息的动态响应来处理信息的。简单地讲,人工神经网络是一个数学模型,可以用电子线路来实现,也可以用计算机程序来模拟,是人工智能研究的一种方法。

1.1.2 神经网络的分类

到目前为止,神经网络已经有几十种不同的模型,按照不同的原则,可以对神经网络进行不同的分类,通常有以下 5 类^[1~4]:

- (1) 按照网络的结构进行分类:前馈网络、反馈网络 and 自组织网络。
- (2) 按照学习方式进行分类:有导师学习网络和无导师学习网络。
- (3) 按照网络的性能进行分类:连续型网络和离散型网络、随机型网络和确定型网络。
- (4) 按照突触性质进行分类:一阶线性并联网和高阶非线性并联网。
- (5) 按照对生物神经系统的层次模拟分类:神经元层次模型、组合式模型、网

络层次模型、神经系统层次模型和智能型模型。

在人工神经网络的设计与应用过程中,人们较多地考虑神经网络的互联结构,包括 5 种典型结构^[5],如图 1.1 所示。

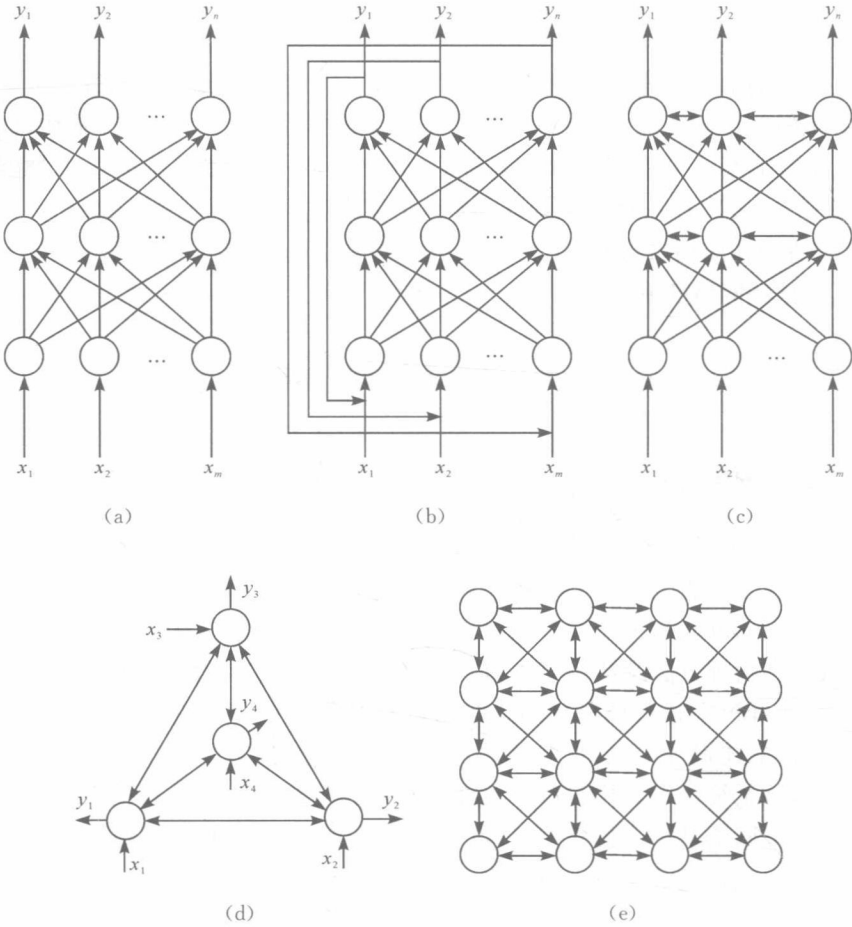


图 1.1 神经网络的拓扑结构

(1) 前馈网络。前馈网络中,神经元是分层排列的,每个神经元只与前一层的神经元相连,如图 1.1(a)所示。最上一层为输出层,最下一层为输入层,还有中间层,中间层也称为隐层,隐层的层数可以是一层或多层。

