



基于排他约束的去噪偏多标记学习方法

高奕^{1,2}, 李佳祺¹, 张敏灵^{1,2*}

1. 东南大学计算机科学与工程学院, 南京 211189

2. 计算机网络和信息集成教育部重点实验室 (东南大学), 南京 211189

* 通信作者. E-mail: zhangml@seu.edu.cn

收稿日期: 2025-12-04; 修回日期: 2026-04-17; 接受日期: 2026-05-21; 网络出版日期: 2026-10-28

国家自然科学基金 (批准号: 62225602, 624B2042) 资助项目

摘要 在偏多标记学习中, 每个训练样本对应一个候选标记集合, 该集合包含真实标记与噪声标记. 消歧是解决偏多标记学习问题的主要策略, 现有方法多依赖建模候选标记集合内的标记预测概率或特定假设 (如平滑性、稀疏性等), 但这类假设在实际场景中往往难以满足, 同时忽略了非候选标记所蕴含的排他性信息. 为有效抑制噪声并提升模型对真实标记的辨别能力, 本文提出基于排他约束的去噪偏多标记学习框架 DeNo. 该框架在候选集合内部构建集中式伪标记分布, 逐轮强化高置信度标记并实现渐进式去噪, 同时在训练过程中采用平方损失保持模型输出与更新分布持续对齐, 从而避免误差累积. 此外, DeNo 利用非候选标记所蕴含的排他性信息抑制无关标记激活, 挖掘可靠负信号并进一步减少无关信息的干扰. 在 11 个多标记数据集上的实验结果表明, 所提方法在抑制噪声和提升模型性能方面均优于对比方法, 从而验证了其有效性. 论文代码链接: <https://github.com/gaoyi439/DeNo>.

关键词 多标记学习, 偏多标记学习, 去噪, 平方损失, 排他性信息

1 引言

在许多实际应用中, 一个样本往往同时对应多个标记. 例如, 在图像标注任务中, 一张街景图像可能同时被标注为“路灯”、“建筑”或“城市”等^[1,2]; 在生物信息学中, 一段基因序列也可能具备“代谢”、“转录”、“蛋白质合成”等多种功能^[3,4]. 多标记学习 (multi-label learning, MLL) 旨在研究如何从具有多个标记的样本中学习出能够预测未见样本真实标记集合的分类模型^[5~8].

由于获取大规模且准确标注的多标记数据需要投入大量人力成本, 众包平台逐渐成为常用的数据收集途径^[9,10]. 然而, 部分标记的准确识别依赖于专业领域知识, 非专业标注者往往难以胜任, 可能分配无效标记, 导致样本被过度标注. 如图 1 中的“Husky”可能被误标为“Wolf”, 使候选标记集合中混入无效标记, 从而引发偏多标记学习 (partial multi-label learning, PML) 问题^[11,12]. 在该学习范式中, 每个样本与一个候选标记集合相关联, 其中仅部分标记为真实标记^[13~15]. PML 的目标是在不准确的监督信息下, 学习一个能够逼近传统 MLL 性能的分类模型^[16,17]. 针对 PML 问题, 早期研究通常将候选标记集合中的所有标记视为真实标记, 并直接采用传统 MLL 算法进行训练^[7,13,18]. 然而, 这类直接策略忽略了候选标记集合中普遍存在的大量噪声标记, 导致

引用格式: 高奕, 李佳祺, 张敏灵. 基于排他约束的去噪偏多标记学习方法. 中国科学: 信息科学, 2026, doi: 10.1360/SSI-2025-0525

Gao Y, Li J-Q, Zhang M-L. A denoising framework with exclusivity constraints for partial multi-label learning. Sci Sin Inform, 2026, doi: 10.1360/SSI-2025-0525



Candidate label set

True labels are **red**

Husky
Wolf
Sun
Snowfield
Car
Forest

图 1 (网络版彩图) 偏多标记样本示例。

Figure 1 (Color online) Example of a partial multi-labeled sample.

模型在训练过程中容易受到错误监督信号的干扰, 从而显著降低泛化性能. 尤其在采用平方损失 (square loss) 进行优化时, 该策略对所有候选标记一视同仁, 难以根据预测概率动态聚焦于更可靠的监督信息. 为减轻噪声 (无效) 标记对求解 PML 问题带来的影响, 后续研究主要从消歧 (disambiguation) 的角度进行建模. 由于预测概率能够反映候选标记成为真实标记的可能性, 基于消歧的方法通常依据预测概率对候选标记进行校正或加权学习, 进而从候选标记中复原真实标记信息^[13,19]. 根据建模思想的不同, 这类方法可大致被分为两类. (1) 基于平滑假设的方法: 通过样本间的结构相似性来传播标记置信度分数, 如标记传播 (label propagation)^[20,21]、标记增强 (label enhancement)^[22,23]、簇分配 (cluster assignment)^[12,24] 等. (2) 基于稀疏假设的方法: 假设候选标记中的噪声占比较少, 通过稀疏正则来学习噪声标记矩阵^[25,26] 或进行噪声变量识别^[27] 以抑制噪声干扰. 然而, 上述方法依赖特定的启发式假设 (如平滑性或稀疏性), 在低质量标注数据下可能表现不稳定^[13]. 此外, 这类方法主要关注候选集合内部的置信度建模, 而忽略了候选标记之外无关标记所蕴含的排他性信息, 而这些信息同样能够为模型学习提供有价值的监督信号.

为解决上述问题, 本文从监督信息建模的角度出发, 对 PML 中的消歧机制进行重新刻画. 不同于现有方法主要基于独立置信度建模、负标记推断或显式噪声建模的思路, 本文将候选标记集合视为一个整体, 在其内部构建归一化分布以引入显式竞争关系, 从而将学习目标由单标记概率估计转化为集合内的概率质量分配与排序. 同时, 本文不对噪声进行显式识别, 而是通过分布演化机制使噪声标记在竞争中逐步被抑制, 并结合非候选标记的排他约束提供稳定的负监督信号, 从而实现更加稳健的去噪过程. 基于上述思想, 本文提出一种基于排他约束的去噪偏标记学习框架 DeNo (DeNoising partial multi-Label learning with exclusive constraints). 具体而言, 不同于传统 MLL 中使用独立的概率映射方式, DeNo 显式引入分布集中机制, 通过采用集中式的归一化分布, 使模型的预测概率集中于候选标记内部, 减少非候选标记被误激活对训练带来的负面影响. 随后, 通过逐轮更新这一分布, DeNo 会持续强化具有高预测概率的标记并抑制低预测概率标记, 从而形成渐进式的去噪过程. 与此同时, 为确保逐轮更新的集中式伪标记分布在训练过程中保持稳定并避免误差积累, DeNo 在整个训练阶段引入平方损失, 对模型输出与更新后的分布进行持续对齐, 为去噪过程提供平滑而可靠的优化信号. 此外, DeNo 进一步利用非候选标记所蕴含的排他性信息, 通过排他正则惩罚无关标记激活, 从全局角度挤压噪声空间, 使模型能够同时挖掘更多可靠的负信号. 在上述策略的协同作用下, DeNo 能自动聚焦于候选标记集合中的潜在真实标记, 从而有效缓解噪声干扰并提升泛化性能. 本文在 11 个 MLL 数据集上进行实验, 并与六种最先进方法进行对比. 结果表明, 所提出的基于排他约束的去噪偏标记学习框架 DeNo 能够更有效地抑制噪声干扰, 从而提升模型性能, 进一步验证本文所提方法的有效性.

本文其余部分安排如下. 第 2 节简要介绍相关工作, 第 3 节详细阐述 DeNo 的技术细节, 第 4 节展示并分析实验结果, 第 5 节对全文进行总结.

2 相关工作

2.1 多标记学习

在传统 MLL 中, 每个样本可能同时关联多个相关标记, 因此有效建模标记间的相关性是提升模型性能的关键之一 [7,18]. 根据标记相关性的建模方式, 现有方法可大致被分为三类: 一阶 (first-order) 方法 [28,29]、二阶 (second-order) 方法 [30,31] 和高阶 (high-order) 方法 [32,33]. 一阶方法通常将 MLL 任务分解为若干相互独立的二分类问题, 对每个类别分别训练独立的二元分类器 [28,29]. 这类方法结构简单且易于扩展, 但忽略了标记间的依赖关系, 限制了模型对结构信息的利用, 从而影响了模型性能的提升. 为克服一阶方法的局限性, 后续研究开始考虑建模二阶标记相关性, 即成对 (pair-wise) 标记之间的关联, 例如, 基于标记对排序或条件随机场的方法, 在类别间关系相对明确的任务中取得更好的表现 [30,31]. 然而, 实际场景中的标记依赖通常比二阶关系更为复杂 [10,34]. 因此, 高阶方法进一步探索标记子集或整个标记空间内标记之间的高阶相关性 [33,35]. 例如, Singh 等 [33] 构建标记之间的图结构, 并结合自适应图结构学习与注意力机制, 通过学习邻接矩阵来动态建模高阶标记关系. 尽管高阶方法在表达能力上更强, 能够捕获复杂的结构信息, 但与一阶和二阶方法相比, 其计算成本也显著增加 [10]. 由于 MLL 方法通常依赖完备且可信的标注数据进行学习, 而 PML 仅提供具有不准确监督信息的候选标记集合, 使任务难度显著提升. 因此, 依赖准确标注的传统 MLL 方法无法直接应用于 PML 场景.

