

研究、探讨

新的基于最近邻聚类的属性离散化算法

王杰, 姜国强

郑州大学 电气工程学院, 郑州 450001

收稿日期 2005-5-13 修回日期 2008-7-30 网络版发布日期 接受日期

摘要 连续属性离散化是知识发现研究中重要的预处理过程, 基于最近邻聚类和粗集的相关理论, 提出一种新的有监督的多属性离散化方法。该算法分两个阶段来处理, 首先利用最近邻聚类动态调整聚类的类别数, 生成初始聚类。然后基于类信息的相似性定义合并相似区间, 减少了聚类区间。通过实例分析, 该算法是非常有效的。

关键词 [离散化](#) [最近邻聚类](#) [粗集](#) [近似分类质量](#)

分类号 [TP301.6](#)

Novel attribute discretization algorithm based on nearest neighbor-clustering

WANG Jie, JIANG Guo-qiang

College of Electrical Engineering, Zhengzhou University, Zhengzhou 450001, China

Abstract

Discretization of continuous attribute is an important pretreatment process in the knowledge discovery study. A novel algorithm of supervised discretization of continuous attributes based on nearest neighbor-clustering and the theories of rough set is introduced. The algorithm is dealt with in two stages, in the first stage, the number of clusters is adjusted dynamically by using the nearest neighbour-clustering, and then the initial clusters are determined. In the second stage, the adjacent region is merged by adopting the definition of merging similar range, and then the clusters are reduced. The algorithm is proved effectively by the case analysis.

Key words [discretization](#) [nearest neighbor-clustering](#) [rough set](#) [approximate classification quality](#)

DOI: 10.3778/j.issn.1002-8331.2009.24.013

通讯作者 王杰 wj@zzu.edu.cn

扩展功能

本文信息

- ▶ [Supporting info](#)
- ▶ [PDF\(370KB\)](#)
- ▶ [\[HTML全文\]\(0KB\)](#)
- ▶ [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](#)
- ▶ [复制索引](#)
- ▶ [Email Alert](#)
- ▶ [文章反馈](#)
- ▶ [浏览反馈信息](#)

相关信息

- ▶ [本刊中 包含“离散化”的相关文章](#)
- ▶ [本文作者相关文章](#)
- [王杰](#)
- [姜国强](#)