

博士论坛

基于粗糙集理论的属性离散化算法

陈昊^{1,2}, 张旻^{1,2}, 杨俊安^{1,2}

1. 解放军电子工程学院, 合肥 230037

2. 安徽省电子制约技术重点实验室, 合肥 230037

收稿日期 2007-12-25 修回日期 2008-1-28 网络版发布日期 2008-4-24 接受日期

摘要 决策系统中连续属性离散化, 即将一个连续属性分为若干属性区间并为每个区间确定一个离散型数值, 对后继阶段的机器学习具有重要的意义。首先研究了满足决策系统最优划分的一种计算候选断点集合的算法, 然后在基于条件属性重要度和贪心算法的基础上提出了一种确定结果断点子集的新启发式算法。所提出的属性离散化算法考虑并体现了粗糙集理论的基本特点和优点, 并能取得较理想的连续属性离散化结果。

关键词 [粗糙集](#) [最优划分](#) [离散化](#) [候选断点](#) [结果断点](#)

分类号

Method of data discretization based on rough set theory

CHEN Hao^{1,2}, ZHANG Min^{1,2}, YANG Jun-an^{1,2}

1. Electronic Engineering Institute, Hefei 230037, China

2. Key Laboratory of Electronic Restriction, Anhui Province, Hefei 230037, China

Abstract

The discretization of continuous attributes values of a decision system which divides continuous values into different space and allocates some discrete values to each space is always with great contribution to the machine learning. This paper studies a new algorithm of computing candidate cuts for best partition in decision system at first, and proposes one heuristic method based on the importance of condition attributes and greedy algorithm. The two algorithms consider specialty of rough set and embody the advantages of this theory. Moreover, excellent discretization results may be expected from them.

Key words [rough set](#) [best partition](#) [discretization](#) [candidate cuts](#) [result cuts](#)

DOI:

通讯作者 陈昊 ritian99@21cn.com

扩展功能

本文信息

- ▶ [Supporting info](#)
- ▶ [PDF\(463KB\)](#)
- ▶ [\[HTML全文\]\(0KB\)](#)
- ▶ [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“粗糙集”的相关文章](#)
- ▶ 本文作者相关文章

- [陈昊](#)
- [张旻](#)
- [杨俊安](#)