

# 偏多标记分类学习

王小碧

南昌航空大学数学与信息科学学院, 江西 南昌

收稿日期: 2024年3月19日; 录用日期: 2024年4月17日; 发布日期: 2024年4月24日

## 摘要

在偏多标记学习中, 以往研究者常将偏多标记分类和标记消歧过程独立进行, 并且未考虑标记相关性的先验知识难以事先获得。为此, 本文提出了偏多标记分类和标记消歧联合学习的算法(JPMLLD)。JPMLLD建立一个从样本特征到标记的映射; 其次通过将观测标记矩阵分解的方式进行消歧; 最后建立了一个偏多标记分类和标记消歧统一的联合学习框架。大量实验结果验证了本文提出的JPMLLD算法的有效性。

## 关键词

偏多标记学习, 标记相关性, 特征选择, 低秩表示, 稀疏表示

# Partial Multi-Label Classification Learning

Xiaobi Wang

School of Mathematics and Information Sciences, Nanchang Hangkong University, Nanchang Jiangxi

Received: Mar. 19<sup>th</sup>, 2024; accepted: Apr. 17<sup>th</sup>, 2024; published: Apr. 24<sup>th</sup>, 2024

## Abstract

In partial multi-label learning, researchers used to separate partial multi-label learning classification and label disambiguation processes, and prior knowledge without considering label correlation is difficult to obtain in advance. In this paper, we propose a joint learning method of partial multiple label classification and label disambiguation (JPMLLD). First, a mapping from sample features to labels is established. Secondly, the disambiguation is carried out by the decomposition of observation label matrix. Finally, a unified learning framework for partial multiple label classification and label disambiguation is established. The experimental results justify the effectiveness of the proposed JPMLLD algorithm.

文章引用: 王小碧. 偏多标记分类学习[J]. 应用数学进展, 2024, 13(4): 1425-1432.

DOI: 10.12677/aam.2024.134133

## Keywords

Partial Multi-Label Learning, Label Correlations, Feature Selection, Low-Rank Representation, Sparse Representation

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

现代社会中会产生大量的数据，同类数据之间会存在相似的特征，而现代人工智能具有强大计算和学习能力，利用机器可以将这些特征学习并且将学习的内容迁移到其他数据，帮助人类标记信息。多标记学习最初主要用于文档分类[1]，随后推广到各种各样的领域中，例如，由于蛋白质[2]、基因等生物学[3]领域的信息存在多义性，而机器学习非常适用于这种多义性信息的分析[4]。所以机器学习慢慢在生物学中得到应用。但是，现实生活中充斥大量信息，仅靠人工标注是非常费时费力的，大部分信息经人工标注后，训练者得到的是一个候选标记集，该候选标记集包括相关标记以及少量的不相关的噪声标记。这对以往的多标记学习提出了挑战，学者需要对这类问题进行研究，由此偏多标记学习(Partial Multi-label Learning, PML) [5] [6]应运而生。PML 主要是研究如何利用候选标记集和样本特征进行学习。

本文提出了基于矩阵分解的偏多标记学习方法。首先，建立一个从样本特征到标记的分类器，使每个标记不仅依赖样本特征也依赖其他标记；其次通过利用标记矩阵分解的方式进行消歧，最后建立了统一分类和消歧的学习框架。

本文的贡献在于：构建了一个偏多标记分类和标记消歧在统一框架下学习的偏多标记学习算法，丰富了偏多标记学习问题的研究。

## 2. 相关工作

如何消除噪声标记的影响是偏多标记的一个关键问题，研究人员将消除标记歧义的过程定义为消歧。已有许多学者研究出处理偏标记数据的消歧方法。有一部分学者[7]在处理候选标记时为每一个候选标记都赋予均等概率，采用监督学习的方式进行偏多标记学习。这种方式没有考虑候选标记中每个标记的差异性，消歧的效果容易受到假阳性标记干扰，真实标记被假阳性标记淹没。为了减轻平均消歧策略中的假阳性标记干扰的问题，有学者也提出使用加权平均[8]作为最终输出，研究者们提出了 PL-KNN [9]方法，根据距离远近对  $K$  近邻进行不同的权重赋值并按数值进行排序，排序最大作为真实标记。除此之外，有学者[10]提出通过最大似然估计进行学习，第一步消除候选标记集中的假阳性标签，第二步通过后验概率最大化得到真实标记。

