Rough Set Theory and Method for Fault Diagnosis

面向故障诊断应用的粗糙集理论及方法

刘金福　白明亮　著

科学出版社

面向故障诊断应用的粗糙集理论及方法

Rough Set Theory and Method for Fault Diagnosis

刘金福　白明亮　著

科学出版社

北京

内 容 简 介

粗糙集方法具有强大的不一致信息处理能力，在故障征兆约简、诊断知识获取和知识度构建等方面表现出巨大优势。然而，关于粗糙集方法的泛化性能研究不足制约该方法在故障诊断中的实际应用。本书分别针对一般故障诊断问题以及多类故障诊断、类不平衡故障诊断和代价敏感故障诊断几种特定故障诊断问题，对粗糙集方法的泛化性能展开深入系统的论述，给出了基于结构风险最小化的粗糙集泛化性能提升方法、基于两分类器设计的多类故障类间干扰抑制方法、基于加权粗糙集的类不平衡故障诊断方法以及代价敏感粗糙集故障诊断方法，为粗糙集理论和方法在故障诊断应用中泛化性能的提高提供了支撑。

本书可供故障诊断和机器学习领域的研究人员及高等院校师生参考。

图书在版编目(CIP)数据

面向故障诊断应用的粗糙集理论及方法 = Rough Set Theory and Method for Fault Diagnosis/刘金福，白明亮著.—北京：科学出版社，2019.6

ISBN 978-7-03-061546-6

Ⅰ. ①面… Ⅱ. ①刘… ②白… Ⅲ. ①集论-研究 Ⅳ. ①0144

中国版本图书馆CIP数据核字(2019)第111959号

责任编辑：范运年 梁晶晶 / 责任校对：王瑞
责任印制：师艳茹 / 封面设计：铭轩堂

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

天津新科印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2019 年 6 月第 一 版 开本：720 × 1000 1/16
2019 年 6 月第一次印刷 印张：10
字数：200 000

定价：98.00 元

(如有印装质量问题，我社负责调换)

前　言

粗糙集方法具有强大的不一致信息处理能力，并且得到的规则型知识容易理解，所以，该方法已经广泛应用于工业过程及设备的故障诊断中，在故障征兆约简、诊断知识获取和知识库构建等诸多方面表现出巨大的优势。然而，到目前为止，粗糙集方法在实际故障诊断中的泛化性能却很少有研究，为此，本书分别针对一般故障诊断问题以及多类故障诊断、类不平衡故障诊断和代价敏感故障诊断几种特定故障诊断问题，对粗糙集方法的泛化性能进行深入系统的研究。

通过将机器学习领域中广泛应用的控制机器学习方法泛化性能的基本理论——结构风险最小化原则，引入粗糙集方法中，本书提出基于结构风险最小化的粗糙集方法，该方法不仅保证了粗糙集方法在现有故障实例集上具有较低的经验风险，而且有效地控制了粗糙集方法的复杂度，使获得的故障诊断知识具有更强的统计特性，系统的实验表明，这一方法能够明显提高粗糙集方法对新故障实例的泛化性能。进一步的实验发现，最小规则集中规则的数目能够有效地控制粗糙集方法的复杂度，基于这一复杂度指标设计的基于遗传多目标优化的结构风险最小化算法和启发式的结构风险最小化算法能够获得较好的性能。

当利用粗糙集方法对多类故障诊断问题进行处理时，某些类故障的关键征兆可能被认为是冗余征兆而被删除，以至于这些类故障的诊断知识由其他类故障的关键征兆表达，这势必影响粗糙集方法对这些类新故障实例的泛化性能。基于将多类问题转化为两类问题，本书提出基于两类分类器设计的多类故障诊断类间干扰抑制方法，使得构建的分类器最大限度地保留了各类故障的关键征兆。系统的实验表明，该方法能够有效地降低多类故障诊断的类间干扰，提高粗糙集方法在多类故障诊断问题中的泛化性能。进一步的对比实验发现，基于一对一的两类分类器构建策略和基于决策结果投票的分类器协同决策策略能够获得最好的性能，可以作为多类故障诊断类间干扰抑制方法的最佳配置。