2.2 偏多标记学习

PML 旨在处理不准确标注条件下的 MLL 任务, 其核心挑战在于候选标记集合中同时存在真实标记与噪声标记, 模型需在训练过程中自动完成消歧并准确识别真实标记. 根据设计假设与优化目标的不同, 现有基于消歧的 PML 方法大致可分为三类: 基于平滑假设的方法、基于低秩约束的方法和基于稀疏正则的方法 [13]. 基于平滑假设的方法认为, 特征空间中相邻样本应具有相似的标记分布, 因此利用样本结构信息辅助消歧. 早期代表性工作通过交替优化预测模型与标记预测概率来逐步恢复真实标记, 其核心思想是基于潜在的簇结构对候选标记进行校准 [12]. 后续研究进一步沿此方向发展, 包括基于图结构的标记传播方法 [20,36]、将消歧视作软标记恢复的标记增强方法 [22,37], 以及通过联合嵌入实现隐式消歧的表示学习方法 [38,39] 等. 基于低秩约束的方法利用标记结构的低秩性来消除歧义, 例如, Yu 等 [11] 构建低秩标记表示矩阵, 以从候选标记集合中自动恢复潜在真实标记; Lyu 等 [40] 则基于自表示模型在潜在子空间中学习样本间的低秩结构, 并通过先验知识约束获得更加稳定的子空间表示, 以提取可信标记并实现有效消歧. 这类方法通过低秩结构约束提升了消歧过程的稳健性, 但效果依赖于数据是否满足低秩性质. 基于稀疏正则的方法假设候选标记中的噪声占比较低, 通过对隐藏在候选标记集合中的噪声标记施加稀疏约束来实现间接去噪. 例如, 对噪声标记矩阵引入稀疏惩罚或通过稀疏结构识别无效候选标记 [25~27]. 然而, 这类方法的性能高度依赖于稀疏性假设的有效性, 当标注质量较低或噪声比例较高时, 可能导致次优结果. 上述方法普遍依赖平滑性、低秩或稀疏性等结构性假设, 而这些先验在复杂真实场景中往往难以满足. 同时, 这类方法主要关注候选标记集合内部的标记预测概率建模, 对非候选标记中蕴含的排他性负信息利用不足, 容易导致无关标记被误激活. 为弥补这些不足, 本文提出 DeNo 框架, 通过联合利用候选与非候选标记的监督信息, 并借助分布集中机制逐轮强化高预测概率标记, 使模型输出持续与更新的监督分布对齐, 从而实现更稳健的去噪消歧.

3 DeNo 方法

3.1 符号说明

设输入空间为 $\mathcal{X} \in \mathbb{R}^d$, 其中每个样本由 d 维特征表示; 标记空间为 $\mathcal{Y} = \{1, 2, \dots, K\}$, 包含 K 个不同类别. 在 PML 任务中, 每个训练样本 $\mathbf{x}_i \in \mathcal{X}$ 与一个候选标记集合 $S_i \subseteq \mathcal{Y}$ 关联. 根据 PML 任务的基本假设, 每个样本 $\mathbf{x}_i$ 所对应的真实标记集合 Y_i 隐藏在候选标记集合 S_i 中, 即 $Y_i \subseteq S_i$, 且 Y_i 在整个训练过程中不可观测. 给定

偏多标记数据集 $\tilde{D} = \{(\mathbf{x}_i, S_i) \mid 1 \leq i \leq n\}$, PML 任务的目标与传统 MLL 任务一致, 旨在学习一个多标记分类器 $h: \mathcal{X} \rightarrow 2^{\mathcal{Y}}$, 可以预测未见样本的所有真实标记. 为简化符号表示, 我们使用一个 K 维指示向量 $\mathbf{s}_i \in \{0, 1\}^K$ 表示候选标记集合 S_i : 当标记 $j \in S_i$ 时, 向量 $\mathbf{s}_i$ 的第 j 个元素 $s_{ij} = 1$, 否则 $s_{ij} = 0$. 同理, 真实标记集合 Y_i 也可表示为 K 维指示向量 $\mathbf{y}_i \in \{0, 1\}^K$.

3.2 分布集中机制

在 PML 中, 候选标记集合同时包含真实标记与大量噪声标记, 使得模型在训练初期难以准确区分不同标记的重要性. 现有基于神经网络模型的 PML 方法普遍在输出层使用 sigmoid 函数, 将每个类别的输出独立映射至 $[0, 1]$ 区间, 转化为对标记的预测概率^[13, 41]. 然而, sigmoid 函数独立建模的特性导致不同标记之间缺乏竞争关系, 难以在候选集合内部形成可靠的相对预测概率排序, 从而限制了消歧效果. 相比之下, softmax 函数的归一化特性更适合用于标记筛选, 可在候选标记内部构建稳定的相对预测概率排序. 基于这一优势, 本文利用 softmax 函数构建了一个基于模型输出的分布集中机制, 用于逐步强化候选标记集合中的潜在真实标记并抑制不可靠标记, 从而实现有效去噪. 此外, 本文从理论上证明了该机制具有噪声抑制能力.

具体而言, 多标记分类模型对于输入样本 $\mathbf{x}_i$ 的原始输出记为 $\mathbf{z}(\mathbf{x}_i) \in \mathbb{R}^K$, 其中 $z_j(\mathbf{x}_i)$ 表示在标记 j 上的原始输出. 令 $\mathbf{z}(\mathbf{x}_i)$ 通过 softmax 函数得到的预测概率分布记为 $\phi(\mathbf{x}_i) \in [0, 1]^K$, 其中第 j 个元素的计算方式如下:

$$\phi_j(\mathbf{x}_i) = \frac{e^{z_j(\mathbf{x}_i)}}{\sum_{k=1}^K e^{z_k(\mathbf{x}_i)}}. \quad (1)$$

为进一步将概率质量集中于候选标记集合 S_i 内, 并逐步过滤候选标记集合中的噪声标记, 我们对上述分布进行重归一化, 得到集中伪标记分布 $\mathbf{p}(\mathbf{x}_i)$. 其中 $\mathbf{p}(\mathbf{x}_i)$ 的第 j 个元素的计算方式为

$$p_j(\mathbf{x}_i) = \begin{cases} \phi_j(\mathbf{x}_i) / \sum_{k \in S_i} \phi_k(\mathbf{x}_i), & j \in S_i, \\ 0, & \text{其他}. \end{cases} \quad (2)$$

该集中伪标记分布会代替候选标记向量 $\mathbf{s}_i$ 在损失函数中作为监督信息, 促使模型在候选标记集合内建立相对预测概率排序, 使潜在真实标记的预测概率在多次迭代中被逐步放大, 并自然形成概率分布的集中趋势, 从而实现渐进式消歧. 定理 1 给出在这一策略下, 噪声标记所能获得的总预测概率质量的上界.

