



结合局部标记序关系的弱监督标记分布学习

秦天, 滕齐发, 贾修一

引用本文:

秦天,滕齐发,贾修一. 结合局部标记序关系的弱监督标记分布学习[J]. 智能系统学报, 2023, 18(1): 47–55.

QIN Tian,TENG Qifa,JIA Xiuyi. Weakly supervised label distribution learning by maintaining local label ranking[J]. *CAAI Transactions on Intelligent Systems*, 2023, 18(1): 47–55.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202204018>

您可能感兴趣的其他文章

自适应标记关联与实例关联诱导的缺失多视图弱标记学习

Adaptive label correlation and instance correlation guided incomplete multiview weak label learning

智能系统学报. 2022, 17(4): 670–679 <https://dx.doi.org/10.11992/tis.202106017>

结合谱聚类的标记分布学习

Label distribution learning based on spectral clustering

智能系统学报. 2019, 14(5): 966–973 <https://dx.doi.org/10.11992/tis.201809019>

代价敏感数据的多标记特征选择算法

Multi-label feature selection algorithm for cost-sensitive data

智能系统学报. 2019, 14(5): 929–938 <https://dx.doi.org/10.11992/tis.201807027>

应用k-means算法实现标记分布学习

Label distribution learning based on k-means algorithm

智能系统学报. 2017, 12(3): 325–332 <https://dx.doi.org/10.11992/tis.201704024>

基于分布先验的半监督FCM的肺结节分类

Classification of pulmonary nodules by semi-supervised FCM based on prior distribution

智能系统学报. 2017, 12(5): 729–734 <https://dx.doi.org/10.11992/tis.201706018>

DOI: 10.11992/tis.202204018

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20221013.1445.006.html>

结合局部标记序关系的弱监督标记分布学习

秦天, 滕齐发, 贾修一

(南京理工大学 计算机科学与工程学院, 江苏 南京 210094)

摘要: 标记分布学习 (label distribution learning, LDL) 是一种用于解决标记多义性的新颖学习范式。现有的 LDL 方法大多基于完整数据信息进行设计, 然而由于高昂的标注成本以及标注人员水平的局限性, 很难获取到完整标注数据信息, 且会导致传统 LDL 算法性能的下降。为此, 本文提出了一种新型的结合局部序标记关系的弱监督标记分布学习算法, 通过维持尚未缺失标记之间的相对关系, 并利用标记相关性来恢复缺失的标记, 在数据标注不完整的情况下提升算法性能。在 14 个数据集上进行了大量的实验来验证算法的有效性。

关键词: 标记分布学习; 标记多义性; 弱监督学习; 标记排序; 弱监督标记分布学习; 多标记学习; 标记相关性; 局部标记序关系

中图分类号: TP181 文献标志码: A 文章编号: 1673-4785(2023)01-0047-09

中文引用格式: 秦天, 滕齐发, 贾修一. 结合局部标记序关系的弱监督标记分布学习 [J]. 智能系统学报, 2023, 18(1): 47-55.

英文引用格式: QIN Tian, TENG Qifa, JIA Xiuyi. Weakly supervised label distribution learning by maintaining local label ranking[J]. CAAI transactions on intelligent systems, 2023, 18(1): 47-55.

Weakly supervised label distribution learning by maintaining local label ranking

QIN Tian, TENG Qifa, JIA Xiuyi

(School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: Label distribution learning (LDL) is a novel learning paradigm for solving labeling polysemy. Most existing LDL methods are designed based on complete data information; however, because of high labeling costs and the limitation of labelers' level, complete labeling data information is difficult to obtain, which leads to performance degradation in traditional LDL algorithms. In this paper, we propose a novel weakly supervised LDL by maintaining a local label ranking (WSLDL-MLLR) algorithm. We improve algorithm performance under incomplete data labeling by maintaining relative relationships between the not-yet-missing labels and using label correlation to recover missing labels. Extensive experiments conducted on 14 datasets verified the effectiveness of the algorithm.

Keywords: label distribution learning; label polysemy; weakly supervised learning; label ranking; weakly supervised label distribution learning; multi-label learning; label correlation; local label ranking relation

标记分布学习^[1]是一种泛化能力和表征能力更强的多义机器学习范式。相较于多标记学习, 标记分布学习对训练数据集的标注更为严格, 即对标记空间中的每个标记赋予具体的描述度数值而非简单的 0/1 标注, 这就导致标记分布数据集的标注需要花费更多的时间和人力成本。并且, 由于实际应用中标注对象更为复杂, 标记分布数据集不可避免地会存在部分缺失。如果直接用现

有的基于完整监督信息设计的标记分布学习算法^[2-4]来处理存在缺失的标记分布数据集, 会出现两种情况: 部分算法无法直接处理此任务; 能处理此任务的算法取得的性能达不到理想效果。于是, Miao 等^[5]将此问题定性为弱监督标记分布学习 (weakly supervised label distribution learning, WSLDL), 并设计专用的算法来更好地解决此类问题。

在弱监督标记分布学习任务中, 由于标记分布空间存在一定程度的缺失, 通常需要通过先验假设从弱监督数据集中挖掘出数据固有的结构信

收稿日期: 2022-04-15. 网络出版日期: 2022-10-14.

基金项目: 国家自然科学基金项目 (62176123); 江苏省自然科学基金项目 (BK20191287).

通信作者: 贾修一. E-mail: jiaxy@njjust.edu.cn.

息,进而通过利用样本相关性和标记相关性来提高算法性能。Miao 等^[5]利用了一种基于全局低秩的标记相关性来提高弱监督标记分布算法的预测性能。Jia 等^[6]同时考虑样本相关性和标记相关性,提出了一种基于矩阵补全的直推式弱监督标记分布学习算法。Zeng 等^[7-8]和 Xu 等^[9]也在他们设计的弱监督标记分布学习方法中应用了样本相关性。

通过研究现有的弱监督标记分布学习工作^[10-11],我们发现已有工作在拟合未缺失标记分布时,都忽略了标记分布值之间隐含的相对序关系,即标记描述度大的标记对应的排序值也更大。这一语义信息对学习未缺失标记和恢复缺失标记有着重要作用。

为了将标记序关系引入到弱监督标记分布学习中,本文设计了一个新颖的损失函数来显式地刻画未缺失标记分布间的序关系。同时,考虑到利用标记相关性的必要性,我们通过更新标记间的相关性来引导模型训练,这一过程与常用的低秩近似手段有明显的区别。

本文的主要贡献总结如下:1)首次在弱监督标记分布学习中引入局部标记序关系,并引入专门用于度量标记之间序关系的评价指标;2)为了打破弱监督标记分布学习中标记相关性为低秩特性这一先验假设,我们在模型训练的过程中不断地主动学习并更新标记相关矩阵;3)提出了一个新颖的结合局部标记序关系的弱监督标记分布学