在实际故障诊断问题中，少数类故障一般难以得到应有的重视。常规粗糙集方法倾向于选取多数类故障的关键征兆，并且提取的少数类故障诊断规则通常具有较低的支持度，这势必导致常规粗糙集方法对少数类新故障实例具有较差的泛化性能。通过对数据样本加权，本书提出基于加权粗糙集的类不平衡故障诊断问题处理方法，使得少数类故障实例中蕴含的诊断知识得到加强，系统的实验表明该方法能明显提高粗糙集方法对少数类故障的泛化性能，其获得的受试者工作特征曲线下的面积(area under the cure，AUC)和少数类分类精度明显优于重采样和

过滤方法，且与基于决策树和支持向量机(support vector machine，SVM)的类不平衡问题处理方法的性能相当，因此，证明提出的基于加权粗糙集的类不平衡故障诊断问题的处理方法是有效的。

常规粗糙集方法是基于诊断错误率最小化的，不考虑故障之间的误诊断代价差异，所以，通常不能保证对高代价故障具有较高的泛化性能，从而难以最小化故障诊断的代价。通过引入机器学习领域中广泛采用的代价敏感问题处理技术，本书提出基于加权粗糙集和最小期望代价分类准则的代价敏感故障诊断问题处理方法，同时考虑数据集类分布特性对代价敏感问题处理的影响。基于新提出的不依赖于测试集特性的性能评价指标，开展系统的实验，发现提出的代价敏感问题处理方法能明显地提高粗糙集方法对高代价故障的泛化性能，降低故障诊断的代价，并且采用逆类概率加权的加权粗糙集和最小期望代价分类准则相结合的代价敏感问题处理方法能够获得最好的性能，该方法可以作为代价敏感问题处理的最佳配置。

本书共 7 章，对粗糙集方法在各种实际故障诊断问题中的泛化性能进行深入系统的研究，通过引入相应的处理技术，明显提高粗糙集方法的泛化性能。

作　者

2018 年 12 月于哈尔滨

目　录

第1章 绪 论

1.1 概 述

1.1.1 开展故障诊断的重要性

随着现代化大生产的发展和科学技术的进步，现代工程技术系统与设备的结构越来越复杂，规模越来越庞大，自动化和智能化的程度也越来越高。但是，这类系统和设备一旦发生事故，将会造成巨大的人员及财产损失[1-4]。例如，1979年3月美国三里岛核电站重大泄漏事故造成几十亿美元的经济损失；1984年12月印度博帕尔农药厂甲基异氰酸酯泄漏事故造成2500多人死亡及近20万人中毒受害；1986年1月美国挑战者号航天飞机因密封系统故障造成失事悲剧；1986年4月苏联切尔诺贝利核电站泄漏事故造成2000多人死亡、达30亿美元经济损失和严重的污染公害。这种情况近年来在我国也同样存在。例如，1979年吉林某液化气站球罐破裂事故造成32人死亡，50多人伤残，直接经济损失600多万元，成为当年世界四大事故之一；1985年大同电厂、1988年秦岭电厂以及1999年阜新电厂的200MW汽轮发电机组的严重断轴毁机事故，造成上亿元的直接经济损失和重大的社会影响；另外，据我国对年产30万t合成氨和48万t尿素化肥厂五大透平压缩机组的初步调查结果，仅1977年和1978年的不完全统计，机械故障就达100多次，经济损失数亿，相当于另办一个新厂的年产量收益。因此，切实保证现代工程技术系统和大型复杂设备的可靠性与安全性是一个十分迫切的问题，具有重大的意义。

设备故障诊断技术是20世纪70年代以来，随着计算机和电子技术的飞跃发展，促进工业生产现代化和机器设备大型化、连续化、高速化、自动化而迅速发展起来的一门新技术，它是现代化设备维修技术的重要组成部分，并且正在成为设备维修管理工作现代化的一个重要标志。设备故障诊断技术对确保机械设备安全、降低突发故障、提高设备运行效率以及节约维修费用均有十分重要的作用。据相关资料报道[1,2]：日本采用状态监测诊断技术后，设备的事故率减少了75%，维修费用减少了25%～50%；英国对2000个国营工厂的调查表明，采用诊断技术后，维修费用每年可减少3亿英镑；美国Pekrul发电厂实行设备预知维修后，每年能节约的费用为诊断和预防性维修成本的36倍。不难看出，设备故障诊断技术能带来巨大的经济收益。