定理1 对任意输入 $\mathbf{x}_i$, 其对应的候选标记集合为 S_i , 真实标记集合为 $Y_i \subseteq S_i$. 若存在间隔 $\gamma > 0$, 使得 $\min_{j \in Y_i} z_j(\mathbf{x}_i) - \max_{l \in S_i \setminus Y_i} z_l(\mathbf{x}_i) \geq \gamma$, 则噪声标记的总概率质量 ℓ 满足

$$\ell = \sum_{j \in S_i \setminus Y_i} p_j(\mathbf{x}_i) \leq |S_i \setminus Y_i| e^{-\gamma}.$$

证明 令标记 $j^* = \arg \max_{j \in Y_i} z_j$, 对任意噪声标记 $j \in S_i \setminus Y_i$, 有

$$\frac{p_j(\mathbf{x}_i)}{p_{j^*}(\mathbf{x}_i)} = \frac{\phi_j(\mathbf{x}_i) / \sum_{k \in S_i} \phi_k(\mathbf{x}_i)}{\phi_{j^*}(\mathbf{x}_i) / \sum_{k \in S_i} \phi_k(\mathbf{x}_i)} = e^{(z_j(\mathbf{x}_i) - z_{j^*}(\mathbf{x}_i))} \leq e^{-\gamma}.$$

因此, 有 $p_j(\mathbf{x}_i) \leq e^{-\gamma} p_{j^*}(\mathbf{x}_i)$. 对等式两边所有的 $j \in S_i \setminus Y_i$ 求和, 得

$$\ell = \sum_{j \in S_i \setminus Y_i} p_j(\mathbf{x}_i) \leq |S_i \setminus Y_i| e^{-\gamma} p_{j^*}(\mathbf{x}_i).$$

由于 $p_{j^*}(\mathbf{x}_i) \leq 1$, 则得到 $\ell \leq |S_i \setminus Y_i| e^{-\gamma}$. □

由此可见, 在分布集中机制下, 候选标记的概率总和为 1, 从而在内部形成竞争. 当模型在真实标记上的原始输出与噪声标记上的原始输出存在间隔 γ 时, 噪声标记的概率将呈指数级衰减, 确保模型在迭代中持续削弱噪声影响, 实现渐进式去噪, 而基于 sigmoid 的独立建模并不具备类似的性质. 值得注意的是, 上述间隔假设主

要用于刻画分布集中机制在理想条件下的噪声抑制能力, 是一种充分条件而非必要条件. 在实际训练过程中, 尤其是在训练初期或高噪声场景下, 该间隔可能尚未形成. 尽管如此, 由于 softmax 归一化引入的相对竞争关系, 即使在预测差异较小时, 不同候选标记之间仍会产生初步的概率分配差异. 在迭代优化过程中, 这种差异会通过集中伪标记分布的逐轮更新被逐步放大, 从而推动潜在真实标记与噪声标记之间间隔的形成与扩大. 此外, 本文引入的平方损失对齐项也有助于缓解早期预测不稳定带来的影响, 从而保证分布集中机制在弱间隔条件下仍能稳定发挥作用. 下一小节会介绍如何利用集中伪标记分布信息监督模型进行训练.

3.3 损失函数设计

在获得集中伪标记分布后, 模型仍需通过合适的优化目标, 将消歧信息有效传递至网络参数更新过程中. 为此, 本文设计了由三部分组成的联合损失函数, 将集中伪标记分布作为监督信息之一, 同时实现: 逐轮强化候选集中的可信标记、保持模型输出与更新分布的持续对齐以及利用非候选标记的排他信息约束无关标记激活. 先前研究表明, 在训练过程中直接使用模型当前预测结果作为伪标记进行自监督, 会导致模型输出与监督分布相互依赖, 进一步放大已有偏差, 造成模型性能退化, 并可能使所有预测结果收敛到同一类别而出现退化解 (collapsed solution) [42~44]. 为避免上述问题, DeNo 在每轮迭代中使用上一轮生成的集中伪标记分布作为当前轮次的监督信号. 该策略一方面保持监督分布的相对稳定, 避免因模型瞬时波动造成集中伪标记分布剧烈变化; 另一方面能够将逐轮强化的消歧信息持续传递, 从而实现渐进式去噪并确保训练过程的稳健收敛.

基于上述考量, 本文首先构建基于候选标记的似然项, 使模型输出在候选标记内部形成竞争关系, 并在当前第 t 轮训练中使用上一轮得到的集中伪标记分布 $\mathbf{p}^{(t-1)}(\mathbf{x}_i)$ 作为监督信号. 该候选集中似然项定义如下:

$$L_{\text{intra}}^{(t)} = -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^K p_j^{(t-1)}(\mathbf{x}_i) \log \left(\phi_j^{(t)}(\mathbf{x}_i) \right), \quad (3)$$

其中 $\phi_j^{(t)}(\mathbf{x}_i)$ 表示模型当前原始输出经 softmax 函数转换后的归一化结果. 集中伪标记分布 $\mathbf{p}^{(0)}(\mathbf{x}_i)$ 根据样本的候选标记集合 S_i 进行初始化, 其第 j 个元素为

$$p_j^{(0)}(\mathbf{x}_i) = \begin{cases} 1/|S_i|, & j \in S_i, \\ 0, & \text{其他}. \end{cases} \quad (4)$$

该似然项通过集中式归一化使预测概率质量局限于候选标记内部, 同时采用上一轮更新得到的集中伪标记分布作为监督, 有效避免监督信号随模型瞬时波动而不稳定, 使高可信度标记在迭代过程中逐步增强, 噪声标记逐步被抑制.

尽管集中伪标记分布在迭代过程中能够逐步强化候选集合中的高可信度标记, 但其本质上是一种由模型预测驱动的递归式更新机制, 因此容易受到局部估计偏差的影响. 由于似然项作用于 softmax 概率分布, 其内部固有的竞争结构会放大微小的预测扰动, 从而使逐轮更新得到的伪标记分布出现结构性波动, 导致整体优化路径不够稳定. 为缓解上述问题, 我们在损失函数中进一步引入平方损失, 用于持续对齐模型输出与更新后的集中伪标记分布, 使伪标记的演化过程更加平滑且可控. 该项能够有效抑制监督漂移带来的累积误差, 使去噪过程在逐步强化可信标记的同时保持优化稳定性, 从而为模型在训练阶段提供连续、可靠的对齐信号. 设模型的原始输出 $\mathbf{z}(\mathbf{x}_i)$ 经 sigmoid 函数后得到预测置信度向量为 $\boldsymbol{\delta}(\mathbf{x}_i) \in [0, 1]^K$, 第 t 轮迭代中的平方损失定义为

$$L_{\text{align}}^{(t)} = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^K \left(\delta_j^{(t)}(\mathbf{x}_i) - p_j^{(t-1)}(\mathbf{x}_i) \right)^2, \quad (5)$$

其中 $\delta_j^{(t)}(\mathbf{x}_i)$ 表示第 j 个维度的 sigmoid 预测概率.

前述两项损失主要作用在候选标记部分, 但对于任意一个 PML 样本 $\mathbf{x}_i$ 来说, 非候选标记同样蕴含着天然可靠的排他性信息. 与候选标记集合同时包含真实标记与噪声标记不同, 非候选标记必然不属于样本的真实标

记, 其激活应被显式压制. 若在训练过程中忽略这类排他信息, 模型可能在高维标记空间中产生不必要的输出激活, 从而增加噪声扩散的风险. 基于上述考虑, 本文在损失函数中进一步引入排他约束惩罚项, 用以利用非候选标记所携带的明确负反馈来约束模型输出. 受 Gao 等^[34]的方法启发, 该惩罚项通过对非候选标记维度施加负对数似然约束, 使模型在这些维度上保持低置信度, 从而有效抑制无关标记的激活. 与前述两项损失不同的是, 非候选标记提供的监督虽然较弱, 但该信息完全无噪, 因此可作为稳定可靠的负监督信息源. 在第 t 轮迭代中, 排他约束惩罚项定义为

$$L_{\text{excl}}^{(t)} = -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^K \tilde{s}_{ij} \log(1 - \delta_j^{(t)}(\mathbf{x}_i)), \quad (6)$$

其中 $\tilde{s}_{ij}$ 为非候选标记向量 $\tilde{\mathbf{s}} = \mathbf{1} - \mathbf{s}_i$ 的第 j 个元素, 若标记 $j \notin S_i$, 则 $\tilde{s}_{ij} = 1$, 否则为 0. 该机制通过抑制非候选标记的预测概率, 减少无关标记在预测分布中的占比, 从而限制其对候选标记判别过程的干扰, 使模型能够同时从候选集合内部的渐进式去噪与非候选区域的排他监督中获益, 从而形成互补的双重去噪效果. 综上, 排他约束惩罚项作为候选集合之外的稳定负监督, 与前两部分损失共同构成 DeNo 的完整优化目标, 其在第 t 轮的总损失定义为

$$L^{(t)} = L_{\text{align}}^{(t)} + \alpha L_{\text{intra}}^{(t)} + \beta L_{\text{excl}}^{(t)}, \quad (7)$$

其中 α 和 β 分别控制 $L_{\text{intra}}^{(t)}$ 和 $L_{\text{excl}}^{(t)}$ 的权重. 上述三项损失协同作用, L_{align} 提供稳定对齐信号、 L_{intra} 实现候选内部的递进式消歧、 L_{excl} 利用非候选标记的排他性进一步强化去噪能力, 从而实现渐进式的标记消歧过程, 方法的完整流程如算法 1 所示. 值得注意的是, 参数 θ 在不同模型中对应不同形式, 在线性模型中指线性模型的权重矩阵, 在多层神经网络模型中指所有层的可学习参数集合.