习算法。

1 相关知识

在标记分布学习框架中, $\mathbf{X} = [x_1 \ x_2 \ \cdots \ x_n] \in \mathbf{R}^{n \times d}$ 表示特征空间,其中 d 是特征维度, n 为样本数目, x_i 表示第 i 个样本。假设, $\mathbf{D} = [\mathbf{D}_1 \ \mathbf{D}_2 \ \cdots \ \mathbf{D}_n] \in \mathbf{R}^{n \times l}$ 为标记空间,其中 $\mathbf{D}_i = [d_{x_i}^{y_1} \ d_{x_i}^{y_2} \ \cdots \ d_{x_i}^{y_l}]$ 表示样本 x_i 的标记分布矩阵, $d_{x_i}^{y_j}$ 表示标记 y_j 对实例 x_i 的描述程度, l 表示标记的数目,此外鉴于标记分布本身的特点, $d_{x_i}^{y_j}$ 应满足两个基本特性,即 $d_{x_i}^{y_j} \in [0, 1]$ 且 $\sum_{j=1}^l d_{x_i}^{y_j} = 1$ 。

在处理弱监督标记分布学习任务时,为了生成弱监督的数据集信息,我们假设标记分布矩阵 \mathbf{D} 中的元素值是均匀随机缺失的。我们从 \mathbf{D} 中随机选择一些已知值和缺失值来创建 $\tilde{\mathbf{D}}$ 。我们保持弱监督标记分布矩阵 $\tilde{\mathbf{D}}$ 的维度大小与 \mathbf{D} 相同,并将缺失值设为 0。特别地,令 $\mathbf{\Omega} \in \mathbf{R}^{n \times l}$ 为索引矩阵, $\mathbf{\Omega}$ 的取值为 $\tilde{\mathbf{D}}$ 中均匀随机选择的已知值的位置索引,其具体定义如下:

$$\Omega_{i,j} = \begin{cases} 0, & \tilde{D}_{i,j} \text{ 缺失} \\ 1, & \text{其他} \end{cases}$$

2 算法构造

2.1 算法框架

本小节介绍的结合局部标记序关系的弱监督标记分布学习算法的整体框架如图 1 所示。

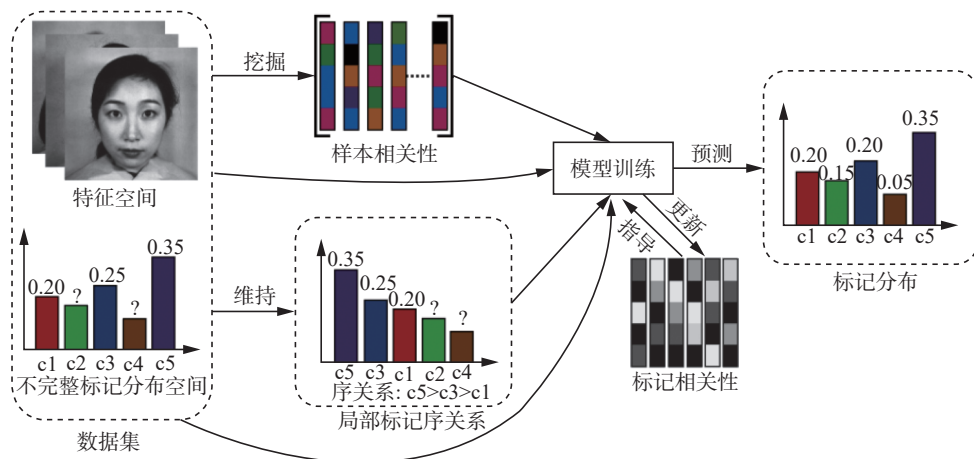


图 1 结合局部标记序关系的弱监督标记分布学习框架

Fig. 1 Framework of weakly supervised distribution learning by maintaining local label ranking

为了简化模型设计,我们选择最小二乘^[12-13]损失作为损失函数,具体计算方式为

$$\min_w \frac{1}{2} \left\| \mathbf{\Omega} \odot (\mathbf{XW} - \tilde{\mathbf{D}}) \right\|_F^2$$

式中: $\mathbf{\Omega}$ 是 $\tilde{\mathbf{D}}$ 中尚未缺失值所对应的索引矩阵;

$\tilde{\mathbf{D}}$ 表示弱监督标记分布矩阵; \odot 为哈达玛积; $\|\cdot\|_F$ 表示矩阵的 F 范数。 \odot 运算规则为: 设 $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbf{R}^{n \times d}$, 令 $\mathbf{C} = \mathbf{A} \odot \mathbf{B}$, 则 $C_{ij} = A_{ij} \times B_{ij}$ 。

尽管弱监督标记分布数据集存在不同程度的缺失,但其数据集本身所固有的标记相关性

对提升模型性能有一定的辅助作用。现有的弱监督标记分布学习算法大部分是基于低秩近似的思想来刻画标记相关性, 这种方法不可避免地会导致模型预测的标记分布值趋于平均, 这就破坏了标记间原本的相对序关系。为了缓解这个问题, 受到文献 [14] 和 [15] 的启发, 我们通过主动学习标记间相关性的方式, 来增强弱监督标记分布矩阵 $\tilde{\mathbf{D}}$, 并在模型的训练损失中加入标记相关性损失项:

$$\min_{\mathbf{W}, \mathbf{C}} \frac{1}{2} \left\| \hat{\mathbf{Q}} \odot (\mathbf{XW} - \tilde{\mathbf{D}}\mathbf{C}) \right\|_F^2$$

式中: $\hat{\mathbf{Q}} = \mathbf{E} - \mathbf{Q}$, $\mathbf{E} = \mathbf{1}_{n \times l}$; $\mathbf{C} \in \mathbf{R}^{l \times l}$ 为标记相关性矩阵。具体地说, 标记相关性矩阵是对称阵, 每个元素反映两个标记之间的相关程度, 标记正相关性越高, 对应的输出就越相近, 反之亦然。现有的弱监督标记分布学习算法只关注未缺失标记分布 \mathbf{Q} 的回归损失, 而忽视了对缺失标记分布 $\hat{\mathbf{Q}}$ 预测值的学习更新, 该损失项的加入可以很好地弥补这一缺陷。

为了使模型预测的缺失标记分布值也满足这种相关性, 在模型中加入损失项来实现这种约束, 即 $\min_{\mathbf{W}, \mathbf{C}} \frac{1}{2} \left\| \mathbf{XWC} - \mathbf{XW} \right\|_F^2$ 。

