

论文

广义粗糙集理论及实值属性约简

肖迪 张军峰

南京工业大学 自动化学院 南京航空航天大学 自动化学院

摘要:

针对经典粗糙集理论仅能处理离散化数据的局限性, 提出属性和属性子集的广义重要度的概念以及空间中的广义近邻关系, 并提出了广义近邻关系下的广义粗糙集扩展模型。广义粗糙集理论利用广义近邻关系在全局中划分相容模块, 构成集合的下、上近似集, 避免了经典粗糙集理论必须量化数据的麻烦。另外, 提出了广义粗糙集的实值属性约简的一种贪心算法, 并分析了约简属性集合的质量。最后通过实例验证了所提方法的正确性和有效性。

关键词: 数据挖掘 广义粗糙集理论 广义重要度 近似约简

Generalization rough set theory and real-valued attributes reduction

Abstract:

Considering that the classical rough set theory can only process the discrete data, the degree of general importance of an attribute and attribute subsets was presented. And then a generalization rough set theory was proposed based on the general near neighborhood relation. The theory partitioned the universe into the tolerant modules and formed lower approximation and upper approximation of the set under general near neighborhood relationship which avoided the discretization in Pawlak's rough set. Furthermore, the definition of attribute reduction in generalization rough set and its greedy algorithm were proposed. Finally, results of some examples show the correctness and validity of this method.

Keywords: data mining general rough set theory degree of general importance approximation reduction

收稿日期 2007-12-21 修回日期 2008-02-02 网络版发布日期

DOI:

基金项目:

通讯作者: 肖迪

作者简介:

参考文献:

本刊中的类似文章

1. 周炎涛; 郭如冰; 李肯立; 吴正国. 基于前馈多层感知器的网络入侵检测的多数据包分析[J]. 计算机应用, 2006,26(4): 806-808
2. 张文字; 张铭华. 基于面向属性泛化及信息增益的数据挖掘方法研究[J]. 计算机应用, 2006,26(4): 861-863
3. 崔晓军 薛永生. 模糊时序关联规则挖掘[J]. 计算机应用, 2007,27(3): 561-564
4. 夏卓群 程昱 梁涤青 . 挖掘多粒度时间下异步周期的模式[J]. 计算机应用, 2006,26(12): 2985-2987
5. 刘以安 羊斌. 关联规则挖掘中对Apriori算法的一种改进研究[J]. 计算机应用, 2007,27(2): 418-420
6. 刘云生 阙小锋 . 基于模糊依赖的实时数据库的快速恢复[J]. 计算机应用, 2007,27(1): 74-76
7. 唐常杰 陈瑜 张欢 段磊 . 基于转基因GEP 的公式发现[J]. 计算机应用, 2007,27(10): 2358-2360
8. 曹洪其 孙志挥 . 基于网格技术的高维大数据集离群点挖掘算法[J]. 计算机应用, 2007,27(10): 2369-2371