1.1.2　智能故障诊断方法的优势及挑战

由于现代工程技术系统和设备的复杂性与耦合性，其故障普遍具有多层次性、随机性等特点，一般难以通过理论分析方法在故障类别与征兆之间建立起对应关系。近年来，以专家系统[5-8]、人工神经网络[9-11]、支持向量机[12-16]、模糊理论[17-19]、Petri 网络[20-22]、贝叶斯网络[23-25]等方法为代表的智能故障诊断技术已广泛地应用于工业过程及设备的故障诊断，取得了良好的诊断效果。

同传统的理论分析方法相比，智能故障诊断方法直接以故障实例为基础，通过各种数据归纳和学习算法来建立故障类别与征兆之间的对应关系，发现隐含在故障实例中的故障诊断知识，进而利用这些发现的知识对新故障实例进行故障诊断，从而克服了传统的理论分析方法在故障诊断过程中对复杂的故障机理进行数学建模的困难。

不难看出，故障实例是智能故障诊断方法的基础。在实际的故障诊断问题中，故障实例不可避免地会带有某种程度的不一致性，即具有相同征兆取值的故障实例具有不同的故障类别。归纳起来，故障实例的不一致性主要来源于如下两个方面：一方面，故障产生机理不完全清楚，故障表现形式不唯一，人们对故障征兆的提取通常带有一定的盲目性，从而可能遗漏某些征兆，或者由于现有技术手段的局限，无法获取某些征兆，诸如此类原因引起的故障征兆缺失必然导致利用现有的故障征兆无法对故障实例进行精确分类，以致出现故障实例的不一致；另一方面，故障实例的征兆值在测量过程中不可避免地存在测量噪声，另外在对故障实例的征兆值及故障类别进行记录、处理和整理的过程中也可能会引入计算误差与人为错误，这些测量噪声、计算误差和人为错误等统称噪声，噪声的存在最终也将导致故障实例出现不一致。

由此可见，在实际的故障诊断问题中，故障实例的不一致是普遍存在的，这是各种智能故障诊断方法必须要面临的重要挑战。如何使各种智能方法在故障实例存在不一致的情况下，仍能进行故障诊断知识的有效提取，并应用提取的诊断知识对新故障实例做出尽可能正确的诊断，这无疑是很有实际意义的。

1.1.3　粗糙集方法在不一致信息处理方面的优势

粗糙集方法是波兰数学家 Pawlak[26-28]于 1982 年提出的一种用于处理不完备、不精确、不一致信息的新型数学工具，与其他不一致信息处理工具相比，粗糙集方法不需要数据以外任何初始的或附加的信息，如统计学中的统计概率分布、D-S 证据理论中的基本概率指派函数以及模糊集理论中的模糊隶属度函数等[29-31]。粗糙集方法能够直接对数据进行分析处理，从中提取有用征兆，发现隐含规律，得到简明扼要的规则型知识表达形式。

故障诊断的粗糙集方法以故障实例在现有征兆知识水平下的不可区分性为基础，将故障实例粒化为一系列征兆知识等价类，每个等价类中的故障实例相对于现有征兆知识而言是完全等价和不可区分的，这些等价类形成了征兆知识表达的基本粒子；如果一个故障类的实例集是由上述某些征兆知识等价类组成的，则这个故障类是可定义的，称为一个精确集，否则这个故障类是不可定义的，称为一个粗糙集；当一个故障类的实例集为粗糙集时，上、下近似和边界域能用来对这个实例集进行刻画，下近似定义为包含在这个实例集中征兆知识等价类的最大合集，上近似定义为包含这个实例集的征兆知识等价类的最小合集，边界域定义为上、下近似的差集；利用现有的征兆知识，一个故障类下近似中的故障实例能够被确切地分为这一故障类，而边界域中的故障实例只是可能被分为这一故障类，显然边界域的大小刻画了一个粗糙集在现有征兆知识水平下的不确定性程度；基于上、下近似和边界域的定义，粗糙集方法从故障实例中提取的故障诊断知识能够被区分为确定性知识和可能性知识。

归纳起来，故障诊断的粗糙集方法在不一致信息处理方面具有如下优势。

(1)粗糙集方法在不一致信息处理过程中不需要故障实例以外任何初始的或附加的信息，故障实例的不一致通过边界域表达和处理，由于边界域能够用确定的数学公式来描述，完全由故障实例决定，所以粗糙集方法对故障实例不一致的处理更加客观，也更易于操作。