通过实验发现, 仅用标记相关性来指导模型训练, 取得的效果并不能达到预期。为了充分利用数据本身的结构信息, 又引入样本相关性损失项来进一步提升模型效果: $\min_{\mathbf{W}} \text{tr}((\mathbf{XW})^T \mathbf{L}(\mathbf{XW}))$, 这可以将样本相关性迁移到标记空间中。此处结合了流形正则化的思想, 即对于样本相似的数据, 其标记也具有相似性。我们利用 $n \times n$ 维 Laplacian 矩阵 $\mathbf{L} = \text{Diag}(\mathbf{S} \times \mathbf{1}_n) - \mathbf{S}$ 来表示约束, 其中 $\mathbf{1}_n$ 为长度为 n 的所有元素值为 1 的向量 Diag 表示对角矩阵构造, \mathbf{S} 具体定义为

$$S_{ij} = \begin{cases} e^{-\frac{\|X_i - X_j\|_2^2}{\sigma^2}}, & X_i \in \mathbf{N}_k(X_j) \text{ 或 } X_j \in \mathbf{N}_k(X_i) \\ 0, & \text{其他} \end{cases}$$

式中: σ 是高斯核的方差, 通常根据样本点之间的平均距离来取值; $\mathbf{N}_k(X_i)$ 是样本点 X_i 的 k 近邻集合。

除此之外, 为了显式地刻画出局部标记序关系, 受文献 [16, 17] 的启发, 我们结合标记分布的特性, 设计了一种新颖的用来度量标记之间相对序关系的损失项:

$$\min_{\mathbf{W}} \frac{1}{c^2} \sum_{i=1}^n \sum_{j,k=1}^l (\tilde{D}_{ij} - \tilde{D}_{ik})^2 \cdot \ln \left(1 + e^{-(X_i \cdot \mathbf{W} \cdot_j - X_i \cdot \mathbf{W} \cdot_k)^2} \right)$$

在预测空间中, 对于不同间距的真实标记对, 其距离越大, 错误预测标记序关系受到的惩罚力度理应越强。因此, 将两个真实标记间的距离平

方作为权重系数, 同时, 为了构造一个凸问题, 引入了 Sigmoid 函数。

为了防止模型过拟合, 同时保证标记分布值满足大于 0 且和为 1 的约束条件, 引入一个额外损失:

$$\min_{\mathbf{W}, \mathbf{C}} \frac{1}{2} \left\| \mathbf{XW} \times \mathbf{1}_l - \mathbf{1}_n \right\|_F^2$$

式中 $\mathbf{1}_l$ 和 $\mathbf{1}_n$ 分别为长度为 l 和 n 的所有元素值为 1 的向量。

最后, 结合本节所有优化项, 得到模型最终的目标函数为

$$\begin{aligned} \min_{\mathbf{W}, \mathbf{C}} \frac{1}{2} \left\| \mathbf{Q} \odot (\mathbf{XW} - \tilde{\mathbf{D}}) \right\|_F^2 + \frac{\lambda_1}{2} \left\| \hat{\mathbf{Q}} \odot (\mathbf{XW} - \tilde{\mathbf{D}}\mathbf{C}) \right\|_F^2 + \\ \frac{\lambda_2}{2} \left\| \mathbf{XWC} - \mathbf{XW} \right\|_F^2 + \lambda_3 \text{tr}((\mathbf{XW})^T \mathbf{L}(\mathbf{XW})) + \\ \frac{\lambda_4}{c^2} \sum_{i=1}^n \sum_{j,k=1}^l (\tilde{D}_{ij} - \tilde{D}_{ik})^2 \cdot \ln \left(1 + e^{-(X_i \cdot \mathbf{W} \cdot_j - X_i \cdot \mathbf{W} \cdot_k)^2} \right) + \\ \frac{\lambda_5}{2} \left\| \mathbf{XW} \times \mathbf{1}_l - \mathbf{1}_n \right\|_F^2 \end{aligned} \quad (1)$$

式中 λ_1 、 λ_2 、 λ_3 、 λ_4 和 λ_5 为每项的权重系数。

2.2 优化求解

式 (1) 中的最小化问题在迭代次数为 t 时, 可以分解为如下几个子问题:

$$\begin{aligned} \mathbf{W}^{t+1} = \arg \min_{\mathbf{W}^t, \mathbf{C}^t} \frac{1}{2} \left\| \mathbf{Q} \odot (\mathbf{XW}^t - \tilde{\mathbf{D}}) \right\|_F^2 + \\ \frac{\lambda_1}{2} \left\| \hat{\mathbf{Q}} \odot (\mathbf{XW}^t - \tilde{\mathbf{D}}\mathbf{C}^t) \right\|_F^2 + \\ \frac{\lambda_2}{2} \left\| \mathbf{XW}^t \mathbf{C}^t - \mathbf{XW}^t \right\|_F^2 + \lambda_3 \text{tr}((\mathbf{XW}^t)^T \mathbf{L}(\mathbf{XW}^t)) + \\ \frac{\lambda_4}{c^2} \sum_{i=1}^n \sum_{j,k=1}^l (\tilde{D}_{ij} - \tilde{D}_{ik})^2 \cdot \ln \left(1 + e^{-(X_i \cdot \mathbf{W}^t \cdot_j - X_i \cdot \mathbf{W}^t \cdot_k)^2} \right) + \\ \frac{\lambda_5}{2} \left\| \mathbf{XW}^t \times \mathbf{1}_l - \mathbf{1}_n \right\|_F^2 \end{aligned} \quad (2)$$

$$\begin{aligned} \mathbf{C}^{t+1} = \arg \min_{\mathbf{W}^{t+1}, \mathbf{C}^t} \frac{\lambda_1}{2} \left\| \hat{\mathbf{Q}} \odot (\mathbf{XW}^{t+1} - \tilde{\mathbf{D}}\mathbf{C}^t) \right\|_F^2 + \\ \frac{\lambda_2}{2} \left\| \mathbf{XW}^{t+1} \mathbf{C}^t - \mathbf{XW}^{t+1} \right\|_F^2 \end{aligned} \quad (3)$$

对于式 (2) 和式 (3), 我们可以通过限制内存的拟牛顿方法 (limited-memory broyden-fletcher-goldfarb-shanno method, L-BFGS) [18] 来优化求解。该优化方法需要得到目标函数的一阶导数, 式 (2) 和式 (3) 的一阶导数分别为

$$\begin{aligned} \nabla \mathbf{W}^{t+1} = \mathbf{X}^T (\mathbf{Q} \odot (\mathbf{XW}^t - \tilde{\mathbf{D}})) + \\ \lambda_1 \mathbf{X}^T (\hat{\mathbf{Q}} \odot (\mathbf{XW}^t - \tilde{\mathbf{D}}\mathbf{C}^t)) + \\ \lambda_2 (\mathbf{X}^T \mathbf{M}^T \mathbf{C}^t - \mathbf{X}^T \mathbf{M}^t) - \\ \frac{2 \times \lambda_3}{c^2} \sum_{i=1}^n \sum_{j,k=1}^l \frac{(\tilde{D}_{ij} - \tilde{D}_{ik})^2 \times \alpha^t \times (X_i^T \cdot (\mathbf{e}_j - \mathbf{e}_k))}{1 + e^{-(\alpha^t)^2}} + \\ \lambda_4 \mathbf{X}^T (\mathbf{L} + \mathbf{L}^T) \mathbf{XW}^t + \lambda_5 \mathbf{X}^T (\mathbf{XW} \times \mathbf{1}_l - \mathbf{1}_n) \times \mathbf{1}_l^T \end{aligned}$$

$$\nabla C^{t+1} = -\lambda_1 \bar{\mathbf{D}}^T (\bar{\mathbf{Q}} \odot (\mathbf{XW}^{t+1} - \mathbf{DC}^t)) + \lambda_2 (\mathbf{XW}^{t+1})^T (\mathbf{XW}^{t+1} \mathbf{C}^t - \mathbf{XW}^{t+1})$$

