

江苏省自然科学基金项目(No.BK20130209)

国家自然科学基金项目(No.61379101)

中国博士后科学基金面上项目(No.2014M560460)

粒度神经网络原理及应用

■ 许新征 ○ 著

LIDU SHENJING WANGLUO

Yuanli Ji Yingyong

中国矿业大学出版社

University of Mining and Technology Press

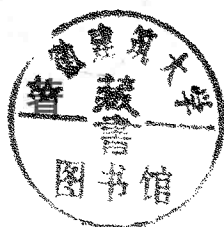
江苏省自然科学基金项目(No. BK20130209)

国家自然科学基金项目(No. 61379101)

中国博士后科学基金面上项目(No. 2014M560460)

粒度神经网络原理及应用

许新征



中国矿业大学出版社

内 容 提 要

本书主要从两个方面研究了基于粗糙集的粒度神经网络模型。一种方式是使用粗糙集作为前置系统,利用属性约简算法对数据集进行粒度约简,以简化神经网络的结构,提高神经网络的训练速度和预测精度。另一种方式是利用粗糙集及其扩展模型来提取决策规则,根据提取的规则来定义粒度神经元及其连接权值,实现粗糙集和神经网络的无缝融合。

本书力求科学化、实用化和前沿化,使读者能够在有限的时间内了解基于粗糙集的粒度神经网络原理及其应用。

本书可作为高等院校计算机专业和其他信息类专业研究生的参考书,也可以供相关领域的科技工作者参考。

图书在版编目(CIP)数据

粒度神经网络原理及应用/许新征著. —徐州:中国矿业大学出版社,2015.11

ISBN 978 - 7 - 5646 - 2831 - 4

I. ①粒… II. ①许… III. ①算法语言—应用—人工神经网络 IV. ①TP183

中国版本图书馆 CIP 数据核字(2015)第 220420 号

书 名 粒度神经网络原理及应用

著 者 许新征

责任编辑 王美柱

出版发行 中国矿业大学出版社有限责任公司

(江苏省徐州市解放南路 邮编 221008)

营销热线 (0516)83885307 83884995

出版服务 (0516)83885767 83884920

网 址 <http://www.cumtp.com> E-mail: cumtpvip@cumtp.com

印 刷 徐州中矿大印发科技有限公司

开 本 880×1230 1/32 印张 4.375 字数 122 千字

版次印次 2015 年 11 月第 1 版 2015 年 11 月第 1 次印刷

定 价 25.00 元

(图书出现印装质量问题,本社负责调换)

前 言

作为粒计算的三个主要模型之一,粗糙集可以直接对数据进行分析 and 推理,从中发现隐含的知识,揭示潜在的规律。而神经网络是一种模仿生物神经网络行为特征,进行分布式并行信息处理的智能计算模型。鉴于粗糙集和神经网络的诸多互补之处,将两者集成的粒度神经网络模型已成为该领域的研究热点问题之一。本书的主要工作是提出了几种基于粗糙集的粒度神经网络模型及其学习算法,并通过实验验证了网络结构及其学习算法的有效性。

全书共 8 章,主要包括:

第 1 章:绪论,介绍了粒度神经网络和极速学习机的研究现状,以及基于粗糙集的粒度神经网络集成中存在的若干问题。

第 2 章:粗糙集理论基础,简要介绍了粗糙集理论中的基本概念,粗糙集中的知识表示方法,知识约简原理,以及决策表的几种常用属性约简算法。

第 3 章:基于粗糙集和 QPSO 算法的粒度 BP 神经网络,主要研究了粒度 BP 神经网络的构建方法,提出了适用于粒度 BP 神经网络的学习算法。

第 4 章:基于粗糙集和 AP 聚类的粒度 RBF 神经网络,提出了基于 AP 聚类的粒度 RBF 神经网络构建方法,实现了粒度 RBF 神经网络结构的自动确定。