算法 1 DeNo 算法流程

输入: 偏多标记训练数据集 $\bar{D}$, 迭代次数 T , 权衡因子 α 和 β , 优化算法 $\mathcal{A}$;

输出: 模型参数 θ ;

```

1: 根据式 (4) 初始化集中伪标记分布  $\mathbf{p}^{(0)}$ ;
2: for  $t = 1$  到  $T$  do
3:   根据式 (7) 计算经验风险 (损失值)  $L^{(t)}$ ;
4:   设置梯度为  $-\nabla_{\theta} L^{(t)}$ ;
5:   根据式 (2) 更新集中伪标记分布  $\mathbf{p}^{(t)}$ ;
6:   根据优化算法  $\mathcal{A}$  更新模型参数  $\theta$ ;
7: end for
```

4 实验

本节开展一系列实验来验证所提方法 DeNo 的有效性. 本文采用 MLL 中常用的四类评价指标 ranking loss、one error、coverage 和 average precision 来评估本文方法及各对比方法的性能, 评价指标的具体定义可参考文献 [7]. 其中, average precision 值越高表明性能越优, 而其余三个评价指标的值越小则代表方法表现越优. 所有实验均基于 PyTorch^[45] 实现, 并在 NVIDIA RTX 3090 GPU 上运行.

4.1 实验设置

数据集与预处理方法. 本文选取 3 个真实世界 PML 数据集¹⁾和 11 个常用 MLL 数据集²⁾开展实验, 其详细统计信息见表 1. 其中, 前三个数据集为真实世界 PML 数据集, 包含来自网络用户的不准确标注以及经人工

1) 下载链接: <https://palm.seu.edu.cn/zhangml/>.

2) 下载链接: <https://mulan.sourceforge.net/datasets-mlc.html>.

表 1 实验数据集具体特征.
Table 1 The details of datasets.

Dataset	Examples	Features	Labels	Domain
YeastCC	6139	6139	50	Biology
YeastBP	6139	6139	217	Biology
Music_style	6839	98	10	Music
scene	2407	294	6	Images
yeast	2417	103	14	Biology
bookmark	87856	2150	208	Text
delicious	16105	500	983	Text
Corel16k	11153	500	153	Images
tmc2007	28596	981	22	Text
rcv1subset1	5815	944	101	Text
rcv1subset2	5252	944	101	Text
rcv1subset3	5410	944	101	Text
rcv1subset4	5761	944	101	Text
rcv1subset5	5532	944	101	Text

校验的真实标注. 为构建合成 PML 数据集, 本文参照文献 [12, 13], 在 MLL 数据集上注入随机噪声, 从而满足 PML 的实验设定. 具体而言, 按照先前研究 [12, 13] 的设定, 对于类别数量超过 100 的 MLL 数据集, 仅保留出现频率最高的前 15 个类别, 并剔除包含其他类别标记的样本. 随后, 在生成任意样本的候选标记集合时, 随机翻转其无关标记, 直至候选集合中无关标记的数量达到真实标记数量的 $\gamma\%$. 为系统探索本文所提出方法在不同噪声含量下的鲁棒性, 本文通过调节噪声比例 γ 控制候选标记集合的噪声含量, 其取值设置为 $\{100, 150, 200, 250\}$.

对比方法. 为验证 DeNo 的有效性, 本文选取 PML 相关领域中具有代表性的六种方法作为对比对象, 并按照其所属研究领域对其进行分组介绍, 具体如下.

- PML 方法: 本文选择三种典型的 PML 方法作为对比, fpml 方法 [11]、PARD 方法 [13] 和 PASE 方法 [41]. fpml 方法是一种特征驱动的消歧方法, 通过在特征空间中建模标记的低秩结构以缓解候选标记中的噪声干扰, 并在此基础上构建特征诱导的判别模型以实现消歧. PARD 方法则基于概率图模型, 通过刻画候选标记的生成过程来恢复潜在的真实标记, 从而实现标记消歧. PASE 方法是一种基于标记特征修正的消歧方法, 通过为每个候选标记学习独立的特征表示空间来刻画其判别性模式, 并在此基础上进行特征层面的修正与去噪, 从而实现标记消歧与真实标记识别.

- 补标记学习 (complementary label learning) 方法: 不同于 PML 中每个样本对应候选标记集合, 在补标记学习任务中, 每个样本对应一个补标记集合, 即样本所不属于的标记类别集合. 本文采用两种典型的补标记学习算法: GDF 方法和 $\bar{\mathcal{L}}_{\text{BCE}}$ 方法 [34], 这类方法将样本的非候选标记作为补标记进行训练. 具体来说, $\bar{\mathcal{L}}_{\text{BCE}}$ 方法是基于 Gao 等 [34] 推导的无偏风险估计器, 对传统二元交叉熵 (binary cross-entropy) 损失进行修正, 使其能够适用于补标记情形. 然而, 该无偏损失在部分情况下会导致梯度无界, GDF 方法则通过梯度友好损失函数来避免这一问题, 从而提高在训练过程中的稳定性.

- MLL 方法: CCMN 方法 [46] 是一个通用的多标记噪声学习框架, 能够在类条件噪声假设下对多种噪声类型进行统一建模与鲁棒学习. 该方法通过显式估计噪声转移矩阵来刻画真实标记与观测标记之间的映射关系, 并将噪声建模过程与分类器训练解耦, 以有效求解多标记噪声学习问题.

参数设置. 对于基于神经网络的方法, 本文采用线性分类模型与多层感知机 (multi-layer perceptron, MLP) 作为分类器. 其中, MLP 采用 $d-512-K$ 结构, 隐层单元数为 512, dropout 率为 0.2. 除模型结构外, 线性分类器与 MLP 的其余实验设置保持一致. 本文采用随机梯度下降 (stochastic gradient descent, SGD) 作为优化器, 动量

表 2 DeNo 超参设置.
Table 2 The hyper-parameters of DeNo.

Dataset	α	β
YeastCC	0.1	1
YeastBP	10	0.3
Music_style	0.8	0.8
scene	10	2
bookmark	2	2
tmc2007	10	10
delicious	10	5
yeast	10	10
Corel16k	5	5
rcv1subset1	0.3	0.8
rcv1subset2	2	0.8
rcv1subset3	2	1
rcv1subset4	2	2
rcv1subset5	2	5

设为 0.9, 训练批次大小为 256, 训练迭代次数为 200. 所有数据集的权重衰减均设为 10^{-3} , 学习率根据数据集特性从 $\{10^{-3}, 5 \times 10^{-2}, 10^{-2}, 10^{-1}\}$ 中选择, 并按照文献 [34] 的设置在第 100 轮与第 150 轮分别降至原来的 10%. 数据集划分方面, 首先随机选取 10% 的样本作为验证集用于超参数选择, 其余 90% 的数据按 9:1 的比例再划分为训练集与测试集, 分别用于模型训练与性能评估. 其中, 训练集仅包含候选标记, 而验证集与测试集使用真实标记. 对于每个数据集重复上述划分过程 10 次, 并统计 10 组实验的平均结果, 以获得具有统计意义的性能估计. 对于本文损失中的权衡因子 α 和 β , 其搜索范围均设定为 $\{0.1, 0.3, 0.5, 0.8, 1, 2, 5, 10\}$, 最终选取的具体参数如表 2 所示.

4.2 实验结果与分析

整体实验结果分析. 表 3 汇总了线性分类模型在 3 个真实世界 PML 数据集上的实验结果. 表 4 和 5 分别报告了在线性模型作为基分类器时, 11 个数据集在不同候选标记噪声比例下, 各方法在 ranking loss 和 average precision 两项指标上的表现 (one error 和 coverage 的结果详见附录 A). 表 6 汇总了在显著性水平 0.05 下, DeNo 与各对比方法在四项评价指标上的成对 t 检验“胜/平/负”统计情况.

从表 3 的实验结果可以看出, DeNo 在不同评价指标上均表现出稳定性能: 在 one error 指标上取得最优或接近最优结果, 在 coverage 和 ranking loss 等指标上整体优于 GDF 和 $\bar{\mathcal{L}}_{\text{BCE}}$ 方法, 同时在 average precision 指标上保持具有竞争力的表现. 上述结果表明, 在真实偏多标记数据分布下, DeNo 仍能够有效抑制噪声干扰并保持良好的泛化性能. 由表 4 和 5 的结果可以观察到, DeNo 在绝大多数数据集与噪声设置下均取得了显著优于所有对比方法的性能表现. 该结果充分验证了本文提出的分布集中机制、排他性约束以及持续对齐信号在复杂噪声环境下的有效性, 使模型能够更稳定地识别候选集合中潜在的真实标记. 下面重点从 ranking loss 与 average precision 两个指标展开分析, 并结合表 6 的统计结果对 DeNo 的整体性能表现进行归纳与讨论.