## 3. 偏多标记分类学习

$X = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^{n \times d}$ ，这里的  $d$  是指特征的维度， $n$  是指有  $n$  个训练示例，将  $Y = [y_1, y_2, \dots, y_k] \in \{0, 1\}^{n \times k}$  代表每个示例对应的标签，其中  $k$  是标签个数，矩阵里面的取值都是  $\{0, 1\}$ ，比如，如果示例  $i$  具有标签  $j$ ，那么  $y_{ij} = 1$ ，否则， $y_{ij} = 0$ ，则目标函数可以表示为：

$$\min \|T - (XW + 1_n b)S\|_F^2 + \lambda_1 \|W\|_{21} + \lambda_2 \|S\|_F^2 + \lambda_3 \text{rank}(T) + \lambda_4 \|E\|_0$$

为方便求解，我们将求  $T$  稀疏的秩转换为求  $T$  的核范数，将求  $E$  的 0-范数转换成其相似的 1-范数进行求解。

$$\begin{aligned} & \min \|T - (XW + 1b)S\|_F^2 + \lambda_1 \|W\|_{21} + \lambda_2 \|S\|_F^2 + \lambda_3 \|T\|_* + \lambda_4 \|E\|_1 \\ & \text{s.t. } Y = T + E \\ & \quad S_{qq} = 1 \end{aligned}$$

根据增广拉格朗日方法可变成：

$$\begin{aligned} & \min \|T - (XW + 1_n b)S\|_F^2 + \lambda_1 \|W\|_{21} + \lambda_2 \|S\|_F^2 + \lambda_3 \|T\|_* + \lambda_4 \|E\|_1 + \frac{\mu_1}{2} \left\| Y - T - E + \frac{Y_1}{\mu_1} \right\|_F^2 \\ & \text{s.t. } S_{qq} = 1 \end{aligned}$$

由于 1 范数非凸，21 范数是不可微，为求解式，为了便于表示，我们使用  $\theta$  来表示通过合并所有参数而形成的参数列向量，可将式子转换为两部分进行求解：

$$\begin{aligned} f(\theta) &= \sum_{k=1}^Q \sum_{i=1}^N \|x^i W_{.k} S_{.k} + 1_n b_k S_{.k} - T_{.k}\|_F^2 + \lambda_2 \sum_{k=1}^Q \|S_{.k}\|_F^2 + \frac{\mu_1}{2} \sum_{k=1}^Q \left\| Y_{.k} - T_{.k} - E_{.k} + \frac{Y_1}{\mu_1} \right\|_F^2 \\ \min_{\theta} F(\theta) &= f(\theta) + g(\theta) \end{aligned}$$

其中  $g(\theta) = \lambda_1 \|W\|_{21} + \lambda_3 \|T\|_* + \lambda_4 \|E\|_1$ ，目标函数  $F(\theta)$  是两个函数的和，其中第一个函数  $f(\theta)$  是凸光滑函数，第二个函数是 1-范数和 2-范数的和 1 范数是凸的，但不是光滑的。为了发展一种近似优化方法，我们考虑了目标函数  $F(\theta)$  的二阶近似。对于任何  $L > 0$ ，在一个给定的点  $\theta^{(t)}$ ，我们定义：

$$Q_L(\theta, \theta^{(t)}) = f(\theta^{(t)}) + (\theta - \theta^{(t)})^T \nabla f(\theta^{(t)}) + \frac{L}{2} \|\theta - \theta^{(t)}\|^2 + g(\theta)$$

这里  $\nabla f(\theta^{(t)})$  是  $f(\bullet)$  在  $\theta$  在  $\theta^{(t)}$  的梯度函数。