第 5 章:具有自适应极速学习能力的单隐层粒度神经网络,提出了一种改进的自适应极速学习算法,建立了具有自适应极速学习能力的单隐层粒度神经网络。

第6章:基于粗规则的粒度神经网络,提出了一种基于决策规则的粗规则粒度神经网络模型,并利用极速学习机算法进一步调整其参数。

第7章:基于变精度粗糙集的粒度双神经元网络,提出了一种粒度双神经元网络模型的构建方法,以及基于AP聚类粒化的优化方法。

第8章:总结与展望。

本书包含了笔者多年的研究成果,也借鉴和吸收了国内外专家学者的研究成果和公开出版的论文资料,在此谨向他们致以崇高的敬意。本书的顺利撰写得到了中国矿业大学计算机科学与技术学院、中国矿业大学出版社等各级领导和同事们的支持和帮助,在此表示感谢,特别是丁世飞教授给予的指导和帮助。

本书得到了江苏省自然科学基金项目(No. BK20130209)、国家自然科学基金项目(No. 61379101)、中国博士后科学基金面上项目(No. 2014M560460)和中央高校基本科研业务费专项资金项目(Nos. 2013QNA24 and 2015XKMS088)的支持。此外,本书还得到了中国矿业大学“第九批校级优秀青年骨干教师”和江苏省矿山机电装备重点实验室专项资金的资助。

由于粒计算和神经网络的不断发展,新的理论方法和技术不断涌现,再加上笔者的学识水平所限,书中难免有错误和不足之处,敬请读者批评指正。

著 者

2015年8月

目 录

1 绪论	1
1.1 研究背景	1
1.2 国内外研究现状	2
1.3 存在的问题及本书研究的重点	13
1.4 本书的主要研究内容和组织结构	15
2 粗糙集理论基础	17
2.1 概述	17
2.2 粗糙集的基本概念	19
2.3 粗糙集中的知识表示	23
2.4 知识约简原理	25
2.5 决策表属性约简算法	28
2.6 小结	30
3 基于粗糙集和 QPSO 算法的粒度 BP 神经网络	32
3.1 BP 神经网络概述	33
3.2 基于粗糙集的粒度约简	38
3.3 粒度 BP 网络的结构与算法	40
3.4 实验与分析	46
3.5 小结	52

4 基于粗糙集和 AP 聚类算法的粒度 RBF 神经网络	53
4.1 RBF 神经网络概述	54
4.2 粒度 RBF 神经网络的结构与学习算法	56
4.3 实验与分析	60
4.4 小结	66
5 具有自适应极速学习能力的单隐层粒度神经网络	67
5.1 ELM 概述	68
5.2 单隐层粒度神经网络的结构	70
5.3 单隐层粒度神经网络的自适应极速学习算法	72
5.4 实验与分析	72
5.5 小结	79
6 基于粗规则的粒度神经网络	80
6.1 决策规则的提取	81
6.2 粗规则粒度神经网络的结构和学习算法	88
6.3 实验与分析	93
6.4 小结	97
7 基于变精度粗糙集的粒度双神经元网络	98
7.1 变精度粗糙集模型概述	99
7.2 粒度双神经元结构	101
7.3 粒度双神经网络模型	104
7.4 粒度双神经网络的进一步优化	106
7.5 实验与分析	109
7.6 小结	113

目 录

8 总结与展望	114
8.1 总结	114
8.2 展望	116
参考文献	118

1 绪 论

1.1 研究背景

作为典型的分类模型,人工神经网络(Artificial Neural Networks, ANN)因具有很强的非线性映射能力及自学习、并行性、抗噪能力和容错能力一直以来都是机器学习与数据挖掘领域的一个重要分支。但在处理数据时,神经网络存在着学习时间较长、表达的知识隐藏且难理解等固有缺陷。近年来提出的粒计算(Granular Computing, GrC)理论与方法可对复杂问题进行抽象和简化,以较低的计算代价得到问题的满意近似解;也可以通过改变粒度的大小来隐藏或者揭示问题的一些细节,以有利于问题的求解。粒计算改变了传统的计算观念,使信息的处理更科学、合理、经济和易操作,在智能系统的设计和实现中将有着重要的作用。利用粒计算思想优化神经网络而形成的粒度神经网络(Granular Neural Networks, GNN)拥有更强的泛化能力与抗噪声性。作为典型的粒计算模型,粗糙集(Rough Sets, RS)与神经网络相结合是当今智能信息处理研究领域的热点之一。粗糙集理论的基本思想是在保持分类能力不变这一粒度划分目标的前提下,对数据集进行粒度约简,并根据约简的结果提取决策或者分类的规则。但粗糙集的主要缺点是在进行数据分析时对噪声比较敏感,因而根据不含噪声的训练样本进行推理和学习得到的分析结果,不适合应用于含有噪声的环境中,这也导致了粗糙集的泛化和推广能力较