式中: $\mathbf{M}^t = \mathbf{XW}^t \mathbf{C}^t - \mathbf{XW}^t$; $\alpha^t = \mathbf{X}_i \cdot \mathbf{W}_{\cdot j}^t - \mathbf{X}_i \cdot \mathbf{W}_{\cdot k}^t$; \mathbf{e}_j 和 \mathbf{e}_k 分别表示第 j 和第 k 个元素值为 1 的 c 维单位向量。

算法 1 给出了 WSLDL-MLLR 算法的伪代码描述。

算法 1 基于局部标记序关系的弱监督标记分布学习方法 (WSLDL-MLLR)

输入 数据集 $(\mathbf{X}, \bar{\mathbf{D}})$, 参数 $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$;

输出 测试样本标记分布 $\hat{\mathbf{D}}$ 。

- 1) 初始化 $\mathbf{W}^0, \mathbf{C}^0$;
- 2) 令 $t = 1$;
- 3) repeat
- 4) 利用式 (2) 计算 \mathbf{W}^{t+1} ;
- 5) 利用式 (3) 计算 \mathbf{C}^{t+1} ;
- 6) $t = t + 1$;
- 7) until 收敛或达到最大迭代次数;
- 8) 通过公式 $\hat{\mathbf{D}} = \mathbf{XW}$ 返回预测的标记分布 $\hat{\mathbf{D}}$ 。

3 实验与结果分析

3.1 数据集

实验主要基于 14 个标记分布数据集, 包括 10 个酵母菌基因数据集 (Yeast^[19])、2 个面部表情数据集 (SJAFFE^[20] 和 SBU_3DF^[21]) 和 2 个图片情感分析数据集 (Emotion6^[22] 以及 Twitter^[23])。全部数据集的基本信息总结在表 1 中。

表 1 14 个数据集的基本信息
Table 1 Basic information of the 14 data sets

编号	数据集	样本个数	特征维数	标记个数
1	Yeast-alpha	2465	24	18
2	Yeast-cdc	2465	24	15
3	Yeast-cold	2465	24	4
4	Yeast-diau	2465	24	7
5	Yeast-dtt	2465	24	4
6	Yeast-elu	2465	24	14
7	Yeast-heat	2465	24	6
8	Yeast-spo	2465	24	6
9	Yeast-spo5	2465	24	3
10	Yeast-spoem	2465	24	2
11	SJAFFE	213	243	6
12	SBU_3DFE	2500	243	6
13	Emotion6	1980	168	7
14	Twitter	10045	168	8

3.2 评价指标

为了检验算法预测的标记分布与真实标记分

布的一致性, 一个很自然的想法是度量每个标记的平均距离和相似度。本文实验选择了 5 个标记分布学习的常用评价指标^[24], 它们分别为切比雪夫距离、克拉克距离、堪培拉距离、余弦距离和内积距离。其中前 3 个指标的数值越小表明该算法取得的性能越好, 后两个指标的数值越大越好。我们在表 2 中总结了所有指标的数学公式。

表 2 标记分布学习方法的评价指标
Table 2 Evaluation measures of label distribution learning

类别	名称	计算形式
距离	切比雪夫距离	$\max_j d_j - \hat{d}_j $
	克拉克距离	$\sqrt{\sum_{j=1}^L \frac{(d_j - \hat{d}_j)^2}{(d_j + \hat{d}_j)^2}}$
	堪培拉距离	$\sum_{j=1}^L \frac{ d_j - \hat{d}_j }{d_j + \hat{d}_j}$
相似度	余弦距离	$\frac{\sum_{j=1}^L d_j \hat{d}_j}{\sqrt{\sum_{j=1}^L d_j^2} \sqrt{\sum_{j=1}^L \hat{d}_j^2}}$
	内积距离	$\sum_{j=1}^L \min(d_j, \hat{d}_j)$

此外, 为了验证我们提出的弱监督标记分布学习算法 (WSLDL-MLLR) 在保持标记序关系方面的优势, 我们另外引入了一个在标记排序领域使用频繁且适用性广的度量指标, 即斯皮尔曼等级相关系数^[25], 其取值范围为 $[-1, 1]$, 结果越大表明算法在维持标记排序方面效果越好, 当预测的标记排序与真实标记排序完全一致时, 斯皮尔曼等级相关系数取 1。对于长度为 l 的真实标记分布 d_i 和预测标记分布 \hat{d}_i , 斯皮尔曼等级相关系数的公式定义为

$$s = 1 - \frac{6 \sum_i (d_i - \hat{d}_i)^2}{l(l^2 - 1)}$$

3.3 实验设置

为了获得弱监督标记分布学习任务所需的数据集, 我们在训练集中随机且均匀地选择数据, 并人为地将缺失数据置 0, 而测试集的标记分布是完全未知的。实验中, 数据集标记空间的缺失率 ω 分别设定为 30%、20% 和 10%。我们在这里使用了一个直推式的学习设置。本章实验随机选取数据集的 90% 用作训练, 10% 用于测试。最后, 通过 3.2 节中介绍的 6 个评价指标从不同角度衡量各种算法的性能差异。

实验中, 超参数 λ_1 、 λ_2 、 λ_3 、 λ_4 、 λ_5 的取值集合是 $\{10^{-6}, 10^{-5}, \dots, 10^4\}$ 。此外, 矩阵 \mathbf{W} 和 \mathbf{C} 的初始化是随机的, 模型训练的最大迭代次数为 100。

为了验证 WSLDL-MLLR 算法的有效性, 我们与 6 种不同的算法做了对比, 分别是: IncomLDL-a^[5]、IncomLDL-p^[5]、WSLDL-MCSC^[6]、IncomLDL-LR^[7]、IncomLDL-SNI^[8] 以及 GRME^[9]。这些弱监督标记分布学习算法中的超参数设定规则均遵循原文的建议。

3.4 实验结果

为了尽可能消除随机性带来的影响, 本文在每个数据集上进行 10 次 10 折交叉验证实验。具体地, 每次实验中, 每个数据集的样本被随机均匀划分为 10 份, 其中 9 份用于训练, 1 份用于测