首先, 在 ranking loss 指标上, DeNo 在所有数据集上均获得最优或并列最优结果. 这说明 DeNo 通过集中式伪标记分布在候选标记内部强化高预测概率标记, 能够有效抑制噪声标记带来的排序结构破坏. 随着候选集合噪声比例 $\gamma\%$ 从 100% 提升到 250%, fpml、PASE、GDF、CCMN 等方法性能出现明显退化, 而 DeNo 在大部分数据集上的性能下降幅度较小, 展现出更强的噪声鲁棒性. 特别是在高噪声比例为 200/250 (%) 时, DeNo 的 ranking loss 结果较次优方法平均降低 20%~40%, 说明引入分布集中机制与排他性约束能够有效压缩噪声空间, 使模型保持稳定的消歧能力. 与对比方法中表现最优的 PARD 方法相比, DeNo 的 average precision 平均提升

表 3 真实世界 PML 数据集上线性分类模型的实验结果 (均值 $\pm$ 标准差). 最优结果加粗表示, 其中 $\bullet/\circ$ 分别表示在成对 t 检验 (显著性水平 0.05) 下, DeNo 显著优于/劣于相应对比方法.

Table 3 Results (mean $\pm$ std) on the real-world PML datasets with the linear classification model. The best performance is highlighted in boldface, where $\bullet/\circ$ indicates whether DeNo is superior/inferior to baselines with pairwise t -test (at the 0.05 significance level).

Dataset	fpml	PARD	PASE	GDF	$\tilde{\mathcal{L}}_{\text{BCE}}$	CCMN	DeNo
One error $\downarrow$							
YeastCC	0.943 $\pm$ 0.030 $\bullet$	0.791 $\pm$ 0.025	0.779 $\pm$ 0.021	0.958 $\pm$ 0.010 $\bullet$	0.962 $\pm$ 0.009 $\bullet$	0.936 $\pm$ 0.027 $\bullet$	0.774$\pm$0.014
YeastBP	0.924 $\pm$ 0.054 $\bullet$	0.751 $\pm$ 0.023 $\bullet$	0.717$\pm$0.017	0.958 $\pm$ 0.006 $\bullet$	0.962 $\pm$ 0.009 $\bullet$	0.950 $\pm$ 0.017 $\bullet$	0.740 $\pm$ 0.030
Music_style	0.457 $\pm$ 0.164 $\bullet$	0.373 $\pm$ 0.034 $\bullet$	0.333 $\pm$ 0.006 $\bullet$	0.934 $\pm$ 0.010 $\bullet$	0.934 $\pm$ 0.010 $\bullet$	0.923 $\pm$ 0.006 $\bullet$	0.309$\pm$0.012
Coverage $\downarrow$							
YeastCC	0.495 $\pm$ 0.185 $\bullet$	0.113 $\pm$ 0.017	0.084$\pm$0.018	0.300 $\pm$ 0.060 $\bullet$	0.305 $\pm$ 0.057 $\bullet$	0.284 $\pm$ 0.044 $\bullet$	0.113 $\pm$ 0.007
YeastBP	0.251 $\pm$ 0.105 $\bullet$	0.335 $\pm$ 0.016 $\bullet$	0.246$\pm$0.008$\circ$	0.684 $\pm$ 0.073 $\bullet$	0.696 $\pm$ 0.070 $\bullet$	0.693 $\pm$ 0.066 $\bullet$	0.299 $\pm$ 0.012
Music_style	0.379 $\pm$ 0.019 $\bullet$	0.265 $\pm$ 0.017 $\bullet$	0.195$\pm$0.008	0.700 $\pm$ 0.011 $\bullet$	0.700 $\pm$ 0.011 $\bullet$	0.633 $\pm$ 0.011 $\bullet$	0.214 $\pm$ 0.014
Ranking loss $\downarrow$							
YeastCC	0.499 $\pm$ 0.054 $\bullet$	0.187 $\pm$ 0.014 $\bullet$	0.143$\pm$0.020	0.462 $\pm$ 0.015 $\bullet$	0.477 $\pm$ 0.006 $\bullet$	0.436 $\pm$ 0.037 $\bullet$	0.171 $\pm$ 0.016
YeastBP	0.483 $\pm$ 0.080 $\bullet$	0.236 $\pm$ 0.013 $\bullet$	0.187$\pm$0.006	0.477 $\pm$ 0.005 $\bullet$	0.494 $\pm$ 0.010 $\bullet$	0.485 $\pm$ 0.006 $\bullet$	0.205 $\pm$ 0.007
Music_style	0.308 $\pm$ 0.021 $\bullet$	0.198 $\pm$ 0.018 $\bullet$	0.133$\pm$0.006	0.664 $\pm$ 0.003 $\bullet$	0.664 $\pm$ 0.003 $\bullet$	0.622 $\pm$ 0.009 $\bullet$	0.142 $\pm$ 0.012
Average precision $\uparrow$							
YeastCC	0.295 $\pm$ 0.012 $\bullet$	0.532 $\pm$ 0.026 $\bullet$	0.587$\pm$0.012	0.149 $\pm$ 0.015 $\bullet$	0.142 $\pm$ 0.008 $\bullet$	0.189 $\pm$ 0.032 $\bullet$	0.563 $\pm$ 0.022
YeastBP	0.179 $\pm$ 0.037 $\bullet$	0.304 $\pm$ 0.018 $\bullet$	0.372$\pm$0.011$\circ$	0.068 $\pm$ 0.004 $\bullet$	0.065 $\pm$ 0.004 $\bullet$	0.071 $\pm$ 0.004 $\bullet$	0.335 $\pm$ 0.019
Music_style	0.589 $\pm$ 0.066 $\bullet$	0.698 $\pm$ 0.021 $\bullet$	0.748$\pm$0.004$\circ$	0.259 $\pm$ 0.004 $\bullet$	0.259 $\pm$ 0.004 $\bullet$	0.264 $\pm$ 0.005 $\bullet$	0.707 $\pm$ 0.009

3%~8%. 同时, 随着噪声比例增大, DeNo 的性能下降幅度显著小于 GDF, 这进一步体现了集中式伪标记分布的稳定监督作用, 以及平方损失带来的平滑对齐机制能够抑制噪声累积.

此外, DeNo 在 ranking loss 与 average precision 两个指标上稳定超过 fpml 与 PARD, 说明基于预测概率建模的 PML 方法在复杂噪声下存在固有局限, 而 DeNo 通过结合排他信息与渐进式去噪策略能够更有效地实现标记消歧. 与 GDF 和 $\tilde{\mathcal{L}}_{\text{BCE}}$ 相比, DeNo 在 rcv1 系列数据集上取得了更加明显的优势, 表明其在噪声结构更复杂的数据上具备更强的适应性, 也印证了集中式伪标记分布的可靠性. 最后, 从表 6 的“胜/平/负”汇总结果来看, DeNo 在四个指标上的成对 t 检验共获得 1055 次胜利, 仅有 7 次失败, 充分展示了本文所提方法在不同噪声水平和数据分布下的稳定性与一致性. 综上, DeNo 之所以能够在广泛的实验设置下稳定优于现有方法, 得益于其三大关键策略的协同作用: 候选集合内部的分布集中去噪、排他性负标记信息挖掘以及持续对齐的稳定优化信号, 这些设计共同确保模型能够在高噪声条件下依然维持准确的标记辨识能力与优越的泛化性能.

消融实验. 为验证 DeNo 框架中各核心组件的有效性, 本文在噪声比例为 200% 的设置下开展了消融实验. 通过对框架的核心结构进行拆解与重组, 共构建 7 种退化版本, 并与完整 DeNo 框架 (最后一行) 进行对比. 在实验中, “ $\checkmark$ ” 表示该版本保留对应组件; 其中, $L_{\text{align w/o } p}$ 表示该损失项去除了集中式伪标记分布, 此时 $L_{\text{align}} = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^K (\delta_j^{(t)}(\mathbf{x}_i) - s_{ij})^2$. 从表 7 可见, 任意组件的移除都会导致性能下降, 且退化趋势在不同指标与数据集之间高度一致, 说明 DeNo 的三个模块具有明显的互补性和不可替代性.

具体而言, 仅保留排他约束 L_{excl} 时, 性能显著下降, 表明仅依赖非候选标记提供的排他性信息无法有效求解 PML 问题. 仅保留候选内集中似然损失 L_{intra} 时, 尽管在 scene 等部分数据集上仍保持一定性能, 但在 delicious、Corel16k 等复杂数据集上退化明显, 说明缺乏分布对齐或排他约束会导致性能不稳定. 对于仅保留分布对齐项 L_{align} 的版本, 性能进一步下降, 表明单纯依赖分布对齐无法防止噪声扩散. 对于任意两项损失的组合, 其性能仍显著劣于完整模型: 缺少 L_{intra} 会削弱真实标记的排序优势; 缺少 L_{align} 会导致优化过程波动; 缺少 L_{excl} 则降低对无关标记的约束, 使其更易被错误激活, 从而影响整体效果. 在保留三项损失但去除集中式伪标记分布的版本中, 由于缺乏有效的对齐目标, 其在四项指标上的性能均有所下降, 进一步证明集中式伪标记分布是

表 4 关于 ranking loss 评价指标在线性分类模型上的实验结果 (均值 $\pm$ 标准差). 最优结果加粗表示, 其中符号 $\bullet/\circ$ 分别表示在成对 t 检验 (显著性水平 0.05) 下, DeNo 显著优于/劣于相应对比方法.