弱。而神经网络具有并行处理、高容错性、泛化能力强和较好的抑制噪声能力等特点,特别适合于分类预测、故障诊断、复杂对象建模和智能控制等领域。但神经网络处理数据时,一般不能对输入数据进行降维或约简,也不知道哪些知识是冗余的哪些是有用的。因此,当输入样本的规模较大时,容易导致神经网络结构比较复杂,学习速度缓慢,训练时间较长等问题。而粗糙集理论却可以不利用任何先验知识就可以发现数据间的关系,不仅可以去掉冗余输入信息,而且可以简化输入信息的维数。由此可见,粗糙集和神经网络既有各自的特点,又有很多互补之处,探索两者的有机结合、融入两者优点而建立的粒度神经网络将成为智能信息处理领域分析和处理问题的有效研究方法,具有重要的研究意义。

1.2 国内外研究现状

ANN 在处理小规模数据集时,表现出较好的并行处理能力、容错性、自学习和自适应等特点,具有很强的非线性映射能力。但是,当 ANN 处理大规模数据集或 Internet 环境下的高维复杂数据时,具有极大的局限性,其网络性能受到极大的影响,甚至导致网络不能工作。而粒计算理论是用来处理高维复杂数据的有效方法,其基本思想是“分而治之”,即在不同粒度下对高维复杂数据进行粒化。粒化后,再使用传统的数据挖掘方法,如神经网络,在各个粒层上对数据进行处理。因此,结合粒计算和神经网络两者优点而建立的粒度神经网络模型,成为智能信息处理领域的一个研究热点,引起了越来越多研究人员的关注。

1.2.1 粒度神经网络研究现状

20 世纪 60 年代,美国著名控制论专家 L. A. Zadeh 提出了模糊集合论,在此基础上,于 1979 年首次提出了模糊信息粒化问

题^[1],他认为粒化(Granulation,全体分解为部分)、组织(Organization,部分集成为整体)和因果(Causation,因果的关联)是人类认知的3个主要特征^[2-4]。J. R. Hobbs于1985年提出了粒度(Granularity)的概念^[5],R. R. Yager和D. Filev进一步指出“人们已形成了一个关于世界的粒度观点……在此观点下,人类的观察、度量、概念化和推理都是在粒度意义下进行的”^[6]。1996年,T. Y. Lin在UC-Berkeley大学访问时,向L. A. Zadeh提出了粒计算(Granular Computing, GrC)的概念^[7]。目前,有关粒度计算的理论与方法主要有三个:L. A. Zadeh教授^[1]提出的“词计算理论”(Theory of Words Computing),Z. Pawlak教授^[8]的“粗糙集理论”(Theory of Rough Set),我国学者张钹院士和张铃教授^[9]提出的“商空间理论”(Theory of Quotient Space)。

ANN是一种模拟人脑信息处理机制的智能系统,不但可以处理数值数据,而且还可以处理知识的思维、学习和记忆能力^[10]。ANN具有自组织、自学习、自适应、容错性、大规模并行处理以及联想记忆等特点,已经发展成为解决非线性问题的强有力工具,并被广泛应用于许多科学研究领域。但是,传统的神经网络也存在着很多问题,如易陷入局部极小,训练时间较长,过拟合等问题。为了克服上述不足,许多优化算法被引入到神经网络权值的学习和结构的自动设计中,如基于粒子群算法优化的神经网络^[11],基于进化算法优化的进化神经网络^[12-15]等。这些优化算法虽然被验证是可行且有效的,但是在处理和分析数据时存在训练时间长、学习效率低下及泛化能力差等问题,而这些问题目前仍然没有较好地解决。

随着粒计算理论不断完善和发展,粒计算理论和神经网络结合而建立的粒度神经网络模型,受到了越来越多研究人员的关注与研究。下面概述粒度神经网络和基于粒计算模型的神经网络优化方法的研究现状。