试。我们对 10 次随机实验得到的所有评价指标结果求均值和方差, 由于文章篇幅有限, 我们仅给出部分实验结果。

表 3~5 给出了在不同缺失率下以“均值 \pm 标准差”的形式给出的实验结果, 对每个数据集上最好的算法结果进行加粗显示。为了便于观察, 表格中的所有算法均为简写: WSLDL-MLLR 简写为 MLLR、IncomLDL-SNI 简写为 SNI、IncomLDL-LR 简写为 LR、WSLDL-MCSC 简写为 MCSC。实验结果可以看出, 在距离和相似度指标上, WSLDL-MLLR 在大多数数据集上可以取得较好的性能。值得一提的是, 在排序指标上, 我们的算法在所有数据集上都取得了最优效果。这进一步说明 WSLDL-MLLR 可以在一定程度上帮助维持标记之间的局部序关系。

表 3 WSLDL-MLLR 在切比雪夫距离评价指标上的结果 (缺失率 $\omega=10\%$)
Table 3 Results of WSLDL-MLLR on Chebyshev evaluation (missing rate $\omega=10\%$)

数据集	MLLR	GRME	SNI	LR	MCSC	IncomLDL-a	IncomLDL-p
Yeast-alpha	0.0134\pm0.0000	0.0135 \pm 0.0002	0.0136 \pm 0.0003	0.0136 \pm 0.0000	0.0135 \pm 0.0000	0.0135 \pm 0.0005	0.0135 \pm 0.0000
Yeast-cdc	0.0162 \pm 0.0001	0.0161\pm0.0001	0.0163 \pm 0.0003	0.0164 \pm 0.0004	0.0163 \pm 0.0005	0.0163 \pm 0.0001	0.0163 \pm 0.0006
Yeast-cold	0.0511\pm0.0000	0.0512 \pm 0.0023	0.0510 \pm 0.0003	0.0513 \pm 0.0002	0.0512 \pm 0.0001	0.0520 \pm 0.0002	0.0521 \pm 0.0003
Yeast-diau	0.0366\pm0.0004	0.0370 \pm 0.0013	0.0376 \pm 0.0002	0.0375 \pm 0.0001	0.0371 \pm 0.0003	0.0378 \pm 0.0003	0.0377 \pm 0.0002
Yeast-dtt	0.0357\pm0.0003	0.0367 \pm 0.0011	0.0361 \pm 0.0003	0.0361 \pm 0.0005	0.0363 \pm 0.0002	0.0365 \pm 0.0003	0.0365 \pm 0.0003
Yeast-elu	0.0162\pm0.0002	0.0165 \pm 0.0003	0.0164 \pm 0.0004	0.0163 \pm 0.0002	0.0165 \pm 0.0002	0.0165 \pm 0.0003	0.0167 \pm 0.0002
Yeast-heat	0.0418\pm0.0002	0.0432 \pm 0.0015	0.0424 \pm 0.0004	0.0422 \pm 0.0001	0.0425 \pm 0.0005	0.0426 \pm 0.0002	0.0429 \pm 0.0006
Yeast-spo	0.0581 \pm 0.0001	0.0577\pm0.0023	0.0594 \pm 0.0003	0.0590 \pm 0.0004	0.0581 \pm 0.0025	0.0594 \pm 0.0003	0.0596 \pm 0.0003
Yeast-spo5	0.0907\pm0.0001	0.0912 \pm 0.0025	0.0911 \pm 0.0003	0.0915 \pm 0.0004	0.0910 \pm 0.0003	0.0916 \pm 0.0008	0.0918 \pm 0.0014
Yeast-spoem	0.0866\pm0.0002	0.0870 \pm 0.0005	0.0876 \pm 0.0005	0.0887 \pm 0.0003	0.0879 \pm 0.0005	0.0883 \pm 0.0004	0.0882 \pm 0.0004
SJAFFE	0.0914\pm0.0002	0.1159 \pm 0.0034	0.1163 \pm 0.0034	0.1026 \pm 0.0035	0.1023 \pm 0.0046	0.1223 \pm 0.0016	0.1235 \pm 0.0031
SBU_3DFE	0.1081\pm0.0001	0.1093 \pm 0.0023	0.1134 \pm 0.0025	0.1125 \pm 0.0021	0.1098 \pm 0.0034	0.1247 \pm 0.0034	0.1256 \pm 0.0024
Emotion6	0.3026\pm0.0001	0.3129 \pm 0.0023	0.3201 \pm 0.0045	0.3176 \pm 0.0029	0.3134 \pm 0.0076	0.3256 \pm 0.0023	0.3276 \pm 0.0035
Twitter	0.2929\pm0.0035	0.3154 \pm 0.0045	0.3276 \pm 0.0037	0.3289 \pm 0.0046	0.3167 \pm 0.0043	0.3285 \pm 0.0037	0.3305 \pm 0.0042

表 4 WSLDL-MLLR 在斯皮尔曼等级相关系数评价指标上的结果 (缺失率 $\omega=20\%$)
Table 4 Results of WSLDL-MLLR on Spearman's rank evaluation (missing rate $\omega=20\%$)

数据集	MLLR	GRME	SNI	LR	MCSC	IncomLDL-a	IncomLDL-p
Yeast-alpha	0.1992\pm0.0010	0.1952 \pm 0.0019	0.1894 \pm 0.0032	0.1884 \pm 0.0027	0.1961 \pm 0.0021	0.1742 \pm 0.0037	0.1738 \pm 0.0031
Yeast-cdc	0.1810\pm0.0021	0.1693 \pm 0.0020	0.1682 \pm 0.0025	0.1653 \pm 0.0021	0.1712 \pm 0.0021	0.1578 \pm 0.0043	0.1582 \pm 0.0035
Yeast-cold	0.2284\pm0.0023	0.2193 \pm 0.0017	0.2145 \pm 0.0025	0.2132 \pm 0.0022	0.2216 \pm 0.0026	0.2145 \pm 0.0021	0.2142 \pm 0.0034
Yeast-diau	0.3982\pm0.0019	0.3893 \pm 0.0017	0.3886 \pm 0.0024	0.3871 \pm 0.0023	0.3901 \pm 0.0014	0.3803 \pm 0.0027	0.3802 \pm 0.0014
Yeast-dtt	0.1982\pm0.0023	0.1872 \pm 0.0014	0.1883 \pm 0.0023	0.1871 \pm 0.0023	0.1897 \pm 0.0014	0.1812 \pm 0.0021	0.1813 \pm 0.0026
Yeast-elu	0.1883\pm0.0056	0.1823 \pm 0.0024	0.1836 \pm 0.0035	0.1838 \pm 0.0024	0.1841 \pm 0.0020	0.1785 \pm 0.0023	0.1789 \pm 0.0024