设  $L(f)$  表示  $\nabla f(\theta)$  的 Lipschitz 常数。然后对于任何  $L \geq L(f)$ ，有  $Q_L(\theta, \theta') \geq F(\theta)$ ，基于 Lipschitz 常数的定义和  $f(\theta)$  的凸性和光滑性 [11] [12]。现在，我们假设  $L \geq L(f)$ 。根据 [13] 方法，我们有  $Q_L(\theta, \theta') \geq F(\theta)$ 。通过最小化  $Q_L(\theta, \theta^{(t)})$ ，可以得到一个解决方法：

$$P_L(\theta^{(t)}) = \arg \min_{\theta} Q_L(\theta, \theta^{(t)}) = \arg \min_{\theta} \left\{ \frac{L}{2} \|\theta - \hat{\theta}\|^2 + g(\theta) \right\}$$

此时的  $\hat{\theta} = \left( \theta^{(t)} - \frac{1}{L} \nabla f(\theta^{(t)}) \right)$ ，上述式子最小化可以等价地分解为一组独立的最小化问题，例如：

$$P_L(W^{(t)}) = \arg \min_W \left\{ \frac{L}{2} \|W - \hat{W}\|^2 + \lambda_1 \|W\|_{21} \right\}$$

$$P_L(E_k^{(t)}) = \arg \min_E \left\{ \frac{L}{2} \|e - \hat{E}_{.k}\|^2 + \lambda_2 \|E\|_1^2 \right\}$$

它们有封闭形式的解。其中稀疏范数凸优化问题的闭型解可以分别应用群软阈值和直接软阈值算子计算 [14]，可得最终  $W, E$  更新公式：

$$P_L(W^{(t)}) = \hat{W}_{.j} \left( 1 - \frac{\lambda_1}{L \|\hat{W}_{.j}\|} \right)_+, \forall j;$$

$$P_L(E_{:k}^{(l)}) = \text{sign}(\hat{E}_{:k}) \circ \left( \left| \hat{E}_{:k} \right| - \frac{\lambda_4}{L} \right)_+$$

其中 $\circ$ 表示元素向量乘积, 算子 $(\cdot)_+ = \max(\cdot, 0)$ 。

更新  $S$ , 固定  $W, b, T, E$

$$\min_S \|T - (XW + 1_n b)S\|_F^2 + \lambda_2 \|S\|_F^2$$

求导可得:

$$S = M^{-1} (XW + 1_n b)^T T$$

$$M = (XW + 1_n b)^T (XW + 1_n b) + \lambda_2$$

更新  $b$ , 固定  $W, S, T, E$

求导可得:

$$b = 1_n^T T S^{-1} - 1_n^T XW$$

更新  $T$  固定  $W, S, E, b$

$$\min_T \frac{L}{2} \|T - \tilde{T}_t^s\|_F^2 + \lambda_3 \|T\|_*$$

是一个非光滑凸问题, 可以得到最优解。该定理[15]总结如下:

设  $\tilde{T}_t^s = U\Sigma V^T$  为  $\tilde{T}_t^s$  的奇异值分解(SVD), 其中,  $U \in \mathbb{R}^{d \times r}$  以及  $v \in \mathbb{R}^{Q \times r}$  为列正交矩阵,  $\Sigma \in \mathbb{R}^{r \times r}$  为对角矩阵, 其中,  $r = \text{rank}(\tilde{T}_t^s)$ 。则通过求解以下方程, 可以得到  $T$  的最优解:

$$T = U\tilde{\Sigma}V^T$$

$\tilde{\Sigma} \in \mathbb{R}^{r \times r}$  为对角线矩阵和  $\tilde{\Sigma}_{ii} = \max\left(0, \Sigma_{ii} - \frac{\lambda_3}{L}\right)$ 。