(2)粗糙集方法从故障实例中提取的故障诊断知识能够被区分为确定性知识和可能性知识，因此，基于这些提取的知识做出的故障诊断决策为确定性和可能性决策，人们可以针对这两种决策结果分别对待和处理。

(3)通过使所有故障类的上、下近似和边界域保持不变，可以对冗余的故障征兆进行约简，约简后的征兆不仅能够保持对故障实例的分类能力不变，而且有助于提取简洁的故障诊断知识，从而能够避免获取和处理冗余征兆所造成的资源浪费，另外许多研究还表明对故障征兆进行约简有助于提高对新故障实例的分类精度。

(4)同神经网络、支持向量机等智能故障诊断方法相比，粗糙集方法获取的故障诊断知识是规则型知识，能够被人们理解。

粗糙集方法在不一致性信息处理方面具有诸多优势，所以该方法已被广泛应用于工业过程及设备的故障诊断过程。文献[32]～文献[34]利用粗糙集方法对故障诊断中的不完备、不确定和不一致数据进行处理，取得了良好的不一致信息处理效果。文献[35]～文献[37]利用粗糙集方法对故障征兆进行约简，明显简化了提取的故障诊断知识，并且提高了对新故障实例的分类精度。文献[38]～文献[40]利用粗糙集方法较强的数据分析能力和容错性，通过提取故障诊断知识，建立了故障诊断专家系统的知识库，并对其进行了有效维护。文献[41]～文献[43]针对故障诊

断中某些征兆的连续取值问题，设计了离散化算法，实现了粗糙集方法在征兆连续取值情况下的故障诊断规则提取。文献[44]～文献[49]分别将粗糙集方法与模糊理论、神经网络、支持向量机、贝叶斯网络等智能方法结合，通过融合各种方法的优势，开展了综合故障诊断技术研究，取得了良好的故障诊断效果。

1.1.4　本书的研究动机及目标

粗糙集方法具有强大的不一致信息处理能力，并且得到的规则型知识容易理解，因此，该方法已经被广泛地应用于工业过程及设备的故障诊断，在故障征兆约简、诊断知识获取和知识库构建等诸多方面表现出巨大的优势。然而，到目前为止，粗糙集方法在实际故障诊断中的泛化性能却很少被研究。粗糙集方法在实际故障诊断中的泛化性能是指粗糙集方法利用从故障实例中提取的故障诊断知识，对新故障实例做出正确诊断的性能，故障诊断的最终目的就是利用现有的经验和知识对新故障实例做出尽可能正确的诊断，因此，好的泛化性能一直是故障诊断方法追求和努力的目标。为此，本书将分别针对一般故障诊断问题以及某些特定故障诊断问题中影响粗糙集方法泛化性能的因素进行深入的分析和探讨，通过引入相应的处理技术，改善粗糙集方法在故障诊断中的泛化性能。

1.2　机器学习方法泛化性能控制的研究进展

20 世纪 60 年代初，Rosenblatt[50]提出了第一个学习机器模型——感知器，这标志着人们对机器学习过程进行数学研究的开始。从概念上讲，感知器的思想并不是新的，它已经在神经生理学领域讨论了多年，但是，Rosenblatt 把这个模型表示为一个计算机程序，并且通过简单的实验说明这个模型能够推广。1962 年，Novikoff[51]证明了关于感知器的第一个定理——收敛性定理，这一定理在机器学习理论的创建过程中发挥了十分重要的作用，它在一定意义上将学习机器具有推广能力的原因与训练集上的错误数最小化原则联系起来。在这一定理结论的基础上，很多学者认为，使学习机器具有推广性的唯一因素就是使其在训练集上的错误数最小，这就是众所周知的经验风险最小化（empirical risk minimization，ERM）原则，持这一观点的学者称为机器学习的应用分析学派。而有些学者认为学习机器的推广能力与训练集上的错误数最小化原则之间的关系并不是不言而喻的，而是需要证明的，这些学者称为机器学习的理论分析学派。因而，对机器学习过程的研究随之形成了应用分析和理论分析两个分支。

关于感知器的实验被人们广为知晓后，应用分析学派很快提出了一些其他类型的学习机器，例如，Widrow 等[52]构造的 Madaline 自适应学习机；Steinbuch 等[53]提出的学习矩阵等。为了解决实际问题，人们还开发了许多计算机程序，如最初