续表 4

数据集	MLLR	GRME	SNI	LR	MCSC	IncomLDL-a	IncomLDL-p
Yeast-heat	0.1673±0.0026	0.1612±0.0016	0.1621±0.0031	0.1613±0.0031	0.1625±0.0021	0.1552±0.0021	0.1554±0.0020
Yeast-spo	0.1986±0.0021	0.1932±0.0021	0.1921±0.0023	0.1913±0.0021	0.1942±0.0023	0.1883±0.0021	0.1876±0.0023
Yeast-spo5	0.1101±0.0010	0.1053±0.0013	0.1045±0.0021	0.1026±0.0023	0.1075±0.0035	0.0984±0.0031	0.0989±0.0021
Yeast-spoem	0.1165±0.0034	0.1123±0.0013	0.1110±0.0021	0.1112±0.0021	0.1135±0.0024	0.1096±0.0026	0.1091±0.0032
SJAFFE	0.5120±0.0031	0.4983±0.0021	0.4876±0.0025	0.4884±0.0023	0.4992±0.0031	0.4876±0.0021	0.4883±0.0023
SBU_3DFE	0.3543±0.0012	0.3512±0.0012	0.3501±0.0016	0.3483±0.0025	0.3495±0.0021	0.3432±0.0021	0.3429±0.0026
Emotion6	0.3392±0.0021	0.3321±0.0017	0.3312±0.0021	0.3301±0.0021	0.3367±0.0017	0.3300±0.0025	0.3294±0.0022
Twitter	0.5712±0.0020	0.5543±0.0021	0.5512±0.0023	0.5567±0.0026	0.5642±0.0017	0.5483±0.0045	0.5474±0.0032

表 5 WSLDL-MLLR 在内积距离评价指标上的结果 (缺失率 $\omega=30\%$)Table 5 Results of WSLDL-MLLR on Intersection evaluation (missing rate $\omega=30\%$)

数据集	MLLR	GRME	SNI	LR	MCSC	IncomLDL-a	IncomLDL-p
Yeast-alpha	0.9625±0.0000	0.9624±0.0000	0.9621±0.0007	0.9619±0.0008	0.9624±0.0000	0.9621±0.0002	0.9623±0.0003
Yeast-cdc	0.9575±0.0000	0.9574±0.0000	0.9571±0.0009	0.9572±0.0010	0.9574±0.0001	0.9571±0.0006	0.9573±0.0002
Yeast-cold	0.9408±0.0001	0.9407±0.0001	0.9403±0.0002	0.9400±0.0001	0.9410±0.0006	0.9398±0.0002	0.9396±0.0008
Yeast-diau	0.9402±0.0000	0.9399±0.0002	0.9398±0.0001	0.9389±0.0002	0.9407±0.0003	0.9385±0.0002	0.9386±0.0002
Yeast-dtt	0.9585±0.0000	0.9585±0.0001	0.9581±0.0002	0.9583±0.0001	0.9585±0.0002	0.9577±0.0002	0.9575±0.0003
Yeast-elu	0.9590±0.0002	0.9585±0.0000	0.9582±0.0001	0.9581±0.0000	0.9587±0.0001	0.9583±0.0002	0.9580±0.0001
Yeast-heat	0.9408±0.0002	0.9403±0.0000	0.9395±0.0002	0.9396±0.0001	0.9406±0.0002	0.9390±0.0002	0.9392±0.0001
Yeast-spo	0.9156±0.0003	0.9160±0.0002	0.9152±0.0007	0.9150±0.0001	0.9158±0.0005	0.9152±0.0003	0.9151±0.0002
Yeast-spo5	0.9087±0.0000	0.9093±0.0000	0.9083±0.0001	0.9082±0.0000	0.9097±0.0002	0.9078±0.0003	0.9079±0.0002
Yeast-spoem	0.9132±0.0000	0.9126±0.0002	0.9125±0.0002	0.9124±0.0001	0.9127±0.0001	0.9120±0.0002	0.9121±0.0001
SJAFFE	0.8756±0.0004	0.8502±0.0004	0.8645±0.0003	0.8639±0.0005	0.8621±0.0010	0.8473±0.0010	0.8467±0.0010
SBU_3DFE	0.8632±0.0007	0.8625±0.0032	0.8593±0.0025	0.8578±0.0043	0.8627±0.0032	0.8513±0.0043	0.8505±0.0032
Emotion6	0.5878±0.0002	0.5765±0.0024	0.5734±0.0031	0.5725±0.0024	0.5814±0.0032	0.5712±0.0021	0.5703±0.0019
Twitter	0.6174±0.0037	0.6123±0.0045	0.6102±0.0027	0.6108±0.0031	0.6145±0.0024	0.6043±0.0056	0.6052±0.0041

3.5 显著性检验

为了验证我们提出的 WSLDL-MLLR 算法与其他现有的弱监督标记分布学习算法有显著差异,我们在 6 个评价指标上对参与比较的算法进

行了显著性水平为 0.05 的双尾 t 检验。以显著优于/不显著/显著劣于的形式记录检验结果,如表 6~8 所示。结果表明所提算法在大部分数据集上能够显著优于对比算法。

表 6 显著性水平为 0.05 的双尾 t 检验结果 (缺失率 $\omega=10\%$)Table 6 Results under the pairwise two-tailed t -test with a 0.05 significance level (missing rate $\omega=10\%$)

评价指标	GRME	SNI	LR	MCSC	IncomLDL-a	IncomLDL-p
切比雪夫	11/2/1	14/0/0	14/0/0	12/2/0	13/1/0	13/0/1
克拉克	10/1/3	14/0/0	14/0/0	11/2/1	14/0/0	14/0/0
堪培拉	10/2/2	14/0/0	14/0/0	11/1/2	14/0/0	14/0/0
余弦	14/0/0	12/2/0	13/1/0	13/1/0	14/0/0	14/0/0
内积	12/1/1	14/0/0	14/0/0	12/1/1	14/0/0	14/0/0
斯皮尔曼	14/0/0	14/0/0	14/0/0	14/0/0	14/0/0	14/0/0

表 7 显著性水平为 0.05 的双尾 t 检验结果 (缺失率 $\omega = 20\%$)Table 7 Results under the pairwise two-tailed t-test with a 0.05 significance level (missing rate $\omega = 20\%$)

评价指标	GRME	SNI	LR	MCSC	IncomLDL-a	IncomLDL-p
切比雪夫	11/1/2	14/0/0	14/0/0	13/0/1	14/0/0	14/0/0
克拉克	13/0/1	11/1/2	12/1/1	13/0/1	14/0/0	14/0/0
堪培拉	9/0/5	12/0/2	11/1/2	12/0/2	14/0/0	14/0/0
余弦	11/2/1	14/0/0	14/0/0	12/1/1	14/0/0	14/0/0
内积	12/0/2	14/0/0	14/0/0	13/1/0	14/0/0	14/0/0
斯皮尔曼	14/0/0	14/0/0	14/0/0	14/0/0	14/0/0	14/0/0