## 4. 实验

### 4.1. 实验设置

实验共使用了九个网络公开数据集(六个 ML 数据集和两个 PML 数据)和六个对比算法。通过添加随机噪声方式, 将每个多标记数据集生成多个带有候选标记的 PML 数据集[16]。具体地, 在给定的多标记实例的标记集中添加不相关的标记生成合成 PML 数据集, 通过改变候选标记的平均数量(#CLs)生成了 21 个合成 PML 数据集, 表 1 总结了数据集的特征, 包括特征数量, 平均标记个数(#GLs)、平均候选标记个数(#CLs)和应用领域,  $r = \{3, 4, 5\}$  表示带有假阳性候选标记的个数。其中音乐情感、音乐风格为两个真实的 PML 数据集。

**Table 1.** Description of the multi-labeled dataset

**表 1.** 多标记数据集的描述

Dataset	#Examples	#Features	#Labels	avg.#CLs	avg.#GLs
Music_style	6839	98	10	6.04	1.44
Music_emotion	6833	98	11	5.29	2.42
Emotions	593	72	6	$r = 3, 4, 5$	1.86
arts	5000	462	26	$r = 3, 4, 5$	1.64

续表

social	5000	1047	39	$r = 3, 4, 5$	1.28
slashdot	3782	1079	22	$r = 3, 4, 5$	1.18
Image	2000	294	5	$r = 3, 4, 5$	1.23
yeast	2417	103	14	$r = 3, 4, 5$	4.23
business	5000	438	30	$r = 3, 4, 5$	1.59

本文对比了三种偏多标记算法和三种经典的多标记算法证明 JPMLLD 的有效性。

1) PML-LRS [17]一种利用低秩矩阵分解将标记矩阵分解成真实标记和无关矩阵的偏多标记学习方法, 该方法可以很好的去除噪声标记解决标记信息冗余等问题。

2) LIFT [18]一种通过对每个正负标记进行聚类获取标记特征来训练分类器的偏多标记学习方法, 该方法的特点是使用标记的特征来挖掘标记信息。

3) ML-KNN [19]一种识别训练集  $k$  近邻获得每个可能类的相邻实例信息, 通过最大后验来确定分类的懒惰学习算法, 该方法在多标记学习中常作为对比算法。

4) PRATICLE [20]是一种基于标记传播技术和  $k$  近邻信息来估计置信度矩阵的偏多标记学习方法, 该方法通过阈值将置信度矩阵转化成标记矩阵, 并基于成对排序模型在估计的标记矩阵上训练一个多标记分类器。

5) TRAM [21]利用已知标记和未知标记获得有效模型的转换的一种多标记分类方法并提出了算法的封闭解, 是一种可以通过标记集传播的多标记学习算法。

6) PLDA [22]是一种研究偏多标记学习识别增强的算法, 该方法提出了一种联合优化和估计部分标记置信度优化原型, 同时保证特征全局和标记局部一致性。

## 4.2. 实验结果与分析

在进行对比实验时, 本文使用 60%的数据集进行训练, 40%数据集用于测试, 并采用随机 10 折交叉验证的方式进行训练。为了评估偏多标记分类器的性能, 采用一种常见的评价准则[16] [23]进行评估比较, 即 Hamming Loss [24]; 对于 Hamming Loss 评估准则, 值越小越优。表 2 详细记录了本文方法和其他六种方法在重复 10 次实验后得到的平均值和标准差。