为专家系统设计的决策树[54]，用于语音识别问题的隐马尔可夫模型[55]等。然而，这些方法都没有涉及对一般学习现象的研究，直到1986年，Le Cun[56]提出了利用后向传播技术同时寻找多个神经元的权值，此时才开创了学习机器研究历史的一个新时代。而在构造感知器(1962年)到实现后向传播(1986年)的这段时间里，应用分析学派没有发生特别有重大影响的事情。

与此形成鲜明对比的是，在这段时间里，理论分析学派关于统计学习理论的研究硕果累累。早在1968年Vapnik等[57]就针对指示函数集(即模式识别问题)提出了描述其复杂度的VC熵(vapnik-chervonenkis entropy)和VC维(vapnik-chervonenkis dimension)，并且利用这些概念发现了泛函空间的大数定律，得到了关于收敛速率的非渐近界的主要结论，1971年他们发表了这些工作的完全证明[58]，1974年他们提出了一个全新的机器学习归纳原则——结构风险最小化(structural risk minimization，SRM)原则，指出学习机器的复杂度和训练集上的错误数共同影响了学习机器的推广能力，从而奠定了学习机器泛化性能控制理论的基础，得到了通过控制学习机器复杂度来控制学习机器泛化性能的方法[59]。1976～1981年，最初针对指示函数集得到的结论如大数定律、完全有界和无界函数集一致收敛速率的界以及SRM原则等被推广到了实函数集[60]。

与此同时，学者从其他视角也发现了控制学习机器泛化性能的类似理论。Tikhonov[61]和Ivanov[62]提出了解决不适定问题的正则化理论，这一理论的核心思想是在目标泛函中增加一个关于函数复杂度的正则化项，使目标泛函成为正则化泛函，从而在问题求解过程中能够考虑函数复杂度。密度估计的非参数方法是一个不适定问题，Rosenblatt[63]和Parzen[64]提出的几种解决此类问题的算法所采用的正是正则化技术，密度估计的非参数方法带来了新的统计学算法，弥补了传统的参数方法的缺陷，使人们能从一个较宽的函数集中估计函数。Solomonoff[65]和Kolmogorov[66]在利用信息论方法研究推理问题时，提出了算法复杂度的思想，在此基础上，1978年Rissanen[67]提出了机器学习的最小描述长度归纳原理。1984年Valiant[68]提出了机器学习的可能近似正确学习模型，并利用这一模型分析了学习过程的样本复杂度和推广能力的界，而在这个模型中，学习机器的复杂度VC维扮演着重要的角色。由此可见，在上述这些理论中，对学习机器泛化性能的控制依然是靠对学习机器复杂度的控制来实现的，因此，从本质上讲，这些理论与SRM原则是相同的。

自感知器提出以来，经过20多年对机器学习过程的理论研究，20世纪90年代初，有限样本情况下的机器学习理论逐步成熟起来，形成了较完善的理论体系——统计学习理论。统计学习理论能够在理论上对学习机器的泛化性能提供保证，因此，这一理论已经被广泛地用于分析和控制机器学习方法的泛化性能。

基于SRM原则，1992年Vapnik[69,70]通过构造Δ-间隔分类超平面来控制学习

机器的复杂度，提出了一种新的学习机器——支持向量机。支持向量机能够在理论上对自身的泛化性能提供保障，因此，与以往的基于 ERM 原则的学习机器如神经网络、决策树等相比，具有诸多理论和实践上的优势。目前，支持向量机已经被广泛地应用于解决小样本、非线性、高维模式识别问题，取得了令人满意的效果[71-73]。

此外，SRM 的思想也越来越多地被传统机器学习方法所采用，以改善其泛化性能，一些典型的例子，如决策树学习中被广泛采用的剪枝技术[74,75]、基于最小描述长度和 SRM 的决策树节点规模控制[76-78]、神经网络学习中对神经元规模[79,80]和节点连接权值的控制[81,82]等，实践表明，通过引入相应的复杂度控制技术，这些传统机器学习方法的泛化性能得到了明显的改善。

