

人工智能及识别技术

基于聚类和遗传交叉的少数类样本生成方法

杜娟, 衣治安, 周颖

(大庆石油学院计算机与信息技术学院, 大庆 163318)

收稿日期 修回日期 网络版发布日期 接受日期

摘要 传统的分类算法在处理不均衡样本数据时, 分类器预测倾向于多数类, 样本数量少的类别分类误差大。针对该问题, 提出一种基于聚类和遗传交叉的少数类样本上采样方法, 通过K-means算法将少数类样本聚类分组, 在每个聚类内使用遗传交叉获取新样本, 并进行有效性验证。基于K-最近邻及支持向量机分类器的实验结果证明了该方法的有效性。

关键词 [不均衡数据集](#); [分类](#); [聚类](#); [遗传交叉](#)

分类号 [TP391](#)

DOI:

通讯作者:

作者个人主页: [杜娟](#); [衣治安](#); [周颖](#)

扩展功能

本文信息

- ▶ [Supporting info](#)
- ▶ [PDF](#) (83KB)
- ▶ [\[HTML全文\]](#) (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“不均衡数据集; 分类; 聚类; 遗传交叉” 的相关文章](#)
- ▶ [本文作者相关文章](#)