基于类属特征的高阶关系多标记分类方法研究

一.成果完成人及所在单位： 魏博文，计算机与信息工程学院 指导教师：翁伟

二.成果简介：

以多标记分类为研究对象，将特征选择与关系挖掘方法相结合，从而把多标记分类中类属特征的选择和标记关系融入到一个框架中。为全面地考虑现实世界标记关系，方法融入了全局、局部、高阶标记关系。同时结合类属特征，为每个标记选出最具判别性的特征。通过选取多个不同大小、不同领域的真实多标记数据集，以及当下多种先进的多标记分类方法，在多个评价指标上做对比实验验证了本项目研究方法的先进性和优越性。这项研究对提高在类不平衡和维度灾难等问题下的多标记分类性能提供了思路和方法。

三. 主要研究内容及创新点：

针对在类不平衡和特征维度灾难下的多标记分类为研究对象，建立了融合类属特征和标记关系的多标记分类方法（见图 1）。方法首先采用余弦相似度和标记邻域信息分别建立全局和局部标记关系，同时建立带有 l_1 范数的线性回归模型进行类属特征的选择。随后将全局和局部标记关系抽象成数学表达式，加入到带有类属特征选择的模型中，建立起目标函数，从而将问题转化为目标优化。经验证该目标函数为凸，选用近端加速梯度优化方法进行凸优化求解。通过在类属特征选择和多标记分类两个维度上的充分实验及验证。研究结果指出，本项目提出的模型在一定程度上解决了多标记分类中的类不平衡和维度灾难等问题，兼顾了类属特征和标记关系，同现有多标记分类方法相比展现出更优异的效果。

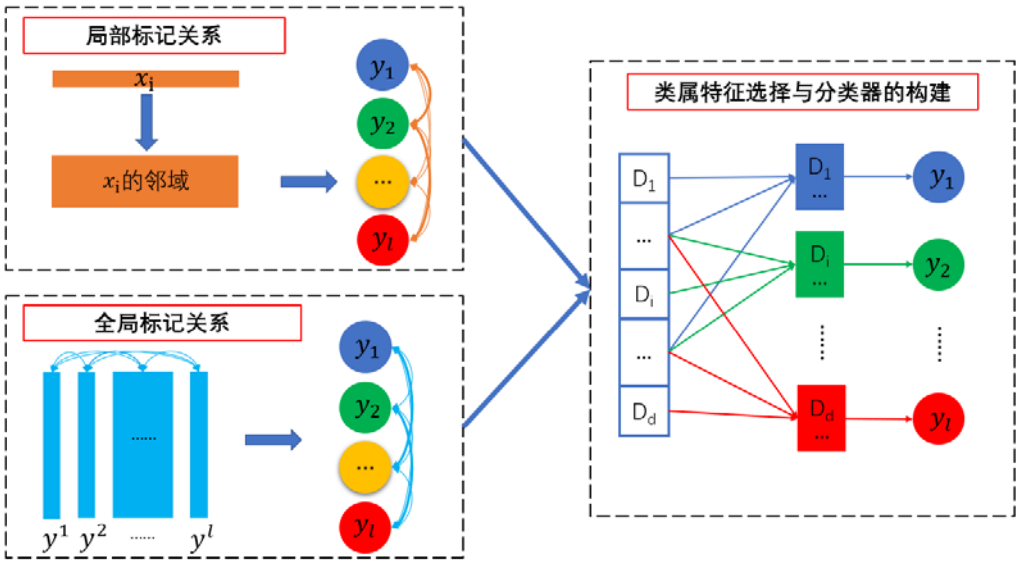


图 1 模型框架图

四. 主要研究成果及应用情况

在《Applied Intelligence》（SCI 3 区）上发表论文 1 篇 Learning label-specific features with global and local label correlation for multi-label classification（和导师共同第一作者）。

联系人：魏博文， 电 话：18031138528 Email:824042710@qq.com