Table 4 Experimental results on ranking loss (mean $\pm$ std) with the linear classification model. The best performance for each dataset is shown in boldface, where $\bullet/\circ$ indicates whether DeNo is superior/inferior to baselines with pairwise t -test (at the 0.05 significance level).

Dataset	$\gamma\%$	fpml	PARD	PASE	GDF	$\tilde{\mathcal{L}}_{\text{BCE}}$	CCMN	DeNo
scene	100	0.471 $\pm$ 0.036 $\bullet$	0.099 $\pm$ 0.009 $\bullet$	0.118 $\pm$ 0.034 $\bullet$	0.139 $\pm$ 0.016 $\bullet$	0.174 $\pm$ 0.018 $\bullet$	0.412 $\pm$ 0.034 $\bullet$	0.083$\pm$0.011
	150	0.474 $\pm$ 0.031 $\bullet$	0.102 $\pm$ 0.013 $\bullet$	0.116 $\pm$ 0.018 $\bullet$	0.142 $\pm$ 0.009 $\bullet$	0.171 $\pm$ 0.011 $\bullet$	0.404 $\pm$ 0.038 $\bullet$	0.083$\pm$0.010
	200	0.471 $\pm$ 0.036 $\bullet$	0.099 $\pm$ 0.011 $\bullet$	0.118 $\pm$ 0.034 $\bullet$	0.139 $\pm$ 0.016 $\bullet$	0.174 $\pm$ 0.016 $\bullet$	0.412 $\pm$ 0.034 $\bullet$	0.083$\pm$0.011
	250	0.485 $\pm$ 0.030 $\bullet$	0.123 $\pm$ 0.034 $\bullet$	0.148 $\pm$ 0.014 $\bullet$	0.145 $\pm$ 0.012 $\bullet$	0.189 $\pm$ 0.007 $\bullet$	0.433 $\pm$ 0.015 $\bullet$	0.091$\pm$0.008
yeast	100	0.212 $\pm$ 0.010 $\bullet$	0.269 $\pm$ 0.035 $\bullet$	0.221 $\pm$ 0.006 $\bullet$	0.224 $\pm$ 0.014 $\bullet$	0.291 $\pm$ 0.018 $\bullet$	0.224 $\pm$ 0.013 $\bullet$	0.189$\pm$0.012
	150	0.218 $\pm$ 0.011 $\bullet$	0.287 $\pm$ 0.064 $\bullet$	0.225 $\pm$ 0.010 $\bullet$	0.220 $\pm$ 0.013 $\bullet$	0.274 $\pm$ 0.016 $\bullet$	0.234 $\pm$ 0.016 $\bullet$	0.199$\pm$0.014
	200	0.225 $\pm$ 0.128 $\bullet$	0.308 $\pm$ 0.043 $\bullet$	0.219 $\pm$ 0.004 $\bullet$	0.224 $\pm$ 0.014 $\bullet$	0.291 $\pm$ 0.018 $\bullet$	0.242 $\pm$ 0.014 $\bullet$	0.204$\pm$0.014
	250	0.356 $\pm$ 0.022 $\bullet$	0.312 $\pm$ 0.033 $\bullet$	0.276$\pm$0.039	0.320 $\pm$ 0.018 $\bullet$	0.447 $\pm$ 0.018 $\bullet$	0.298 $\pm$ 0.018 $\bullet$	0.309 $\pm$ 0.017
bookmark	100	0.403 $\pm$ 0.009 $\bullet$	0.123 $\pm$ 0.048 $\bullet$	0.112 $\pm$ 0.004 $\bullet$	0.194 $\pm$ 0.004 $\bullet$	0.274 $\pm$ 0.017 $\bullet$	0.192 $\pm$ 0.004 $\bullet$	0.094$\pm$0.005
	150	0.396 $\pm$ 0.009 $\bullet$	0.129 $\pm$ 0.010 $\bullet$	0.115 $\pm$ 0.004 $\bullet$	0.196 $\pm$ 0.010 $\bullet$	0.275 $\pm$ 0.016 $\bullet$	0.214 $\pm$ 0.012 $\bullet$	0.095$\pm$0.009
	200	0.401 $\pm$ 0.010 $\bullet$	0.124 $\pm$ 0.013 $\bullet$	0.112 $\pm$ 0.004 $\bullet$	0.194 $\pm$ 0.008 $\bullet$	0.274 $\pm$ 0.014 $\bullet$	0.191 $\pm$ 0.013 $\bullet$	0.094$\pm$0.008
	250	0.391 $\pm$ 0.012 $\bullet$	0.148 $\pm$ 0.014 $\bullet$	0.081$\pm$0.008	0.196 $\pm$ 0.009 $\bullet$	0.266 $\pm$ 0.015 $\bullet$	0.249 $\pm$ 0.016 $\bullet$	0.097 $\pm$ 0.005
delicious	100	0.381 $\pm$ 0.010 $\bullet$	0.273 $\pm$ 0.012 $\bullet$	0.281 $\pm$ 0.009 $\bullet$	0.299 $\pm$ 0.011 $\bullet$	0.351 $\pm$ 0.016 $\bullet$	0.424 $\pm$ 0.011 $\bullet$	0.260$\pm$0.034
	150	0.394 $\pm$ 0.011 $\bullet$	0.303 $\pm$ 0.014 $\bullet$	0.306 $\pm$ 0.007 $\bullet$	0.322 $\pm$ 0.009 $\bullet$	0.387 $\pm$ 0.015 $\bullet$	0.423 $\pm$ 0.012 $\bullet$	0.276$\pm$0.020
	200	0.411 $\pm$ 0.012 $\bullet$	0.335 $\pm$ 0.013 $\bullet$	0.327 $\pm$ 0.008 $\bullet$	0.344 $\pm$ 0.010 $\bullet$	0.403 $\pm$ 0.014 $\bullet$	0.443 $\pm$ 0.011 $\bullet$	0.296$\pm$0.010
	250	0.421 $\pm$ 0.013 $\bullet$	0.325 $\pm$ 0.012 $\bullet$	0.344 $\pm$ 0.006 $\bullet$	0.340 $\pm$ 0.012 $\bullet$	0.406 $\pm$ 0.016 $\bullet$	0.454 $\pm$ 0.013 $\bullet$	0.310$\pm$0.012
Corel16k	100	0.310 $\pm$ 0.061 $\bullet$	0.253 $\pm$ 0.007 $\bullet$	0.269 $\pm$ 0.012 $\bullet$	0.312 $\pm$ 0.011 $\bullet$	0.410 $\pm$ 0.015 $\bullet$	0.354 $\pm$ 0.013 $\bullet$	0.234$\pm$0.011
	150	0.307 $\pm$ 0.011 $\bullet$	0.271 $\pm$ 0.015 $\bullet$	0.283 $\pm$ 0.066 $\bullet$	0.312 $\pm$ 0.012 $\bullet$	0.410 $\pm$ 0.014 $\bullet$	0.367 $\pm$ 0.012 $\bullet$	0.239$\pm$0.012
	200	0.308 $\pm$ 0.012 $\bullet$	0.286 $\pm$ 0.013 $\bullet$	0.293 $\pm$ 0.007 $\bullet$	0.313 $\pm$ 0.011 $\bullet$	0.401 $\pm$ 0.015 $\bullet$	0.382 $\pm$ 0.014 $\bullet$	0.247$\pm$0.011
	250	0.306 $\pm$ 0.013 $\bullet$	0.342 $\pm$ 0.012 $\bullet$	0.295 $\pm$ 0.006 $\bullet$	0.309 $\pm$ 0.013 $\bullet$	0.407 $\pm$ 0.016 $\bullet$	0.376 $\pm$ 0.013 $\bullet$	0.247$\pm$0.013
tmc2007	100	0.183 $\pm$ 0.009 $\bullet$	0.059 $\pm$ 0.010 $\bullet$	0.062 $\pm$ 0.012 $\bullet$	0.142 $\pm$ 0.008 $\bullet$	0.288 $\pm$ 0.014 $\bullet$	0.138 $\pm$ 0.011 $\bullet$	0.059$\pm$0.009
	150	0.183 $\pm$ 0.010 $\bullet$	0.065 $\pm$ 0.011 $\bullet$	0.067 $\pm$ 0.004 $\bullet$	0.137 $\pm$ 0.009 $\bullet$	0.287 $\pm$ 0.013 $\bullet$	0.140 $\pm$ 0.012 $\bullet$	0.059$\pm$0.010