扩展功能

本文信息

- ▶ Supporting info
- ▶ PDF(604KB)
- ▶ [HTML全文]
- ▶ 参考文献

服务与反馈

- ▶ 把本文推荐给朋友
- ▶ 加入我的书架
- ▶ 加入引用管理器
- ▶ 引用本文
- ▶ Email Alert
- ▶ 文章反馈
- ▶ 浏览反馈信息

本文关键词相关文章

- ▶ 数据挖掘
- ▶ 广义粗糙集理论
- ▶ 广义重要度
- ▶ 近似约简

本文作者相关文章

- ▶ 肖迪
- ▶ 张军峰

PubMed

- ▶ Article by
- ▶ Article by

9. 马丽生 邓辉文 齐逸 .一种新的最大频繁项目集挖掘算法[J]. 计算机应用, 2006,26(11): 2670-2673
10. 陈明 史忠植 王文杰 .一种有效的基于图的关联规则挖掘算法[J]. 计算机应用, 2006,26(11): 2654-2656
11. 刘齐宏 唐常杰 李川 刘齐巍 曾涛 蒋永光.基于属性归纳的中药方剂数据挖掘[J]. 计算机应用, 2007,27(2): 449-452
12. 张雨佳 赵会群 吴杰伟 .数据挖掘算法在排球比赛技战术分析中的应用研究[J]. 计算机应用, 2006,26(12): 3027-3029
13. 文娟娟 柴玉梅 .基于免参数据挖掘的相异度度量研究[J]. 计算机应用, 2006,26(12): 2982-2984
14. 张建明; 宋迎清; 周四望; 欧阳竟成.无线传感器网络中数据汇聚技术的研究[J]. 计算机应用, 2006,26(6): 1273-1278
15. 高原; 耿国华; 王怡.基于动态矩形的聚类方法的设计与实现[J]. 计算机应用, 2006,26(4): 870-871
16. 鲁红英; 肖思和.基于改进的遗传神经网络数据挖掘方法研究[J]. 计算机应用, 2006,26(4): 878-879
17. 崔贯勋 朱庆生.一种改进的基于密度的离群数据挖掘算法[J]. 计算机应用, 2007,27(3): 559-561
18. 张慧哲 王坚.多重最小支持度频繁项集挖掘算法研究[J]. 计算机应用, 2007,27(9): 2290-2293
19. 曹明 闪四清 梁海燕 .基于数据挖掘的财务预警模型设计与实现[J]. 计算机应用, 2006,26(10): 2421-2424
20. 石淼磊 冯登国 苏璞睿 .Web浏览器历史数据自动分类取证系统[J]. 计算机应用, 2006,26(10): 2427-2429
21. 吴小兰 王忠群 刘涛 王勇 .基于扩展元胞自动机的在线零售站点的自适应[J]. 计算机应用, 2006,26(10): 2430-2432
22. 唐常杰 张悦 唐良 李川 陈瑜.亚复杂系统中动力学干预规则挖掘技术研究进展[J]. 计算机应用, 2008,28(11): 2732-2736
23. 吴甲 陈峻.一种快速的频繁子图挖掘算法[J]. 计算机应用, 2008,28(10): 2533-2536
24. 张师超 曾德胜 王日凤 谢冲 .大型时态数据库中的Burst 模式挖掘[J]. 计算机应用, 2006,26(10): 2413-2416
25. 陈治平 王雷 .基于密度梯度的聚类算法研究[J]. 计算机应用, 2006,26(10): 2389-2392
26. 邬依林 李中华 毛宗源 .自适应人工免疫算法在数据挖掘中的应用[J]. 计算机应用, 2006,26(8): 1943-1946
27. 崔晓军 薛永生 .基于日历的时序关联规则挖掘算法[J]. 计算机应用, 2006,26(8): 1898-1899
28. 唐常杰 刘威 温粉莲 乔少杰 .社会网络分析和社团信息挖掘的三项探索——挖掘虚拟社团的结构、核心和通信行为[J]. 计算机应用, 2006,26(9): 2020-2023
29. 周钦亮; 李玉忱; 公爱国.一种新的高效生成FP-Tree条件模式基的算法[J]. 计算机应用, 2006,26(6): 1418-1421
30. 刘博; 彭宏; 郑启伦.一种新的数据预处理算法——NLCA[J]. 计算机应用, 2006,26(6): 1406-1408
31. 戴维迪; 何丕廉; 刘振华; 张倩.一种基于窗口的快速聚类算法——F-CABDET[J]. 计算机应用, 2006,26(5): 1127-1129
32. 江伟 陈龙 王国胤 .用户行为异常检测在安全审计系统中的应用[J]. 计算机应用, 2006,26(7): 1637-1639
33. 吴江 唐常杰 段 磊 李太勇.基于基因表达式编程的信用评估模型挖掘方法[J]. 计算机应用, 2007,27(4): 877-880
34. 孙林 徐久成 马媛媛.基于新的条件熵的决策树规则提取方法[J]. 计算机应用, 2007,27(4): 884-887
35. 周晓庆 李明东 肖顺文.基于KW、MW及DW的决策支持系统架构研究[J]. 计算机应用, 2007,27(4): 1027-1029
36. 曾小宁 肖水晶.一种增量更新算法在数据挖掘中的应用[J]. 计算机应用, 2007,27(6): 1403-1406
37. 沈鸿 颜晓龙.一种适用于高维数据流的子空间聚类方法[J]. 计算机应用, 2007,27(7): 1680-1684
38. 张建锦 吴渝 刘小霞.一种改进的密度偏差抽样算法[J]. 计算机应用, 2007,27(7): 1695-1698
39. 岳静 冯鑫 周永华.一种面向用户的ontology进化模型[J]. 计算机应用, 2007,27(7): 1767-1769
40. 常新功 寇纪淞 李敏强.基于带状态回溯个体进化的子结构发现[J]. 计算机应用, 2007,27(8): 1944-1947
41. 陈鑫影 李雄飞.基于粗糙集理论的并行约简算法[J]. 计算机应用, 2007,27(8): 1964-1966
42. 李爱国 赵华.基于PPR的煤矿瓦斯监测数据相似搜索方法[J]. 计算机应用, 2008,28(10): 2721-
43. 潘薇.基于数据挖掘的四维飞行轨迹预测模型[J]. 计算机应用, 2007,27(11): 2637-2639
44. 常鹏 陈; 耿 朱玉全.一种分布式序列模式挖掘算法[J]. 计算机应用, 2008,28(11): 2964-2966
45. 刘天鹏 周娅.P2P网络中的数据挖掘[J]. 计算机应用, 2008,28(1): 162-164,
46. 蒋盛益 李霞.一种改进的BIRCH聚类算法[J]. 计算机应用, 2009,29(1): 293-296
47. 王自强 钱旭.基于KDA和SVM的文档分类算法[J]. 计算机应用, 2009,29(2): 416-418
48. 郑莉华 陈佳.基于贝叶斯网络的电话费欺诈模型的研究及应用[J]. 计算机应用, 2008,28(2): 511-512