表 8 显著性水平为 0.05 的双尾 t 检验结果 (缺失率 $\omega = 30\%$)Table 8 Results under the pairwise two-tailed t-test with a 0.05 significance level (missing rate $\omega = 30\%$)

评价指标	GRME	SNI	LR	MCSC	IncomLDL-a	IncomLDL-p
切比雪夫	12/0/2	14/0/0	14/0/0	9/2/3	14/0/0	14/0/0
克拉克	11/1/2	14/0/0	14/0/0	8/2/4	14/0/0	14/0/0
堪培拉	10/2/2	14/0/0	14/0/0	9/2/3	14/0/0	14/0/0
余弦	11/1/2	14/0/0	14/0/0	10/2/2	14/0/0	14/0/0
内积	10/1/3	14/0/0	13/1/0	9/0/5	14/0/0	14/0/0
斯皮尔曼	14/0/0	14/0/0	14/0/0	14/0/0	14/0/0	14/0/0

3.6 参数分析

本节分析了模型中 λ_1 、 λ_2 、 λ_3 、 λ_4 和 λ_5 这 5 个超参数的不同设置对算法整体性能的影响。由于在 14 个数据集上的分析结果较为类似, 本文选取 Emotion6 数据集的分析结果作为代表。在该数据集上, 首先通过实验得到最优结果下 5 个参数的设置, 即, $\lambda_1 = 1$ 、 $\lambda_2 = 10^{-2}$ 、 $\lambda_3 = 10^{-1}$ 、 $\lambda_4 = 10^{-2}$ 和 $\lambda_5 = 10^{-1}$ 。接着每次调整 1 个参数的取值并固定其余 4 个参数, 分别对 λ_1 、 λ_2 、 λ_3 、 λ_4 和 λ_5 进行实验分析, 5 个参数的取值集合都为 $\{10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10^1, 10^2, 10^3, 10^4\}$ 。实验结果的折线图如图 2 所示。由于篇幅限制, 同时为了显示本文对维持弱监督数据集中标记序关系的突出效果, 本文只给出斯皮尔曼等级相关系数 ρ 评价指标的参数分析结果, 其余评价指标的参数分析结果与其类似。

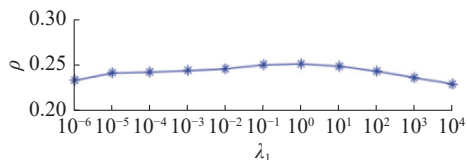
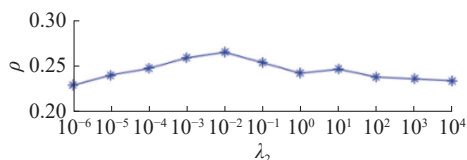
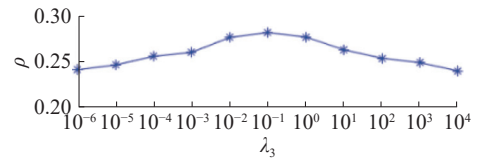
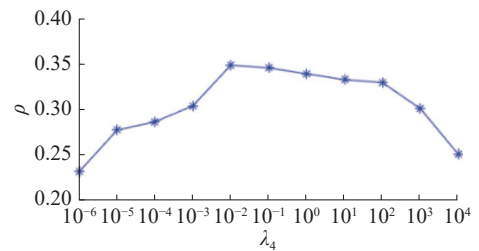
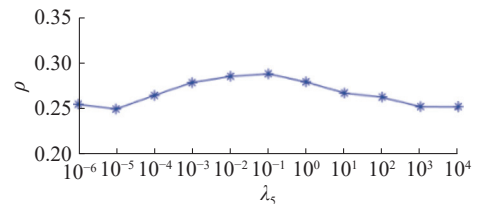
(a) λ_1 对指标的影响(b) λ_2 对指标的影响(c) λ_3 对指标的影响(d) λ_4 对指标的影响(e) λ_5 对指标的影响

图 2 5 个参数在 Emotion6 上关于斯皮尔曼等级相关系数评价指标的影响 (缺失率为 10%)

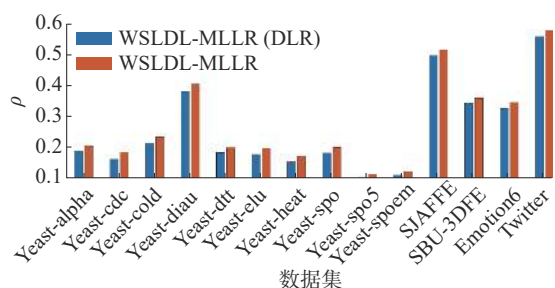
Fig. 2 Influence of 5 parameters on Emotion6 regarding Spearman's rank (missing rate of 10%)

通过分析对比这 5 个参数的不同设置得到的实验结果, 可以发现, 本文提出的局部序关系思想对维持弱监督数据集的标记序关系有明显作用。如图 2 所示, λ_4 对应标记排序损失项, 对比其他

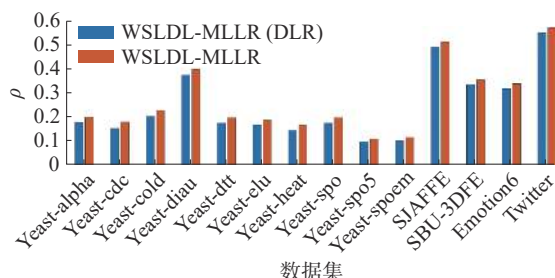
4 个参数对应的结果曲线, λ_4 对整体模型在斯皮尔曼等级相关系数评价指标上的结果影响较大, λ_4 的设置对指标能否取得最优值起着关键作用。

3.7 消融实验

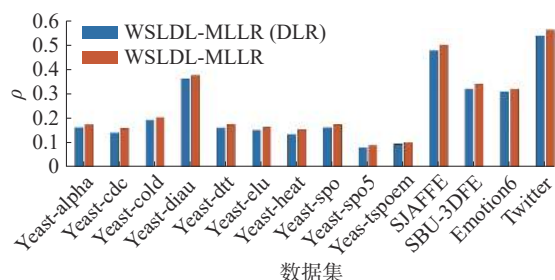
为了进一步验证在 WSLDL-MLLR 中引入局部标记序关系的有效性, 我们对这部分内容进行了消融实验。在 14 个数据集上的实验结果柱状图如图 3 所示。其中, WSLDL-MLLR(DLR) 表示剔除局部标记序关系优化项之后的模型。实验结果表明, 在斯皮尔曼等级相关系数评价指标上, 包含局部序关系优化项的弱监督标记分布学习算法 WSLDL-MLLR 在所有数据集上的结果都优于 WSLDL-MLLR(DLR)。



