

## 数据库与人工智能

### 基于相关性分析及遗传算法的高维数据特征选择

任江涛, 黄焕宇, 孙婧昊, 印鉴

中山大学计算机科学系

收稿日期 2005-12-6 修回日期 网络版发布日期 接受日期

**摘要** 特征选择是模式识别及数据挖掘等领域的重要问题之一。针对高维数据对象, 特征选择一方面可以提高分类精度和效率, 另一方面可以找出富含信息的特征子集。针对此问题, 提出了一种综合了filter模型及wrapper模型的特征选择方法, 首先基于特征与类别标签的相关性分析进行特征筛选, 只保留与类别标签具有较强相关性的特征, 然后针对经过筛选而精简的特征子集采用遗传算法进行随机搜索, 并采用感知器模型分类错误率作为评价指标。实验结果表明, 该算法可有效地找出具有较好的线性可分离性的特征子集, 从而实现降维并提高分类精度。

**关键词** [特征选择](#), [相关性](#), [遗传算法](#)

分类号

**DOI:**

对应的英文版文章: [5126454](#)

通讯作者:

任江涛 [issrjt@mail.sysu.edu.cn](mailto:issrjt@mail.sysu.edu.cn); [renjt@yahoo.com](mailto:renjt@yahoo.com)

作者个人主页: 任江涛; 黄焕宇; 孙婧昊; 印鉴

## 扩展功能

### 本文信息

- ▶ [Supporting info](#)
- ▶ [PDF \(591KB\)](#)
- ▶ [\[HTML全文\]\(0KB\)](#)
- ▶ [参考文献\[PDF\]](#)
- ▶ [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](#)

### 相关信息

- ▶ [本刊中 包含“特征选择,相关性,遗传算法”的 相关文章](#)
- ▶ [本文作者相关文章](#)

· [任江涛](#)  
· [黄焕宇](#)  
· [孙婧昊](#)  
· [印鉴](#)