(2) 输入输出有反馈的前馈网络。如图 1.1(b)所示,在输出层上存在一个反馈回路到输入层,而网络本身还是前馈型的。该种神经网络的输入层不仅接受外界的输入信号,也接受网络自身的输出反馈信号。输出反馈信号可以是原始输出信号,也可以是经过转化的输出信号;可以是本时刻的信号,也可以是经过一定延

迟的信号。此种网络经常用于系统控制、实时信号处理等需要根据系统当前状态进行调节的场合。

(3) 前馈内层互连网络。如图 1.1(c)所示,在同一层内存在互相连接,它们可以形成互相制约,而从外部看还是一个前向网络,很多自组织网络大都存在内层互连的结构。

(4) 反馈型全互连网络。图 1.1(d)所示的网络是一种单层全互连网络,每个神经元的输出都与其他神经元相连,如 Hopfield 神经网络和 Boltzmann 机都是属于这一类网络。

(5) 反馈型局部连接网络。图 1.1(e)所示的网络是一种单层网络,它的每个神经元的输出只与其周围的神经元相连形成反馈的网络,这类网络也可发展为多层的金字塔形的结构。

目前,最具代表性的前馈神经网络模型有:①反向传播(back propagation, BP)神经网络。它是一种多层前馈网络,采用最小均方差的学习方式,是使用最广泛的网络。这个网络的缺点是为有导师训练,训练时间较长,易于陷入局部极小等。②径向基函数(radial basis function, RBF)神经网络。它是一种非常有效的多层前馈网络,其神经元基函数具有仅在微小局部范围内才产生有效的非零响应的局部特性,因而可以在学习过程中获得高速化。这个网络的缺点是由于高斯函数的特性,该网络难以学习映射的高频部分。最具代表性的反馈网络模型是 Hopfield 神经网络,该网络是由相同的神经元构成的单层网络,并且是不具学习功能的自联想网络,它需要对称连接。这个网络可以完成制约优化和联想记忆(associative memory, AM)等功能。

1.2 神经网络的基本特征和基本功能

1.2.1 神经网络的基本特征

神经网络的基本特征可归结为结构特征和能力特征。

(1) 结构特征——并行处理、分布式存储与容错性。人工神经网络是由大量的简单处理单元相互连接构成的高度并行的非线性系统,具有大规模并行性处理特征。虽然每个处理单元的功能十分简单,但大量简单处理单元的并行活动使网络呈现出丰富的功能,并具有较快的速度。结构上的并行性使神经网络的信息存储必然采用分布式方式,即信息不是存储在网络的某个局部,而是分布在网络所有的连接权中。神经网络内在的并行性与分布性表现在其信息的存储与处理都是在空间上分布、时间上并行的,这两个特点必然使神经网络在两个方面表现出良好的容错性:一方面由于信息的分布式存储,当网络中部分神经元损坏时不会

对系统的整体性能造成影响,这一点就像人脑中每天都有神经细胞正常死亡而不会影响大脑的功能一样;另一方面当输入模糊、残缺或变形的信息时,神经网络能够通过联想恢复出完整的记忆,从而实现对不完整输入信息的正确识别,这一点就像人可以对不规则的手写字进行正确识别一样。

(2) 能力特征——自学习、自组织与自适应性。自适应性是指一个系统能够改变自身的性能以适应环境变化的能力,它是神经网络的一个重要特征。自适应性包含自学习与自组织两层含义。神经网络的自学习是指当外界环境发生变化时,经过一段时间的训练或感知,神经网络能够通过自动调整网络结构参数,对给定输入能产生期望的输出。训练是神经网络学习的途径,因此,经常将学习与训练两个词混用,神经系统能在外部刺激下按一定规则调整神经元之间的突触连接,逐渐构建起神经网络,这一构建过程称为网络的自组织。神经网络的自组织能力与自适应性相关,自适应性是通过自组织实现的。