1.2.1.1 粒计算模型与神经网络的结合

作为智能信息处理领域的两个独立分支,模糊理论和神经网络理论在各自的应用领域均表现出很强的容错能力。模糊理论是基于人的模糊逻辑思维建立的理论,本身就具有容错性。而神经网络的容错能力则来源于它对人脑的形象思维的模拟。因此,模糊神经网络融合了模糊理论的模糊表达能力与神经网络的并行自学习能力等优点^[16]。与传统神经网络的连接权值没有明确的物理含义不同,模糊神经网络利用神经网络的学习过程来实现模糊推理,其权值可以用来表示模糊逻辑中的推理参数。目前,模糊神经网络模型主要有以下几种类型:① 简单联结。该类型中,模糊控制和神经元网络串联组合,各自独立工作。② 神经网络为主、模糊理论为辅。该类型中,基于模糊理论先将输入空间分割成若干模糊推论组合,经过模糊逻辑判断后,作为神经元网络模型的输入。③ 模糊理论为主、神经网络为辅。该类型中,以模糊控制方法作为神经网络的学习“样本”,通过神经网络的离线学习过程来实现模糊控制的决策过程。④ 神经网络与模糊理论的完全融合。该类型中,两个系统密切结合,不能分离。当用于信息处理时,两种模型直接共同作用于控制对象,同时处理输入信息,实现两者的完全融合。

粗糙集理论可以直接对数据进行分析 and 推理,从中发现隐含的知识,揭示潜在的规律。而作为数据挖掘的另一种经典方法,神经网络是一种模仿生物神经网络行为特征,进行分布式并行信息处理的智能计算模型。鉴于粗糙集和神经网络在信息处理方式、知识获取方式、抑制噪声能力、泛化测试能力等方面有很多互补之处,将两种技术的有效结合是当前的一个研究热点^[17],已引起许多学者的广泛关注。目前,常用的有如下几种结合方式:① “松耦合”方式:粗糙集作为神经网络的预处理,主要是用来消除数据集中的冗余属性,同时用来简化神经网络结构,提高神经网络学习效

率;②“紧耦合”方式:通过粗糙集分析决策表得到初始决策规则,然后用于决策规则构建粗糙神经网络;③粗神经网络:在传统神经网络结构中的输入层和隐层之间加一个粗神经元(全互连接),以抑制输入层数据的波动。

目前,商空间理论用于神经网络优化的研究还比较少,在此不作介绍。

1.2.1.2 粒度神经网络研究现状

2000年,Y. Q. Zhang等^[18]面向数据库中的数字和语言数据问题提出了一个新颖的粒度神经网络(GNN)用来处理数据库中的粒度知识。该网络能学习输入和输出间的粒度联系,并预测新的联系。GNN能处理粒状数据(如数字和语言数据),提取粒状的IF-THEN规则,融合粒状数据组,压缩粒状数据库,并且能预测新的数据。粒度神经网络克服了人工神经网络的缺陷,可以较好地处理复杂数据,但是处理海量数据时,算法的速度和效率还需要进一步提高。2002年,M. Syeda和Y. Q. Zhang等^[19]提出了并行粒度神经网络,用来处理信用卡欺诈行为的监测,加快了数据挖掘和知识发现的过程。2005年,A. Vasilakos和D. Stathakis^[20]将粒度神经网络应用于陆地分类,通过对卫星图片的处理,获得了较好的结果。2007年,Y. Q. Zhang等^[21]提出了遗传粒度神经网络。Y. Q. Zhang等^[22]使用粒度神经网络开发了用于基于网页的股票预测代理。粒度神经网络使用各个股票数据进行训练,用于发现模糊规则。2008年,M. Milan和M. Dušan^[23]将粒度神经网络应用到工资时序数据的预测,其建立了一种称为模糊逻辑RBF神经网络,将标准的高斯激活函数替换为基于标准云概念(Normal Cloud Concept)的函数,这使得数据固有的不确定性和随机性被同时包含到预测的过程中。用该方法与其他多种方法进行比较,该方法可建立一个更好的模型。2008年,Y. Q. Zhang等^[24]提出一种基于阴阳方法的使用粒度集合的新的间隔推理(Interval