49. 张亚玲 康立锦.基于数据挖掘的Snort系统改进模型[J]. 计算机应用, 2009,29(2): 409-411
50. 赵小燕 张朝晖.基于分层神经网络模型的数据挖掘算法[J]. 计算机应用, 2009,29(3): 795-797
51. 孙亚 钱洪波 叶亮.数据挖掘算法在交通状态量化及识别的应用[J]. 计算机应用, 2008,28(3): 738-741
52. 谭颖 胡瑞飞 殷国富.多密度阈值的DBSCAN改进算法[J]. 计算机应用, 2008,28(3): 745-748
53. 郑向群 赵政.基于S-CART决策树的多关系空间数据挖掘方法[J]. 计算机应用, 2008,28(3): 749-752
54. 李云 徐涛 田素方 李拓.带兴趣度的序列概念格模型及其构造[J]. 计算机应用, 2008,28(3): 726-728
55. 高集荣 田艳 邵海英.一种新的用户事务算法[J]. 计算机应用, 2009,29(4): 1099-1101
56. 孔鹏程 张继福.基于离散区间的频繁嵌入式子树挖掘算法[J]. 计算机应用, 2009,29(4): 1120-1123
57. 王金龙 耿雪玉.基于研究者发文序列的研究领域扩散[J]. 计算机应用, 2008,28(6): 1424-1426
58. 常新功 寇纪淞 李敏强.进化子结构发现在区域经济研究中的应用[J]. 计算机应用, 2008,28(5): 1173-1176
59. 邱小宁 陈治平.基于人工免疫系统的客户流失分析[J]. 计算机应用, 2008,28(7): 1705-1708
60. 胡彧 闫巧梅.滑动窗口模型下的优化数据流聚类算法[J]. 计算机应用, 2008,28(6): 1414-1416
61. 陈超祥 叶时平 华成 金林樵.基于XML的完全频繁查询模式挖掘算法[J]. 计算机应用, 2008,28(6): 1450-1453
62. 刘造新.基于本体的XML关联规则挖掘方法[J]. 计算机应用, 2008,28(9): 2318-2320

文章评论 (请注意:本站实行文责自负, 请不要发表与学术无关的内容!评论内容不代表本站观点.)

反馈人	<input type="text"/>	邮箱地址	<input type="text"/>
反馈标题	<input type="text"/>	验证码	<input type="text"/> 3000