**Table 2.** Performance comparison based on Hamming Loss (smaller value is better) on different labeled datasets

**表 2.** 在不同标记的数据集上, 基于 Hamming Loss (值越小越好)的性能比较

Datasets		JPMLLD	PML-LRS	LIFT	ML-kNN	PRATICLE	TRAM	PLDA
music_style	6.04	0.329 ± 0.025	0.804 ± 0.004●	0.580 ± 0.033●	0.848 ± 0.006●	0.841 ± 0.002●	0.877 ± 0.002●	0.722 ± 0.108●
music_emotion	5.29	0.302 ± 0.007	0.328 ± 0.003	0.336 ± 0.009	0.387 ± 0.017●	0.761 ± 0.003●	0.838 ± 0.003●	0.654 ± 0.050●
Emotions	$r = 3$	0.191 ± 0.015	0.252 ± 0.009●	0.553 ± 0.026●	0.212 ± 0.031	0.055 ± 0.013○	0.246 ± 0.026●	0.912 ± 0.014●
	$r = 4$	0.259 ± 0.040	0.281 ± 0.020	0.636 ± 0.018●	0.655 ± 0.021●	0.926 ± 0.008●	0.422 ± 0.023●	0.928 ± 0.015●
	$r = 5$	0.270 ± 0.015	0.303 ± 0.000	0.686 ± 0.016●	0.688 ± 0.017●	0.924 ± 0.008●	0.527 ± 0.016●	0.939 ± 0.011●

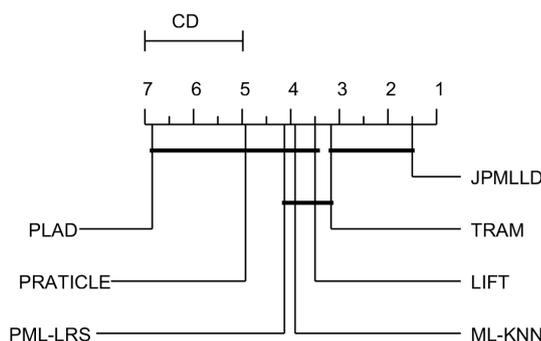
续表

Art	$r = 3$	0.054 ± 0.002	0.060 ± 0.001	0.054 ± 0.002 <sup>○</sup>	0.061 ± 0.002	0.989 ± 0.001●	0.071 ± 0.002	0.994 ± 0.001●
	$r = 4$	0.060 ± 0.000	0.063 ± 0.000	0.060 ± 0.001 <sup>○</sup>	0.061 ± 0.002	0.989 ± 0.001●	0.163 ± 0.002●	0.994 ± 0.001●
	$r = 5$	0.062 ± 0.000	0.064 ± 0.000	0.808 ± 0.018●	0.062 ± 0.001 <sup>○</sup>	0.989 ± 0.001●	0.197 ± 0.002●	0.995 ± 0.001●
Social	$r = 3$	0.020 ± 0.001	0.060 ± 0.000●	0.021 ± 0.001	0.023 ± 0.002	0.575 ± 0.020●	0.090 ± 0.001●	0.992 ± 0.002●
	$r = 4$	0.023 ± 0.001	0.038 ± 0.001	0.021 ± 0.001 <sup>○</sup>	0.024 ± 0.001	0.989 ± 0.001●	0.116 ± 0.001●	0.993 ± 0.002●
	$r = 5$	0.032 ± 0.000	0.059 ± 0.000	0.520 ± 0.009●	0.024 ± 0.001	0.575 ± 0.020●	0.140 ± 0.001●	0.993 ± 0.002●
Slashdot	$r = 3$	0.041 ± 0.002	0.043 ± 0.001 <sup>○</sup>	0.040 ± 0.001 <sup>○</sup>	0.052 ± 0.001	0.988 ± 0.002●	0.163 ± 0.002●	0.993 ± 0.002●
	$r = 4$	0.042 ± 0.002	0.041 ± 0.001 <sup>○</sup>	0.041 ± 0.001 <sup>○</sup>	0.054 ± 0.002	0.993 ± 0.002●	0.208 ± 0.003●	0.993 ± 0.002●
	$r = 5$	0.045 ± 0.002	0.043 ± 0.003 <sup>○</sup>	0.041 ± 0.001 <sup>○</sup>	0.055 ± 0.001	0.990 ± 0.003●	0.224 ± 0.008●	0.995 ± 0.003●
Image	$r = 3$	0.171 ± 0.010	0.205 ± 0.004	0.580 ± 0.033●	0.178 ± 0.018	0.779 ± 0.048●	0.227 ± 0.016●	0.914 ± 0.016●
	$r = 4$	0.225 ± 0.005	0.362 ± 0.014●	0.576 ± 0.042●	0.753 ± 0.006●	0.927 ± 0.006●	0.523 ± 0.020●	0.936 ± 0.010●
	$r = 5$	0.225 ± 0.005	0.599 ± 0.022●	0.752 ± 0.004●	0.753 ± 0.006●	0.903 ± 0.062●	0.608 ± 0.018●	0.959 ± 0.006●
Yeast	$r = 3$	0.207 ± 0.002	0.207 ± 0.003	0.248 ± 0.008	0.192 ± 0.005 <sup>○</sup>	0.946 ± 0.002●	0.327 ± 0.011●	0.953 ± 0.006●
	$r = 4$	0.212 ± 0.004	0.217 ± 0.011	0.316 ± 0.014●	0.241 ± 0.006●	0.941 ± 0.016●	0.270 ± 0.014●	0.954 ± 0.007●
	$r = 5$	0.217 ± 0.003	0.267 ± 0.044●	0.571 ± 0.013●	0.327 ± 0.015●	0.893 ± 0.004●	0.366 ± 0.008●	0.956 ± 0.006●
Business	$r = 3$	0.025 ± 0.009	0.051 ± 0.001●	0.025 ± 0.002 <sup>○</sup>	0.027 ± 0.002	0.971 ± 0.000●	0.114 ± 0.001●	0.971 ± 0.001●
	$r = 4$	0.026 ± 0.000	0.052 ± 0.001●	0.026 ± 0.002 <sup>○</sup>	0.028 ± 0.001	0.971 ± 0.000●	0.144 ± 0.002●	0.972 ± 0.002●
	$r = 5$	0.027 ± 0.000	0.054 ± 0.002●	0.027 ± 0.000	0.028 ± 0.001	0.971 ± 0.000●	0.176 ± 0.002●	0.972 ± 0.002●