Reasoning)方法。为了使得间隔粒度推理有效地优化基于训练数据的间隔激活函数,Y. Q. Zhang 等设计了一种基于新的快速的进化间隔学习的粒度神经网络。将其用于非线性函数逼近和生物信息学,仿真结果表明该方法能够更有效地提取出粒度规则。2009年7月,D. F. Leite 等人提出了进化粒度神经网络(eGNN)模型,该模型是被提出用来解决分类问题持续不断的变化环境的,eGNN 构造上使用了快速增量学习算法,该模型需要较少的内存执行分类任务。模型的主要特点是不断学习,自我组织,适应未知环境^[25]。2010年6月,D. F. Leite 等又提出了一个数据流分类的自适应模糊神经网络的框架使用部分监督的学习算法。该模型包含一个不断发展的粒度神经网络,并且能够处理没有固定的数据流,使用了增量算法。粒度神经网络模型包含模糊超盒和基于神经元分类的空规范数据,学习算法进行结构和参数的适应环境的变化时,输入数据反映。它需要事先没有有关数据和类的统计知识。计算实验表明,针对不同类型的概念漂移颗粒的模糊神经网络是强大的,能够有效地处理未标记的例子^[26]。2011年9月,A. Ganivada等提出了一个基于多层感知器使用反向传播算法的模糊分类模式的模糊粗糙粒度神经网络模型(FRGNN),该模型的主要思路是依靠网络输入向量,并由模糊粗糙集理论确定初始连接权重和目标向量,输入向量是按照模糊粒度值,目标向量按照模糊类隶属度和零点。对每一个决策表的数据惊醒粒化,通过决策表的语法自动确定适当数量的隐藏节点,而所有的决策表的依赖因素是初始权。每个属性和决定类的所有属性的依赖因素的平均程度,被认为是输入层和隐层节点、隐藏层和输出层之间的初始连接权值^[27]。2011年8月,S. F. Ding^[28]等针对特征维数高的大样本,提出基于因子分析和聚类的BP神经网络算法。该算法通过FA对原始数据进行特征降维,以简化网络结构;通过CA对降维后的样本划分成不同子类,进行网络训练,以提高网络的适应

性,应用时将新样本先归类,再利用该类所对应的网络进行预测。2011年7月,X. Z. Xu^[29]等建立了一种新型的基于普通最小二乘(Ordinary Least Squares, OLS)的神经网络集成优化分类算法。将数据先进行特征降维,以方便网络的设计,分别用建立基于主成分分析(Principal Component Analysis, PCA)的 Elman 神经网络算法作为子分类器 I,建立基于因子分析(Factor Analysis, FA)的 RBF 神经网络算法作为子分类器 II,将两个子分类器进行集成学习,用 OLS 原理确定各子分类器的最优权值,则得到集成优化分类算法。2011年7月,K. C. Kwak 构建了级联粒度神经网络的模型,该模型首先构造一个粒度神经网络,作为基础模型。接下来,构建另一个 GNN 用来补偿基础模型产生的误差。这个模型较之前的模糊建模中遇到的以规则为基础的系统相比较,既提高了效率也提高了运算精度^[30]。

1.2.1.3 粗糙神经网络研究现状

基于粗糙集和神经网络集成的粗集神经网络作为智能集成系统的重要分支,在基础理论研究和应用研究等方面未成熟,还处于实验和探索阶段。本书总结了现阶段关于粗集神经网络研究的几种方式,下面介绍几类常见的粗集神经网络模型。

(1) 基于粗糙集数据预处理的粗集神经网络

粗糙集理论通过属性约简算法和值约简算法实现对原始决策信息系统的预处理,删除原始系统中的冗余信息。基于粗糙集的数据预处理分为3部分:形成决策表,连续属性离散化和属性约简^[31]。最终形成最简决策表为后面做数据处理提供依据。粗糙集的数据预处理包括属性约简、缺省数据的处理和错误数据的处理。缺省数据处理方法大概分为两种:用粗糙集补齐数据和直接利用不全的决策进行预处理。由于粗糙集进行属性约简是个 NP 问题,且粗糙集面对的是未知信息,对于错误数据的处理显得乏力。约简可以包括属性约简和值约简,分别对应用来对条件属性

