

<<云模型与粒计算>>

图书基本信息

书名：<<云模型与粒计算>>

13位ISBN编号：9787030350640

10位ISBN编号：7030350642

出版时间：2012-7

出版时间：科学出版社

作者：苗夺谦、王国胤、姚一豫、梁吉业、吴伟志、张燕平

页数：320

字数：427000

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

## <<云模型与粒计算>>

### 内容概要

云模型是研究定性概念与定量数值之间相互转换的不确定性认知模型。

粒计算是当前计算智能研究领域模拟人类思维和解决复杂问题的新方法。

它覆盖了所有有关粒度的理论、方法和技术,是研究复杂问题求解、海量数据挖掘和模糊信息处理等问题的有力工具。

《云模型与粒计算》介绍云模型与粒计算交叉研究的最新进展,由国内外相关领域的华人学者撰文14章,内容涉及云模型、高斯云的数学性质、云模型与相近概念的关系、区间集、区间值信息系统的粒计算模型与方法、多粒度粗糙集、粒计算模型的特性分析与比较、云计算环境下层次粗糙集模型约简算法、基于粒计算的聚类分析、并行约简与F-粗糙集、单调性分类学习、不确定性研究中若干问题的探讨、基于云模型的文本分类应用、数据挖掘算法的云实现。