1.3　粗糙集方法的研究现状

1.3.1　经典粗糙集方法的研究现状

波兰学者 Pawlak[26]于 1982 年正式提出的粗糙集理论是在经典集合论基础上发展起来的处理不完备、不确定、不一致信息的数学工具，与 Zadeh[83]提出来的词计算理论、张铃等[84]提出来的商空间理论合称三大粒度计算理论。1995 年 *Rough Sets*[27]一文的发表，引起了计算机和应用数学领域研究人员的广泛关注，粗糙集理论的研究进入高潮，国内外学者分别从模型性质、算法设计等角度对该方法进行了大量的研究[85,86]。

在模型性质方面，Iwinski 等系统地研究了粗糙集模型的构造化和公理化方法[87-91]，分析了粗糙集模型的数学性质。Bonikowski 等[92]探讨了粗糙集理论的内涵与外延。Wong 等分别比较了粗糙集方法与统计学方法、模糊集方法以及 D-S 证据理论等的异同[93-100]。

对于粗糙集方法的设计，其研究主要围绕如下四方面进行：数值型属性离散化、属性约简、规则提取和推理决策。

在数值型属性离散化方面，等宽和等频是最早与最简单的离散化方法，它们不需要决策类的信息，属于无监督方法，由于属性值的分布通常是不均匀的，并且属性值野点会对这些方法的离散结果产生明显的影响，因此，这些方法在实际应用中通常难以取得令人满意的效果[101]。Holte[102]和 Dougherty 等[103]通过将决策类信息引入等宽和等频离散化方法，分别提出了单规则和最大边缘熵的离散化方法。Catlett[101]通过利用信息熵来度量每个属性离散断点的重要性，进而递归地选取最重要的离散断点，提出了数值型属性离散的 D2 方法。Fayyad 等[104,105]在 D2 方法的基础上，通过引入最小描述长度准则来停止属性离散的递归过程，提出了递归最小熵划分方法，克服了 D2 方法中人为选取停止准则的困难。Kerber[106]通

过利用 χ^2 度量递归地融合相邻的属性离散断点，提出了 ChiMerge 方法，在此基础上，Liu 等[107,108]通过进一步将属性离散过程所引起的数据集不一致程度作为属性离散过程的停止准则，提出了 ChiMerge 方法的自适应版本——χ^2 方法，克服了 ChiMerge 方法中人为选取停止准则的困难，并且 χ^2 方法还有一个显著特点就是能在属性离散过程中删除冗余的属性。除此之外，Mantaras 等[109]和 Cerquides 等[110]提出了基于距离的离散化方法，Ho 等[111]提出了基于 Zeta 度量的离散化方法，Nguyen 等[112-114]提出了基于布尔推理的离散化方法，于达仁等[41]、苗夺谦[115]提出了基于聚类的离散化方法等多种解决数值型属性离散问题的方法。Liu 等[116]对各种属性离散化方法进行了综述和比较，通过实验总结得出：Fayyad 等提出的递归最小熵划分方法通常情况下能获得最好的数值型属性离散效果，而在数值型属性离散化的同时，还需要删除冗余属性，χ^2 方法是一个好选择。

属性约简是粗糙集理论研究的核心问题之一，目前已经提出了许多属性约简算法。Skowron 等[117]在 1991 年提出的基于差别矩阵的属性约简算法揭示了属性约简的结构，这被视为粗糙集理论最为重要的研究成果之一。基于差别矩阵的属性约简方法能够求取全部的属性约简，从而提供了一种发现最小约简的方法[118,119]，然而，该方法具有较大的时间和空间复杂度，并且 Wong 等[120]和 Ziarko[121]已经证明全部属性约简的求取是一个 NP-hard 问题，因此，许多学者提出了属性约简的启发式算法[122-124]。一个经典的启发式属性约简算法是 Michal 等 [125]在他们的粗糙集库(rough set library, RSL)中实现的基于近似质量的启发式算法，该算法首先基于近似质量定义一个属性的重要度，然后递归地选取具有最大重要度的属性，从而得到一个约简。徐章艳等[126]通过分析近似质量在计算属性重要度时存在的问题，设计了一个改进的属性重要度计算公式，构造了一个时间复杂度为 $\max[O(|C||U|), O(|C|^2|U/C|)]$的属性约简算法。Duntsch 等[127]、Miao 等[128]、王国胤等[129]通过对粗糙集模型中知识的不确定性分析，基于信息熵的知识不确定性度量，提出了基于信息熵的启发式属性约简算法，Wang 等[130,131]更进一步地讨论了基于信息熵与基于近似质量的属性约简算法的关系。为了求取最小属性约简，Wang 等[132]和 Wroblewski[133]分别提出了基于粒子群和遗传算法等进化理论的属性约简方法。Min 等[134]发现最小属性约简在很多情况下并不是最好的约简，并不能帮助粗糙集方法获得更好的泛化性能，因此，他们提出了基于最小属性值域空间的属性约简方法，但是在某些情况下仍然不能获得满意的性能。除此之外，叶东毅[135]提出了一种结合启发式算法和差别矩阵算法的属性约简方法，文献[136]～文献[139]将近似质量与信息熵结合起来提出了一种基于粗糙熵的属性约简方法，刘少辉等[140]基于排序思想给出了一种时间复杂度为 $O(|C|^2|U|\log|U|)$ 的属性约简算法，Slezak[141,142]提出了近似属性约简算法，Bazan 等[143,144]提出了动态属性约简算法，