简化和对决策属性进行简化^[32]。常用的属性约简技术主要有以下几种^[33]：差别矩阵、正区域、信息熵、属性重要度和互信息等。由于传统的粗糙集理论存在较多的不足，为了得到比较稳定、泛化能力强的约简，可采用动态技术和优化算法。通过将扰动方法和扰动特征相结合，提出了 Attribute Bagging 算法^[34]，取得了较好的应用效果，但是算法的时间复杂度较高。在此基础上，凌锦江等^[35]提出了一种 RegBag(Relief Bagging)改进算法，该算法引入采样来进行特征提取，减小了算法的时间复杂度。张东波等^[36]将扰动方法引入到样本空间和属性空间，提出基于特征选择的粗糙神经网络集成方法，并取得了很好的测试效果。

(2) 基于粗逻辑的粗集神经网络

粗逻辑结合粗糙集和逻辑学的特点来分析和处理问题，并以此提出了粗逻辑神经网络。粗逻辑神经网络一般分为 4 层：输入层(数据的输入)、模糊化层(对输入数据进行离散化，通过模糊化神经元实现离散区间的模糊化)、规则层(每个节点代表一条规则，规则的获取是粗逻辑神经网络的重点)、输出层(输出结果)。粗逻辑神经网络用规则层代替了原始神经网络的隐含层，从而使神经元具有可理解性，使之不再为黑箱。张东波等人^[37]提出了逻辑决策算法的实现过程：属性域知识编码；知识获取；决策算法的简化；决策算法的实现。这给规则提取提供了很好的方向，简化了粗逻辑神经网络的实现。基于近似域可划分的可变离散精度逻辑神经网络^[38]的提出，解决了由可能性区域信息粒度过大造成的误分类，对可能性区域离散区间进一步细化。基于变精度粗糙集的神经网络^[39]，通过对近似约简条件进行弱化推广，使其具有更强的近似决策和泛化能力。

(3) 基于粗神经元的粗集神经网络

粗神经元是在传统神经元基础上，通过和粗糙集理论相结合而构成的。粗神经元主要是可以用来处理输入为某个范围的情

况,此时可以利用粗糙集理论中的上下近似来解决,以上近似代表输入参数范围的最大值,下近似代表输入参数范围的最小值。如果用经典的神经元,就不容易表示该类问题,产生的误差会比较大;而基于粗神经元构造的粗集神经网络,利用其粗神经元,可以很好地解决这类问题。文献[40-41]表明用粗神经元解决此类问题可以提高网络性能,降低误差。王玮等^[42]提出了粗糙神经网络模型,并与经典神经网络做了比较。它们间主要区别为粗神经网络的输入层和隐含层间,或隐含层和输出层间存在着全连接,或者它们都为全连接,其他则为经典神经元间的连接。现实中往往涉及一些模糊概念和模糊知识,如知识库的知识是清晰的而被近似的概念是模糊,或者两者都是模糊的,这时候一般的粗神经元是无法解决的。知识的模糊性和粗糙性经常是并存的,模糊粗糙神经网络即模糊—粗隶属度函数神经网络可以很好地解决上述问题。模糊粗糙神经网络^[43]分为6层:输入层、模糊化层(用模糊隶属函数对精确值进行模糊化)、规则前件层(用来匹配模糊规则的前件)、综合推理层(完成模糊规则后件的功能)、明层(得到模糊输出结果)、反模糊输出层(反模糊输出结果,得到明确的输出)。把该理论应用到信息处理,实验取得了很好的结果。规则层是模糊粗糙神经网络的重要部分。利用粗糙集提取规则首先要划分论域,确定每个分类的隶属函数,然后形成决策表,再利用粗糙集的核及约简,简化决策表,最后计算规则的置信度。

(4) 粗粒度神经网络

信息是有粒度的^[44],因此可以结合粒计算的多层次划分思想来解决复杂问题,将复杂问题在不同粗细粒度下转化成易求解问题。粗粒度神经网络就是基于粗信息粒度和神经网络结合而形成的^[45-46]。它主要由分布式智能体系统信息粒度的划分和每一个粒层上的近似空间组成。通过对智能体参数空间的参数进行调整来训练网络,通过对近似空间局部参数的调整进行学习,使其获得