1.2.2 神经网络的基本功能

(1) 联想记忆。由于神经网络具有分布存储信息和并行计算的性能,因此,它具有对外界刺激信息和输入模式进行联想记忆的能力,这种能力是通过神经元之间的协同结构及信息处理的集体行为实现的。神经网络是通过其突触权值和连接结构来表达信息的记忆,这种分布式存储使神经网络能存储较多的复杂模式和恢复记忆的信息。神经网络通过预先存储信息和学习机制进行自适应训练,可以从不完整的信息和噪声干扰中恢复原始的完整信息,这一能力使其在图像复原、图像和语音处理、模式识别、分类等方面具有巨大的潜在应用价值。

(2) 非线性映射。在客观世界中,许多系统的输入与输出之间存在复杂的非线性关系,对于这类系统,往往很难用传统的数理方法建立其数学模型。设计合理的神经网络通过对系统输入输出样本对进行自动学习,能够以任意精度逼近任意复杂的非线性映射。神经网络的这一优良性能使其可以作为多维非线性函数的通用数学模型,该模型的表达是非解析的,输入输出数据之间的映射规则由神经网络在学习阶段自动抽取并分布式存储在网络的所有连接中。

(3) 分类与识别。神经网络对外界输入样本具有很强的识别和分类能力。对输入样本的分类实际上是在样本空间找出符合分类要求的分割区域,每个区域内的样本属于一类。传统分类方法只适合解决同类相聚、异类分离的识别与分类问题,但客观世界中,许多事物在样本空间上的区域分割曲面是十分复杂的,相近的样本可能属于不同的类,而远离的样本可能同属一类。神经网络可以很好地解决对非线性曲面的逼近,因此,比传统的分类器具有更好的分类和识别能力。

(4) 优化计算。优化计算是指在已知的约束条件下寻找一组参数组合,使由该组合确定的目标函数达到最小值。某些类型的神经网络可以把待求解问题的

可变参数设计为网络的状态,将目标函数设计为网络的能量函数,神经网络经过动态演变过程达到稳定状态时对应的能量函数最小,从而其稳定状态就是问题的最优解。这种优化计算不需要对目标函数求导,其结果是网络自动给出的。

(5) 知识处理。神经网络获得知识的途径与人类相似,也是从对象的输入输出信息中抽取规律而获得关于对象的知识,并将知识分布在网络的连接中予以存储。神经网络的知识抽取能力使其能够在没有任何先验知识的情况下自动地从输入数据中提取特征,发现规律,并通过自组织过程将自身构建成适合于表达所发现的规律。另外,人的先验知识可以大大提高神经网络的知识处理能力。两者相结合会进一步提升神经网络的智能。

1.3 神经网络的基本性质、优点及应用

1.3.1 神经网络的基本性质

神经网络的基本性质主要包括收敛性、容错性、鲁棒性及推广性等。

神经网络的收敛性是指神经网络的训练算法在有限次迭代之后可收敛到正确的权值或权向量。神经网络良好的容错性保证网络将不完整的、污损的、畸变的输入样本恢复成完整的原型。容错性的研究归结于神经网络动力系统记忆模式吸引域的大小,吸引域越大,网络从部分信息恢复全部信息的能力越大,表明网络的容错性越大。神经网络的高度鲁棒性使网络中的神经元或突触遭到破坏时网络仍然具有学习和记忆能力,从而使网络表现出高度的自组织性。研究表明,如果记忆模式的吸引域越“规则”,那么网络抵抗干扰、噪声或自身损害的能力就越强,即鲁棒性越好。训练好的神经网络应能够对不属于训练样本集合的输入样本正确识别或分类,这种现象常称为神经网络具有良好的推广性。

人工神经网络的操作有两种过程:一是训练学习,二是正常操作或称回忆。训练时,把要教给网络的信息(外部输入)作为网络的输入和要求的输出,使网络按某种规则(训练算法)调节各处理单元间的连接权值,直到加上给定输入后网络就能产生给定输出为止。这时,各连接权已经调节好,网络的训练完成。所谓正常操作,就是对训练好的网络输入一个信号,它就可以正确回忆出相应输出。不论是训练网络还是操作网络,人工神经网络的状态总是变化的。所谓神经网络的状态,是指神经网络所有节点的输出信号值。状态变化可以指某个节点的状态变化,也可指所有节点的状态变化。神经网络的这种动态特性受两种性能的约束,即系统的整体稳定性和收敛性。所谓稳定的神经网络,定义为这样一种非线性动态系统:当在该系统上加入一初始输入时,系统的状态发生变化,但最后达到一固定点(收敛点或均衡点),这些固定点就是存储信息的点。虽然稳定的神经网络总

