

doi:10.1631/FITEE.1700804

题目：联合局部学习和组稀疏回归的无监督特征选择

概要：近十年，特征选择备受关注。通过挑选特征子集，可有效提升学习算法效率。由于难以获取标签信息，无监督特征选择算法相较于有监督特征选择算法应用更为广泛，其关键在于找出更能反映数据分布的特征集合。由于数据集中存在冗余和噪声，使用全部特征并不能很好展现数据的真实分布。为解决这一问题，本文提出联合局部学习和组稀疏回归的无监督特征选择算法。将基于局部学习聚类方法与组稀疏回归算法有机整合，选出有效反映数据流形分布同时保持组稀疏结构的特征。通过迭代算法，回归系数汇聚到重要特征上，选出能得到更优聚类效果的特征。对多个实际数据集（图像、声音和网页）的实验证明了该算法的有效性。

关键词：无监督；局部学习；组稀疏回归；特征选择