从表 2 中数据可以看出,所有方法随着  $r$  的增加性能基本都会降低。在不同的  $r$  值情况下, JPMLLD 在多数情况下优于其他算法。从表 1 的实验中可以看出,与其他方法相比, JPMLLD 算法在 12 个数据集上的 Hamming Loss 值都大于其它算法,且在绝大多数数据集上相对稳定,对于 Business 数据集, JPMLLD 的 Hamming Loss 值最大,对于 Arts 数据集, JPMLLD 取得的平均结果最优。为了验证本文提出的算法是否与其它算法存在差异,在 5% 显著性水平的 t 检验上进行两两算法效果比较,并将结果表示在表 2 中。在表 1 中●表示本文的方法在统计 t 检验上优于比较方法,○则表示比较方法优于本文的方法。

最后,采用显著性为 5%的 Nemenyitest [25]进一步分析算法之间的差异,如两个算法之间差异显著,她们的实验结果排序值就不低于临界差(critical difference, 简称 CD),如图 1 所示,在每个子图 CD 在轴上方,排名越高效果越好的方法就越排在左边,没有显著差异的方法组用粗体实线连接。在图 1 中,JPMLLD 与 TRAM 方法用 CD 线连接,说明基于 Hamming Loss 评价准则时,JPMLLD 与 TRAM 方法性能相当,并且优于其他算法。

在所有 21 个实验中(包括评价指标以及不同数据集),JPMLLD 在 65%的实例中排名第一。值得注意的是,JPMLLD 在所有评价中都达到了最优,这说明本文的方法拥有更好的性能。



**Figure 1.** Performance comparison on classifier evaluation criteria

**图 1.** 分类器评价准则上的性能比较

## 5. 结论

本文提出了一种基于矩阵分解偏多标记分类方法,通过低秩稀疏矩阵分解将标记矩阵分解成精确的标记矩阵和噪声标记矩阵的方式进行消歧;将问题可转化成一个凸优化问题,通过交替求解的方式进行求解,在 9 个数据集上基于 Hamming Loss 评价准则上的实验结果表明,JPMLLD 优于其他 6 种算法。