能保证所有输入被映射到固定点,但不能保证该固定点就是要求的固定点。不难理解,神经网络的稳定性是与反馈网络的回忆操作相联系的,这种反馈网络的稳定性可以用 Lyapunov 准则进行判定。收敛性是指在训练过程中,输出节点的实际输出值与要求的输出值之间的误差,最后能达到可接受的最小值。一般要求收敛过程迅速和精确,即输出能尽快趋于目标值。显然,收敛性是与有指导的训练操作相联系,收敛过程严格依赖于所采用的具体训练算法和训练参数。

1.3.2 神经网络的优点

- (1) 很强的鲁棒性和容错性。这是因为信息是分布存储于网络内的神经元中。
- (2) 并行处理方法。人工神经网络在结构上是并行的,而且网络的各个单元可以同时进行类似的处理过程,使计算快速。
- (3) 自学习、自组织、自适应性。神经元之间的连接多种多样,各神经元之间连接强度具有一定可塑性,使得神经网络可以处理不确定或不知道的系统。
- (4) 可以充分逼近任意复杂的非线性关系。
- (5) 具有很强的信息综合能力。能同时处理定量和定性的信息,能很好地协调多种输入信息关系,适用于处理复杂非线性和不确定对象。

1.3.3 神经网络的应用

神经网络以其独特的结构和处理信息的方法,在许多实际应用领域中取得了显著的成效。主要应用于自动控制、处理组合优化问题、模式识别、图像处理、传感器信号处理、机器人控制、信号处理、卫生保健、医疗、经济、化工、焊接、地理、数据挖掘、电力、交通、军事、矿业、农业和气象等领域。

1.4 神经网络的性能指标及研究内容

1.4.1 神经网络的性能指标

与人脑的作用机理类似,一个神经网络完成任务的过程包括学习(训练)过程和使用(回忆或联想)过程。对一个神经网络学习算法来说,衡量其性能优劣的指标有以下几个方面:

- (1) 泛化能力。一个训练好的神经网络到实际中使用是否有好的效果,这是神经网络最重要的性能指标。
- (2) 时间复杂性。训练一个固定结构的神经网络所需要的时间。
- (3) 空间复杂性。算法计算机实现时所占用的内存空间,一般与神经网络的结构复杂程度密切相关。

(4) 在线学习能力。如果神经网络的学习过程和使用过程是分别进行的,即先学习后使用,则称为离线学习;如果这两个过程是同时进行的,即边学习边使用,则称为在线学习能力。

(5) 其他指标。包括能否用硬件实现、算法的稳定性、神经网络模型的鲁棒性等。

1.4.2 神经网络的研究内容

当前,神经网络研究内容主要包括神经网络理论研究、神经网络实现技术研究和神经网络应用研究三个方面。

(1) 神经网络理论研究。神经网络理论研究侧重于寻找合适的神经网络模型和学习算法。其中,模型研究是指构造合适的单个神经元模型及确定神经元之间的连接方式,并探讨它所适用的场合;学习算法研究是指在神经网络模型的基础上找出一种调整神经网络结构和权值的算法,并满足学习样本的要求,同时具有较快的学习速度。神经网络理论研究的另一个重要内容是从理论上分析常用的神经网络设计方法对泛化能力的影响。

(2) 神经网络实现技术研究。神经网络实现技术研究主要是探讨利用电子、光学、光电、生物等技术实现神经计算机的途径,包括利用传统计算机技术实现模拟神经计算机及新型神经计算机体系结构的研究等。

(3) 神经网络应用研究。神经网络应用研究是探讨如何利用神经网络解决实际工程问题。人们可以在几乎所有的领域中发现神经网络应用的影子。当前,神经网络的主要应用领域有模式识别、故障检测、智能机器人、非线性系统辨识和控制、市场分析、决策优化、物资调用、智能接口、知识处理和认知科学等。