《云模型与粒计算》可供计算机、自动化等相关专业的研究人员、教师、研究生、高年级本科生和工程技术人员参考。

<<云模型与粒计算>>

作者简介

苗夺谦、王国胤、姚一豫、梁吉业、吴伟志、张燕平

## &lt;&lt;云模型与粒计算&gt;&gt;

## 书籍目录

前言第1章 双向认知计算模型——云模型1.1 引言1.2 正态分布与正态隶属函数1.3 云模型1.3.1 云模型的定义1.3.2 云模型的数字特征1.3.3 正态云模型的递归定义及其数学性质1.3.4 云发生器1.3.5 双向认知计算模型1.4 本章小结参考文献第2章 高斯云的数学性质2.1 高斯云分布2.2 高斯云的数学性质2.2.1 高斯云的数字特征2.2.2 高斯云的期望曲线2.2.3 高斯云的雾化特性2.2.4 高阶高斯云的数字特征2.3 高斯云的参数对峰度的影响分析2.4 高斯云的幂律特性实验2.5 本章小结参考文献第3章 云模型与相近概念的关系3.1 二型Fuzzy集3.1.1 二型Fuzzy集的定义3.1.2 二型Fuzzy集的运算3.1.3 区间值Fuzzy集3.1.4 Gaussian二型Fuzzy集3.1.5 二型Fuzzy集的嵌入区间值Fuzzy集3.1.6 m型Fuzzy集与Genuine集3.1.7 区间集与阴影集3.2 直觉Fuzzy集3.2.1 直觉Fuzzy集的定义3.2.2 直觉Fuzzy集的运算3.2.3 区间值直觉Fuzzy集3.3 Neumaier云3.3.1 Neumaier云的定义3.3.2 离散云3.3.3 连续云与潜云3.4 Fuzzy概率集3.4.1 随机集3.4.2 Fuzzy概率集3.4.3 Bifuzzy概率集和区间值概率集3.5 Soft集3.5.1 Soft集的定义3.5.2 Soft集的运算3.6 云模型3.6.1 云模型的定义3.6.2 云模型算法3.6.3 正态云生成的区间值Fuzzy集3.7 云集3.7.1 各类集合的关系3.7.2 云集3.8 本章小结参考文献第4章 区间集4.1 引言4.2 不精确概念及其表示4.3 区间集4.3.1 区间集与部分已知概念4.3.2 区间集与概念近似4.4 区间集代数4.4.1 幂代数4.4.2 区间集运算4.4.3 基于包含序的区间集代数4.4.4 基于知识的区间集代数4.5 基于不完备信息表的区间集构造方法4.6 区间集与其他理论的联系4.6.1 区间集与Kleene三值逻辑4.6.2 区间集与粗糙集4.6.3 区间集与三支决策4.6.4 区间集、模糊集和云模型4.7 本章小结参考文献第5章 区间值信息系统的粒计算模型与方法5.1 引言5.2 基础概念5.2.1 不可分辨关系和近似域5.2.2 决策系统中的不确定性度量5.3 区间值决策系统的不确定性度量5.3.1 区间值的相似关系5.3.2 相似类和决策类5.3.3  $\beta$ -条件熵5.3.4 不确定性度量和粗糙粗糙决策熵5.4 实验5.5 本章小结参考文献第6章 多粒度粗糙集6.1 问题描述6.2 乐观多粒度粗糙集6.2.1 Pawlak粗糙集理论6.2.2 乐观粗糙近似6.2.3 多粒度粗糙集中的几个度量6.2.4 特征选择6.3 悲观多粒度粗糙集6.3.1 悲观粗糙近似6.3.2 粗糙成员函数6.3.3 多粒度粗糙集中的规则6.4 本章小结参考文献第7章 粒计算模型的特性分析与比较7.1 引言7.2 不确定性表示方法7.2.1 隶属度的方法7.2.2 粗糙集的表示方法7.2.3 商空间的表示方法7.3 粒计算表示不确定性方法之间的关系7.3.1 隶属度函数表示方法与商空间链表示方法的关系7.3.2 粗糙集的表示方法与商空间链表示方法之间的关系7.3.3 云模型与二型模糊之间的关系7.4 问题求解方法的比较7.5 本章小结参考文献第8章 云计算环境下层次粗糙集模型约简算法8.1 层次粗糙集模型8.1.1 引言8.1.2 概念层次8.1.3 基于云模型的概念提取及概念提升8.1.4 层次粗糙集模型8.2 云计算技术8.2.1 云计算介绍8.2.2 MapReduce技术8.3 云计算环境下层次粗糙集模型约简算法8.3.1 云计算环境下知识约简算法中的并行性分析8.3.2 云计算环境下计算层次编码决策表算法8.3.3 云计算环境下层次粗糙集模型约简算法的研究8.4 实验与分析8.4.1 理论分析8.4.2 实验结果8.4.3 实验分析8.5 本章小结参考文献第9章 基于粒计算的聚类分析9.1 引言9.2 粒度计算与聚类分析的关系9.3 粒聚类的基本方法9.3.1 模糊聚类分析9.3.2 粗糙集聚类分析9.3.3 商空间聚类分析9.4 基于融合的粒度模型的聚类分析9.4.1 模糊集与粗糙集的结合9.4.2 模糊商空间9.5 多粒度聚类若干问题的研究9.5.1 多粒度聚类中粒子的转换问题9.5.2 约简集粒度的精准性9.5.3 多粒度快速聚类算法9.6 基于多粒度聚类的问题求解应用举例:粗糙RBF神经网络的学习算法9.6.1 粗糙RBF神经网络的学习算法9.6.2 粗糙RBF神经网络的可用性与可靠性实验9.7 本章小结参考文献第10章 并行约简与F-粗糙集10.1 粗糙集基本知识10.2 F-粗糙集10.3 并行约简定义与性质10.4 并行约简算法10.4.1 基于属性重要度矩阵的并行约简算法10.4.2 基于属性重要度矩阵的并行约简算法的优化10.4.3 基于F隶属属性重要度的并行约简算法10.4.4 (F,  $\beta$ )-并行约简10.5 决策系统的分解10.6 本章小结参考文献第11章 单调性分类学习11.1 引言11.2 基于优势关系粗糙集的单调性分类分析11.3 基于模糊偏好粗糙集的单调性分类分析11.4 基于排序熵模型的单调性分类分析11.4.1 Shannon信息熵11.4.2 有序信息熵11.5 基于排序熵的单调性决策树11.5.1 程序描述11.5.2 性质研究11.5.3 在人工数据上的实验11.6 本章小结参考文献第12章 不确定性研究中若干问题的探讨12.1 隶属度的不确定性问题12.2 运算法则的不确定性12.3 模糊运算与逻辑运算问题12.3.1 模糊运算12.3.2 逻辑运算12.4 排序的不确定性12.5 截集水平的不确定性12.6 Fuzzy集合的互补律问题12.7 集合的统一问题12.8 本章小结参考文献第13章 基于云模型的文本分类应用13.1 云模型在文本挖掘中的理论扩充13.1.1 基于VSM模型的文本知识表示13.1.2 基于信息表的文本知识表示13.1.3 基于云模型的文本信息表转换13.1.4 基于云相似度的文本相似度量13.2 文本分类及其常用方