	200	0.183 $\pm$ 0.011 $\bullet$	0.075 $\pm$ 0.012 $\bullet$	0.080 $\pm$ 0.007 $\bullet$	0.143 $\pm$ 0.010 $\bullet$	0.292 $\pm$ 0.012 $\bullet$	0.138 $\pm$ 0.013 $\bullet$	0.068$\pm$0.010
	250	0.185 $\pm$ 0.012 $\bullet$	0.077 $\pm$ 0.013 $\bullet$	0.083 $\pm$ 0.001 $\bullet$	0.137 $\pm$ 0.011 $\bullet$	0.289 $\pm$ 0.014 $\bullet$	0.139 $\pm$ 0.014 $\bullet$	0.068$\pm$0.011
rcv1subset1	100	0.375 $\pm$ 0.007 $\bullet$	0.217 $\pm$ 0.009 $\bullet$	0.168 $\pm$ 0.012 $\bullet$	0.248 $\pm$ 0.008 $\bullet$	0.354 $\pm$ 0.011 $\bullet$	0.219 $\pm$ 0.010 $\bullet$	0.090$\pm$0.007
	150	0.375 $\pm$ 0.008 $\bullet$	0.095 $\pm$ 0.010 $\bullet$	0.229 $\pm$ 0.002 $\bullet$	0.264 $\pm$ 0.009 $\bullet$	0.351 $\pm$ 0.012 $\bullet$	0.262 $\pm$ 0.011 $\bullet$	0.083$\pm$0.008
	200	0.368 $\pm$ 0.009 $\bullet$	0.253 $\pm$ 0.034 $\bullet$	0.272 $\pm$ 0.057 $\bullet$	0.281 $\pm$ 0.084 $\bullet$	0.344 $\pm$ 0.038 $\bullet$	0.153 $\pm$ 0.015 $\bullet$	0.084$\pm$0.023
	250	0.374 $\pm$ 0.013 $\bullet$	0.275 $\pm$ 0.024 $\bullet$	0.265 $\pm$ 0.012 $\bullet$	0.275 $\pm$ 0.033 $\bullet$	0.351 $\pm$ 0.045 $\bullet$	0.232 $\pm$ 0.055 $\bullet$	0.108$\pm$0.010
rcv1subset2	100	0.329 $\pm$ 0.008 $\bullet$	0.196 $\pm$ 0.010 $\bullet$	0.177 $\pm$ 0.012 $\bullet$	0.248 $\pm$ 0.009 $\bullet$	0.351 $\pm$ 0.012 $\bullet$	0.185 $\pm$ 0.011 $\bullet$	0.077$\pm$0.008
	150	0.335 $\pm$ 0.009 $\bullet$	0.094 $\pm$ 0.011 $\bullet$	0.200 $\pm$ 0.007 $\bullet$	0.235 $\pm$ 0.010 $\bullet$	0.340 $\pm$ 0.013 $\bullet$	0.180 $\pm$ 0.012 $\bullet$	0.086$\pm$0.009
	200	0.324 $\pm$ 0.003 $\bullet$	0.254 $\pm$ 0.044 $\bullet$	0.227 $\pm$ 0.046 $\bullet$	0.269 $\pm$ 0.089 $\bullet$	0.374 $\pm$ 0.039 $\bullet$	0.167 $\pm$ 0.034 $\bullet$	0.084$\pm$0.003
	250	0.329 $\pm$ 0.011 $\bullet$	0.229 $\pm$ 0.030 $\bullet$	0.250 $\pm$ 0.055 $\bullet$	0.258 $\pm$ 0.024 $\bullet$	0.366 $\pm$ 0.033 $\bullet$	0.198 $\pm$ 0.046 $\bullet$	0.085$\pm$0.034
rcv1subset3	100	0.333 $\pm$ 0.023 $\bullet$	0.199 $\pm$ 0.009 $\bullet$	0.206 $\pm$ 0.027 $\bullet$	0.292 $\pm$ 0.100 $\bullet$	0.362 $\pm$ 0.049 $\bullet$	0.198 $\pm$ 0.005 $\bullet$	0.085$\pm$0.009
	150	0.335 $\pm$ 0.023 $\bullet$	0.098 $\pm$ 0.009 $\bullet$	0.215 $\pm$ 0.031 $\bullet$	0.285 $\pm$ 0.048 $\bullet$	0.372 $\pm$ 0.014 $\bullet$	0.181 $\pm$ 0.013 $\bullet$	0.076$\pm$0.010
	200	0.337 $\pm$ 0.054 $\bullet$	0.237 $\pm$ 0.064 $\bullet$	0.216 $\pm$ 0.091 $\bullet$	0.259 $\pm$ 0.030 $\bullet$	0.370 $\pm$ 0.028 $\bullet$	0.194 $\pm$ 0.042 $\bullet$	0.113$\pm$0.031
	250	0.331 $\pm$ 0.039 $\bullet$	0.268 $\pm$ 0.074 $\bullet$	0.250 $\pm$ 0.040 $\bullet$	0.262 $\pm$ 0.064 $\bullet$	0.371 $\pm$ 0.024 $\bullet$	0.189 $\pm$ 0.055 $\bullet$	0.101$\pm$0.012
rcv1subset4	100	0.319 $\pm$ 0.009 $\bullet$	0.167 $\pm$ 0.011 $\bullet$	0.164 $\pm$ 0.012 $\bullet$	0.271 $\pm$ 0.010 $\bullet$	0.361 $\pm$ 0.013 $\bullet$	0.208 $\pm$ 0.012 $\bullet$	0.091$\pm$0.009
	150	0.305 $\pm$ 0.010 $\bullet$	0.087 $\pm$ 0.012 $\bullet$	0.202 $\pm$ 0.012 $\bullet$	0.288 $\pm$ 0.011 $\bullet$	0.355 $\pm$ 0.014 $\bullet$	0.189 $\pm$ 0.013 $\bullet$	0.092$\pm$0.010
	200	0.336 $\pm$ 0.011 $\bullet$	0.237 $\pm$ 0.028 $\bullet$	0.216 $\pm$ 0.034 $\bullet$	0.259 $\pm$ 0.017 $\bullet$	0.370 $\pm$ 0.015 $\bullet$	0.195 $\pm$ 0.029 $\bullet$	0.091$\pm$0.022
	250	0.306 $\pm$ 0.037 $\bullet$	0.222 $\pm$ 0.030 $\bullet$	0.179 $\pm$ 0.012 $\bullet$	0.250 $\pm$ 0.018 $\bullet$	0.340 $\pm$ 0.029 $\bullet$	0.181 $\pm$ 0.027 $\bullet$	0.123$\pm$0.027
rcv1subset5	100	0.305 $\pm$ 0.010 $\bullet$	0.185 $\pm$ 0.012 $\bullet$	0.189 $\pm$ 0.012 $\bullet$	0.240 $\pm$ 0.011 $\bullet$	0.366 $\pm$ 0.014 $\bullet$	0.246 $\pm$ 0.013 $\bullet$	0.088$\pm$0.010
	150	0.298 $\pm$ 0.020 $\bullet$	0.093 $\pm$ 0.005 $\bullet$	0.170 $\pm$ 0.039 $\bullet$	0.246 $\pm$ 0.032 $\bullet$	0.350 $\pm$ 0.032 $\bullet$	0.217 $\pm$ 0.037 $\bullet$	0.088$\pm$0.009
	200	0.304 $\pm$ 0.012 $\bullet$	0.207 $\pm$ 0.033 $\bullet$	0.217 $\pm$ 0.038 $\bullet$	0.244 $\pm$ 0.017 $\bullet$	0.336 $\pm$ 0.057 $\bullet$	0.217 $\pm$ 0.041 $\bullet$	0.100$\pm$0.012
	250	0.286 $\pm$ 0.013 $\bullet$	0.192 $\pm$ 0.021 $\bullet$	0.217 $\pm$ 0.029 $\bullet$	0.255 $\pm$ 0.028 $\bullet$	0.357 $\pm$ 0.040 $\bullet$	0.214 $\pm$ 0.523 $\bullet$	0.105$\pm$0.013

对齐机制有效性的关键. 整体来看, 完整的 DeNo 框架在全部四项指标上均取得最佳表现, 表明通过分布集中式去噪、持续对齐、排他约束三者的协同, DeNo 实现了对噪声标记的有效抑制以及对真实标记的稳定识别.

表 5 关于 average precision 评价指标在线性分类模型上的实验结果 (均值 $\pm$ 标准差). 最优结果加粗表示, 其中符号 $\bullet/\circ$ 分别表示在成对 t 检验 (显著性水平 0.05) 下, DeNo 显著优于/劣于相应对比方法.