1.5 神经网络的发展简史、存在问题及发展趋势

1.5.1 神经网络的发展简史

神经网络是一门活跃的边缘性交叉学科,研究它的发展过程和前沿问题具有重要意义。目前,神经网络的理论和应用研究得到了极大的发展,而且已经渗透到几乎所有的工程应用领域。但是,人工神经网络的发展过程并不是一帆风顺的,大致经历了以下几个阶段:

(1) 初始时期。1943年,McCulloch 和 Pitts^[6]提出了 MP 模型,从而给出了神经元的最基本模型及相应的工作方式。1949年,神经生物学家 Hebb^[7]发现,脑细胞之间的通路在参与某种活动时将被加强,这个重要规则给出了生理学与心理学间的联系,被称为 Hebb 学习规则,该规则至今还被许多神经网络学习算法所使

用。1958年, Rosenblatt^[8]提出了感知器模型,这是一个由线性阈值神经元组成的前馈神经网络,可用于分类。1960年, Widrow 和 Hoff^[9]提出了自适应线性单元,这是一种连续取值的神经网络,可用于自适应系统。

(2) 低潮时期。1969年,人工智能的创始人 Minsky 和 Papert^[10]出版了 *Perceptrons*, 在该书中,他们指出单层感知器只能作线性划分,多层感知器不能给出一种学习算法,因此无实用价值。由于 Minsky 和 Papert 在人工智能领域的地位,该书在人工神经网络研究人员间产生了极大的反响,从而使神经网络研究受到了严重影响,自此陷入低潮。但是,即便在神经网络研究的低潮时期,也仍有一些人在兢兢业业地研究神经网络,并取得了一些重要成果。其中,最著名的是1982年加利福尼亚理工大学教授 Hopfield^[11]提出的 Hopfield 神经网络。在这个用运算放大器搭成的反馈神经网络中, Hopfield 借用 Lyapunov 能量函数的原理,给出了网络的稳定性判据,并为著名的组合优化问题——旅行商问题(TSP)提供了一个新的解决方案。Hopfield 神经网络可用于联想存储、优化计算等领域。

(3) 高潮时期。1986年, Rumelhart 等^[12]给出了多层感知器的权值训练 BP 算法,从而解决了 Minsky 认为不能解决的多层感知器的学习问题,自此引导了神经网络的复兴,神经网络的研究也进入了一个崭新的发展阶段。

1.5.2 神经网络的一些问题

随着对神经网络研究的广泛关注,其中的一些问题逐渐暴露出来,并已成为该学科进一步发展的障碍。目前,亟待从以下几个方面予以改进^[13]:

(1) 加快神经网络的学习速度。目前,绝大多数神经网络算法都需要进行耗时的迭代训练,其计算开销相当大,训练速度太慢,难以满足实时性要求较高的在线学习任务的需要。此外,太慢的速度还使得神经计算技术很难用于数据挖掘等领域,因为等到网络训练完成时,数据库的内容可能已经发生了更新变动,网络学习到的知识将无法反映出当前事物的特点。

(2) 增强神经网络的可理解性。神经网络模型的一大特点是其分布式知识表示,即网络中单一的神经元或连接并没有明确的意义,这就决定了神经网络是一种典型的“黑箱”模型,其学习到的知识隐藏在大量的连接权值中,用户无法知道某一个具体的网络能做什么,也无法知道它是怎么做。一般来说,“可解释性”是可靠系统的必备特性,由于通常的神经网络模型都是“不可解释”的,这在一定程度上影响了用户对通过神经计算技术构建智能系统的信心。虽然 Baum 和 Haussler^[14]指出,“如果一个神经网络可以为大量的训练例产生正确的解答,那么,可以相信它们也能为类似于训练例的未知例产生正确解答”,但这并没有抵消用户对可理解性的偏好。此外,训练好的神经网络学习到的知识不能以容易理解的方式提交给决策者,这也是神经计算技术难以用于数据挖掘领域的主要原因