

数据库、信号与信息处理

## 一种改进的决策树分类属性选择方法

王 苗<sup>1</sup>, 柴瑞敏<sup>2</sup>

1. 辽宁工程技术大学 研究生院, 辽宁 葫芦岛 125105

2. 辽宁工程技术大学 电子与信息工程学院, 辽宁 葫芦岛 125105

收稿日期 2009-10-21 修回日期 2009-12-28 网络版发布日期 2010-3-11 接受日期

**摘要** 分析了ID3算法的基本原理、实现步骤及现有两种改进分类算法的优缺点, 针对ID3算法的取值偏向问题和现有两种改进算法在分类时间、分类精确度方面存在的不足, 提出了一种新的分类属性选择方案, 并利用数学知识对其进行了优化。经实验证明, 优化后的方案克服了ID3算法的取值偏向问题, 同时在分类时间及分类精确度方面优于ID3算法及现有两种改进的分类算法。

**关键词** [数据挖掘](#) [决策树](#) [属性选择](#)

**分类号** [TP399](#)

## Improved classification attribute selection scheme for decision tree

WANG Miao<sup>1</sup>, CHAI Rui-min<sup>2</sup>

1. Institute of Graduate, Liaoning Technical University, Huludao, Liaoning 125105, China

2. School of Electronic and Information Engineering, Liaoning Technical University, Huludao, Liaoning 125105, China

### Abstract

Analyze the basic principles and implementation steps of ID3 and point out the advantages and disadvantages of two existing improved classification algorithms. With the shortcoming of inclining to choose attributes having many values for ID3 and the deficiencies of classification time and classification accuracy for existing two improved classification algorithms, a new attribute selection scheme is proposed and optimized with mathematical knowledge. Experiment results show that the optimized scheme can overcome the above disadvantage of ID3 and has the advantages of classification time and classification accuracy over the existing two classification algorithms.

**Key words** [data mining](#) [decision tree](#) [attributes selection](#)

DOI: 10.3778/j.issn.1002-8331.2010.08.036

通讯作者 王 苗 [hsmobei16525@sina.com](mailto:hsmobei16525@sina.com)

### 扩展功能

#### 本文信息

▶ [Supporting info](#)

▶ [PDF\(548KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [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“数据挖掘”的相关文章](#)

▶ 本文作者相关文章

· [王 苗](#)

· [柴瑞敏](#)