文献[145]～文献[148]研究了不一致信息系统的属性约简等。在目前已经提出的各种属性约简方法中，基于差别矩阵、近似质量和信息熵的方法是最为基本的属性约简方法，其他方法基本上都是对这三种方法在算法效率、性能等方面的改进和融合研究，在实际的属性约简中，这三种最基本的方法仍然是最广泛采用的方法，并且它们也是新的属性约简算法设计的基础。

对于规则提取，在机器学习领域已经被广泛研究，并且已经提出了许多算法，如 AQ[149]、PRISM[150]、CN2[151]等。尽管粗糙集方法没有自己特有的规则提取算法，然而，基于粗糙集的规则提取算法有其自身的特点，这主要体现在其对不一致数据的处理上：粗糙集方法不像其他方法那样对不一致数据进行统计、纠正等预处理操作，而是直接利用上、下近似将这些不一致数据转化为一致数据，然后从这些转化的一致数据中提取确定性和可能性规则[152,153]。一些代表性的粗糙集规则提取算法及软件系统可归纳如下。Grzymala-Busse[152]和 Chan 等[154]基于局部覆盖思想，通过定义启发式条件递归地选取最好的基本规则前件，提出了规则提取的 LEM2 算法，并且在此基础上构建了 LERS 粗糙集学习系统。Skowron 等[117,155]提出了基于差别矩阵和布尔推理的规则提取算法，并且通过融合近似约简、动态约简、数据过滤等技术，开发了 Rosetta 粗糙集学习系统的计算核心[156]。文献[157]～文献[159]基于 Explore 算法，提出了一种全部规则提取算法和一种能够满足用户给定要求的满意规则提取算法，并且在此基础上开发了 RoughFamily 粗糙集学习系统。Ziarko 等[160]提出了基于决策矩阵的规则提取算法，并开发了 KDD-R 粗糙集学习系统，主要针对海量数据的强规则提取。Stefanowski[161]对各种粗糙集规则提取算法进行了综述和比较，将目前主要的规则提取算法分为三类：最小规则提取算法、全部规则提取算法和满意规则提取算法。其中 LEM2 是最小规则提取的典型算法，Explore 是全部规则和满意规则提取的典型算法，终止条件的差异决定了 Explore 算法最终提取的规则类型；在这三类算法中，最小规则提取算法对计算资源的占用最小，更适合于构建分类器，而其他两类算法通常需要占用大量的计算资源，经常被用于其他知识发现任务，特别是利用它们从数据中发现具有潜在兴趣的知识；通过构建对比实验，Stefanowski 等发现这三类算法在分类性能方面并不存在明显的优劣差异，并且与 ID3[54]、C4.5[162]等经典机器学习算法的性能相当，然而，最小规则提取算法获得的规则数目最少，全部规则提取算法获得的规则数目远远超过最小规则提取算法，并且含有大量的冗余规则和弱规则，而满意规则提取算法则能够根据用户给定要求进行规则获取，其获得的规则数目介于上述二者之间。

在推理决策方面，粗糙集方法与其他机器学习方法不存在差异，因此，粗糙集方法可以借鉴机器学习领域的研究成果。在机器学习领域，基于规则的推理决策算法大体可以分为两类：一类是基于排序规则的算法，其代表算法是 C4.5[162]