(a) 缺失率为 10% 时斯皮尔曼等级相关系数评价指标的结果



(b) 缺失率为 20% 时斯皮尔曼等级相关系数评价指标的结果



(c) 缺失率为 30% 时斯皮尔曼等级相关系数评价指标的结果

图 3 在 14 个数据集上的消融实验

Fig. 3 Ablation experiments on 14 data sets

4 结束语

为了设计出能更有效处理弱监督数据的标记分布学习算法, 本文刻画了未缺失数据的标记序关系, 并将维持这种局部序关系加入算法优化目

标。此外, 区分于其他现有的弱监督标记分布学习算法, 本文采用主动挖掘、不断更新标记相关性的方式来引导模型训练。从上述两点出发, 本文设计了一个全新的弱监督标记分布学习算法 (WSLDL-MLLR), 并通过大量的实验验证了算法的有效性。未来我们将探索更多贴合实际应用的弱监督标记分布学习算法。

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.
GENG Xin, XU Ning. Label distribution learning and label enhancement[J]. *Scientia sinica: informationis*, 2018, 48(5): 521–530.
- [2] 邵东恒, 杨文元, 赵红. 应用 k-means 算法实现标记分布学习 [J]. *智能系统学报*, 2017, 12(3): 325–332.
SHAO Dongheng, YANG Wenyuan, ZHAO Hong. Label distribution learning based on k-means algorithm[J]. *CAAI transactions on intelligent systems*, 2017, 12(3): 325–332.
- [3] 王一宾, 李田力, 程玉胜. 结合谱聚类的标记分布学习 [J]. *智能系统学报*, 2019, 14(5): 966–973.
WANG Yibin, LI Tianli, CHENG Yusheng. Label distribution learning based on spectral clustering[J]. *CAAI transactions on intelligent systems*, 2019, 14(5): 966–973.
- [4] 刘睿馨, 刘新媛, 李晨. 基于低秩表示的标记分布学习算法 [J]. *模式识别与人工智能*, 2021, 34(2): 146–156.
LIU Ruixin, LIU Xinyuan, LI Chen. Label distribution learning method based on low-rank representation[J]. *Pattern recognition and artificial intelligence*, 2021, 34(2): 146–156.
- [5] MIAO Xu, ZHOU Zhihua. Incomplete label distribution learning[C]//Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence Main track. Melbourne: IJCAI, 2017: 3175–3181.
- [6] JIA Xiuyi, REN Tingting, CHEN Lei, et al. Weakly supervised label distribution learning based on transductive matrix completion with sample correlations[J]. *Pattern recognition letters*, 2019, 125: 453–462.
- [7] ZENG Xueqiang, CHEN Sufen, XIANG Run, et al. Filling missing values by local reconstruction for incomplete label distribution learning[J]. *International journal of wireless and mobile computing*, 2019, 16(4): 314.
- [8] ZENG Xueqiang, CHEN Sufen, XIANG Run, et al. Incomplete label distribution learning based on supervised neighborhood information[J]. *International journal of machine learning and cybernetics*, 2020, 11(1): 111–121.
- [9] XU Chao, GU Shilin, TAO Hong, et al. Fragmentary la-

- bel distribution learning via graph regularized maximum entropy criteria[J]. *Pattern recognition letters*, 2021, 145: 147–156.
- [10] 任婷婷. 基于数据相关性挖掘的标记分布学习算法研究 [D]. 南京: 南京理工大学, 2020.
- REN Tingting. Label distribution learning algorithm based on data correlation mining[D]. Nanjing: Nanjing University of Science and Technology, 2020.
- [11] 游美琳. 基于层次结构的标记分布学习研究 [D]. 武汉: 中南财经政法大学, 2020.
- YOU Meilin. Research on label distribution learning based on hierarchical structure[D]. Wuhan: Zhongnan University of Economics and Law, 2020.
- [12] Abdi H. The method of least squares[J]. *Encyclopedia of measurement and statistics*, 2007, 1: 530–532.
- [13] WE Deming, XI. The application of least squares[J]. *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, 1931, 11(68): 146–158.
- [14] HUANG Jun, QIN Feng, ZHENG Xiao, et al. Improving multi-label classification with missing labels by learning label-specific features[J]. *Information sciences*, 2019, 492: 124–146.
- [15] CHENG Yusheng, QIAN Kun, MIN Fan. Global and local attention-based multi-label learning with missing labels[J]. *Information sciences*, 2022, 594: 20–42.
- [16] LI Xue, SHEN Bin, LIU Baodi, et al. Ranking-preserving low-rank factorization for image annotation with missing labels[J]. *IEEE transactions on multimedia*, 2018, 20(5): 1169–1178.
- [17] JIA Xiuyi, SHEN Xiaoxia, LI Weiwei, et al. Label distribution learning by maintaining label ranking relation[J]. *IEEE transactions on knowledge and data engineering*, 2021(99): 1.
- [18] LIU Dong C, NOCEDAL J. On the limited memory BFGS method for large scale optimization[J]. *Mathematical programming*, 1989, 45(1/2/3): 503–528.
- [19] EISEN M B, SPELLMAN P T, BROWN P O, et al. Cluster analysis and display of genome-wide expression patterns[J]. *Proceedings of the national academy of sciences of the United States of America*, 1998, 95(25): 14863–14868.
- [20] LYONS M, AKAMATSU S, KAMACHI M, et al. Coding facial expressions with Gabor wavelets[C]//Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition. Nara: IEEE, 1998: 200–205.
- [21] YIN Lijun, WEI Xiaozhou, SUN Yi, et al. A 3D facial expression database for facial behavior research[C]//7th International Conference on Automatic Face and Gesture Recognition. Southampton: IEEE, 2006: 211–216.
- [22] PENG Kuanchuan, CHEN T, SADOVNIK A, et al. A mixed bag of emotions: model, predict, and transfer emotion distributions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015: 860–868.
- [23] YANG Jufeng, SUN Ming, SUN Xiaoxiao. Learning visual sentiment distributions via augmented conditional probability neural network[C]//AAAI'17: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. New York: ACM, 2017: 224–230.
- [24] CHA S H. Comprehensive survey on distance/similarity measures between probability density functions[J]. *International journal of mathematical models and methods in applied sciences*, 2007, 1(4): 300–307.
- [25] SPEARMAN C. The proof and measurement of association between two things[J]. *International journal of epidemiology*, 2010, 39(5): 1137–1150.

作者简介:



秦天, 硕士研究生, 主要研究方向为机器学习和数据挖掘。



滕齐发, 硕士研究生, 主要研究方向为机器学习和数据挖掘。



贾修一, 副教授, 博士生导师, 博士, CCF 高级会员, 主要研究方向为机器学习、粒计算和数据挖掘。发表学术论文 60 余篇。