## <<云模型与粒计算>>

法13.2.1 文本分类概述13.2.2 文本分类常用方法13.2.3 性能分析13.2.4 文本分类模型的评估13.3 基于云模型与粒计算的文本分类13.3.1 虚拟泛概念树及概念跃升13.3.2 基于云模型的文本特征自动提取算法13.3.3 基于云概念跃升的文本分类13.4 本章小结参考文献第14章 数据挖掘算法的云实现14.1 在云上实现数据挖掘算法的技术背景14.2 现有基于云计算的数据挖掘平台14.2.1 “大云”系统14.2.2 Mahout开源项目14.2.3 电子科技大学与华为公司合作的云挖掘项目14.3 经典数据挖掘算法的MapReduce实现思路14.4 经典数据挖掘算法在Hadoop平台的实现范例14.4.1 协同过滤算法在Hadoop平台的实现14.4.2 朴素贝叶斯算法在Hadoop平台的实现14.5 云挖掘技术的展望14.5.1 针对Web信息的云挖掘14.5.2 针对图结构的云挖掘14.5.3 针对声音与视频等多媒体信息的云挖掘参考文献

## 章节摘录

版权页：插图：13.3.2 基于云模型的文本特征自动提取算法 在传统数据挖掘中，只能对结构化的数据信息进行处理，能够处理的特征项也较少。

而对于非结构化文本信息则处理困难，这主要是由于文本特征向量的高维性造成的。

基于此，文本的特征表示成为文本挖掘面临的首要任务。

一个合理的文本特征表示方法具有以下两个特征：一方面需要包含足够的信息量，文本信息的关键特征不能丢失太多；另一方面，文本特征项也不能过多，否则会严重影响挖掘算法性能。

出于这两方面的考虑，产生了文本特征的选择。

如何选择适当的特征集是文本挖掘中一个十分重要的问题。

过大或过小的特征空间都会严重影响到文本挖掘性能。

一个有效的文本特征集合必须满足文本内容表示完整、较强的文本区分能力与保证尽量小的特征维度特点[10—12]。

目前，一般采用领域专家的经验知识来进行特征集大小的设定，主要有直接据经验设定特征数（PFC）、按特征比例设定（THR）、通过统计量（MVS）或向量的空间稀疏性（SPA）等选择方法[13—15]。

通过实际应用测试，在特定语料库上这些方法较好地解决了特征集空间的大小问题，但缺少理论基础支撑，往往具有较强的主观因素充斥其中。

当语料库改变或有所调整时，需重新进行设定，对于大规模的自动分类聚类将不再适用。

综上所述，若能找到一种具有较强数据适应能力，能够进行文本特征自动提取的方法，将极大促进文本挖掘的研究进展。

1. 文本特征提取概述 从理论上来说，特征项越多越能更好地进行表示。

但大量实践显示，事实并非总是如此。

过大的特征空间将导致TM时间空间复杂度激增，计算代价更为高昂。

随着Internet的广泛深入与信息资源的日益增加，寻求一种高性能的文本特征提取算法在TM中将变得越来越重要。

定义13—4 特征提取。

特征提取（feature selection，FS）是选择描述目标的最佳特征子集的过程，得到的特征子集包含于原始特征集。

即 $F \rightarrow F'$ ，其中 $F$ 是原始特征集， $F'$ 是选择后的特征子集。

FS不仅能提高文本挖掘算法的运行速度，降低占用的内存空间；另一方面由于去掉了不相关或相关程度低的特征，可提高许多TM任务的性能，如文本分类的分类质量等。

## <<云模型与粒计算>>

### 编辑推荐

《云模型与粒计算》覆盖了所有有关粒度的理论、方法和技术，是研究复杂问题求解、海量数据挖掘和模糊信息处理等问题的有力工具。

《云模型与粒计算》可供计算机、自动化等相关专业的研究人员、教师、研究生、高年级本科生和工程技术人员参考。

<<云模型与粒计算>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介, 请支持正版图书。

更多资源请访问:<http://www.tushu007.com>