## 参考文献

- [1] 耿新, 徐宁. 标记分布学习与标记增强[J]. 中国科学: 信息科学, 2018, 48(5): 521-530.
- [2] 季小皖. 弱监督不平衡多标记分类方法研究[D]: [硕士学位论文]. 舟山: 浙江海洋大学, 2023.
- [3] 李鸿燕. 双因子张量范数正则化低秩张量填充[D]: [硕士学位论文]. 大连: 辽宁师范大学, 2023.
- [4] 薛鹏. 人体运动数据的特征表示与合成研究[D]: [硕士学位论文]. 南京: 南京信息工程大学, 2022.
- [5] 郝秀雁. 特征与标记同时缺失的多标记学习方法研究[D]: [硕士学位论文]. 马鞍山: 安徽工业大学, 2022.
- [6] 熊胤崧. 基于标记增强的特征选择及其在叶片种类识别中的应用研究[D]: [硕士学位论文]. 南昌: 江西农业大学, 2022.
- [7] 吕庚育. 弱监督多标记学习算法关键技术研究[D]: [硕士学位论文]. 北京: 北京交通大学, 2022.
- [8] 余婷婷. 噪声环境下的多标记学习方法研究与应用[D]: [硕士学位论文]. 重庆: 西南大学, 2021.
- [9] 孙利娟. 偏多标记学习算法关键技术研究[D]: [博士学位论文]. 北京: 北京交通大学, 2021.
- [10] 孙悦. 基于消歧策略的偏标记学习算法研究[D]: [硕士学位论文]. 北京: 北京交通大学, 2021.
- [11] 朱雯芳. 基于样本相关性的多标记学习算法及应用研究[D]: [硕士学位论文]. 南京: 南京理工大学, 2021.
- [12] Ming, K.X. and Sheng, J.H. (2021) Partial Multi-Label Learning with Noisy Label Identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **44**, 3676-3687.
- [13] 谢明昆. 偏多标记学习研究[D]: [硕士学位论文]. 南京: 南京航空航天大学, 2021.

- [14] 宋帆. 基于专家的特征选择及缺省多标记学习策略[D]: [硕士学位论文]. 安庆: 安庆师范大学, 2020.
- [15] 李勇. 面向多标签决策系统的特征优化方法研究[D]: [硕士学位论文]. 北京: 中国石油大学(北京), 2020.
- [16] 杭仁龙. 遥感数据分析中的特征表示方法研究[D]: [硕士学位论文]. 南京: 南京信息工程大学, 2017.
- [17] 孙宁朝. 标注受限场景下的多标签学习方法研究[D]: [硕士学位论文]. 北京: 国防科技大学, 2020.
- [18] 张博文. 基于自适应近邻图模型与局部判别分析的特征选择算法研究[D]: [硕士学位论文]. 合肥: 安徽大学, 2020.
- [19] 王一宾, 吴陈, 程玉胜, 等. 不平衡标记差异性多标记特征选择算法[J]. 深圳大学学报(理工版), 2020, 37(3): 234-242.
- [20] 王雨思, 路德杨, 李海洋. 基于分式函数约束的稀疏子空间聚类方法[J]. 计算机工程与应用, 2020, 56(7): 39-47.
- [21] 王礼琴. 半监督多标记特征选择算法研究[D]: [硕士学位论文]. 长沙: 湖南师范大学, 2019.
- [22] 王恒远. 含噪数据的稀疏子空间聚类算法及理论研究[D]: [硕士学位论文]. 长沙: 西安工程大学, 2018.
- [23] 陈晓云, 廖梦真. 基于稀疏和近邻保持的极限学习机降维[J]. 自动化学报, 2019, 45(2): 325-333.
- [24] 曾玉华. 稀疏数据恢复的结构优化模型及其算法研究[D]: [硕士学位论文]. 长沙: 湖南大学, 2016.
- [25] 何志芬. 结合标记相关性的多标记分类算法研究[D]: [博士学位论文]. 南京: 南京师范大学, 2015.