Table 5 Experimental results on average precision (mean $\pm$ std) with the linear classification model. The best performance for each dataset is shown in boldface, where $\bullet/\circ$ indicates whether DeNo is superior/inferior to baselines with pairwise t -test (at the 0.05 significance level).

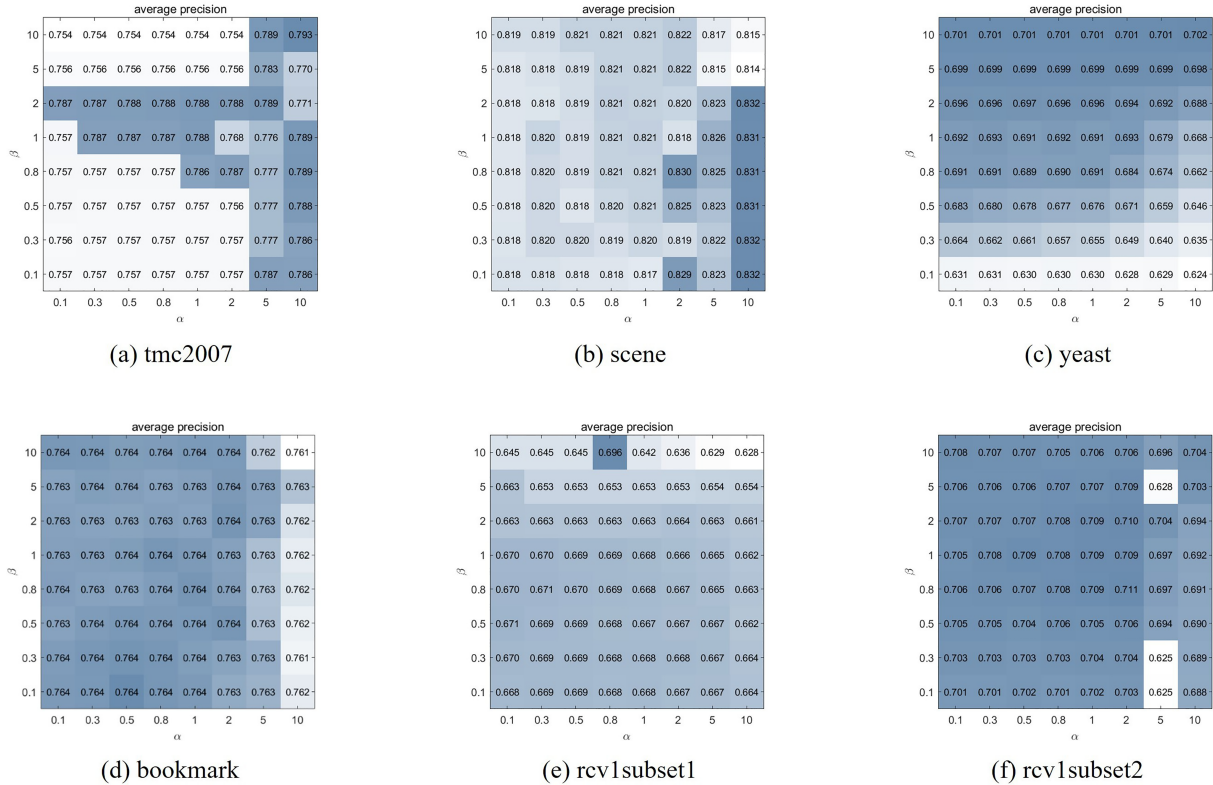
Dataset	$\gamma\%$	fpml	PARD	PASE	GDF	$\tilde{\mathcal{L}}_{\text{BCE}}$	CCMN	DeNo
scene	100	0.438 $\pm$ 0.023 $\bullet$	0.833 $\pm$ 0.015 $\bullet$	0.822 $\pm$ 0.027 $\bullet$	0.778 $\pm$ 0.019 $\bullet$	0.735 $\pm$ 0.018 $\bullet$	0.503 $\pm$ 0.020 $\bullet$	0.856$\pm$0.019
	150	0.437 $\pm$ 0.018 $\bullet$	0.826 $\pm$ 0.022 $\bullet$	0.815 $\pm$ 0.013 $\bullet$	0.774 $\pm$ 0.011 $\bullet$	0.734 $\pm$ 0.010 $\bullet$	0.505 $\pm$ 0.021 $\bullet$	0.856$\pm$0.018
	200	0.438 $\pm$ 0.023 $\bullet$	0.832 $\pm$ 0.017 $\bullet$	0.822 $\pm$ 0.003 $\bullet$	0.777 $\pm$ 0.022 $\bullet$	0.735 $\pm$ 0.018 $\bullet$	0.503 $\pm$ 0.020 $\bullet$	0.856$\pm$0.019
	250	0.429 $\pm$ 0.023 $\bullet$	0.799 $\pm$ 0.010 $\bullet$	0.778 $\pm$ 0.028 $\bullet$	0.774 $\pm$ 0.018 $\bullet$	0.721 $\pm$ 0.012 $\bullet$	0.496 $\pm$ 0.015 $\bullet$	0.841$\pm$0.013
yeast	100	0.704 $\pm$ 0.020 $\bullet$	0.689 $\pm$ 0.021 $\bullet$	0.698 $\pm$ 0.009 $\bullet$	0.705 $\pm$ 0.019 $\bullet$	0.623 $\pm$ 0.017 $\bullet$	0.699 $\pm$ 0.018 $\bullet$	0.743$\pm$0.017
	150	0.702 $\pm$ 0.020 $\bullet$	0.677 $\pm$ 0.044 $\bullet$	0.696 $\pm$ 0.076 $\bullet$	0.715 $\pm$ 0.020 $\bullet$	0.651 $\pm$ 0.016 $\bullet$	0.696 $\pm$ 0.022 $\bullet$	0.733$\pm$0.017
	200	0.686 $\pm$ 0.017 $\bullet$	0.661 $\pm$ 0.032 $\bullet$	0.696 $\pm$ 0.009 $\bullet$	0.705 $\pm$ 0.019 $\bullet$	0.623 $\pm$ 0.017 $\bullet$	0.685 $\pm$ 0.020 $\bullet$	0.721$\pm$0.014
	250	0.529 $\pm$ 0.016 $\bullet$	0.633$\pm$0.035$\circ$	0.637 $\pm$ 0.039 $\bullet$	0.555 $\pm$ 0.018 $\bullet$	0.467 $\pm$ 0.017 $\bullet$	0.614 $\pm$ 0.023 $\circ$	0.574 $\pm$ 0.014
bookmark	100	0.334 $\pm$ 0.008 $\bullet$	0.725 $\pm$ 0.006 $\bullet$	0.739 $\pm$ 0.003 $\bullet$	0.621 $\pm$ 0.005 $\bullet$	0.514 $\pm$ 0.012 $\bullet$	0.606 $\pm$ 0.010 $\bullet$	0.764$\pm$0.007
	150	0.336 $\pm$ 0.008 $\bullet$	0.718 $\pm$ 0.006 $\bullet$	0.738 $\pm$ 0.055 $\bullet$	0.617 $\pm$ 0.006 $\bullet$	0.510 $\pm$ 0.006 $\bullet$	0.562 $\pm$ 0.029 $\bullet$	0.764$\pm$0.008
	200	0.335 $\pm$ 0.007 $\bullet$	0.724 $\pm$ 0.006 $\bullet$	0.739 $\pm$ 0.004 $\bullet$	0.621 $\pm$ 0.005 $\bullet$	0.514 $\pm$ 0.012 $\bullet$	0.606 $\pm$ 0.010 $\bullet$	0.764$\pm$0.007
	250	0.341 $\pm$ 0.009 $\bullet$	0.698 $\pm$ 0.007 $\bullet$	0.713 $\pm$ 0.006 $\bullet$	0.618 $\pm$ 0.005 $\bullet$	0.527 $\pm$ 0.009 $\bullet$	0.527 $\pm$ 0.014 $\bullet$	0.760$\pm$0.008
delicious	100	0.499 $\pm$ 0.007 $\bullet$	0.621 $\pm$ 0.007 $\bullet$	0.609 $\pm$ 0.012 $\bullet$	0.589 $\pm$ 0.004 $\bullet$	0.533 $\pm$ 0.005 $\bullet$	0.474 $\pm$ 0.009 $\bullet$	0.633$\pm$0.004
	150	0.487 $\pm$ 0.009 $\bullet$	0.599 $\pm$ 0.012 $\bullet$	0.584 $\pm$ 0.003 $\bullet$	0.567 $\pm$ 0.005 $\bullet$	0.498 $\pm$ 0.003 $\bullet$	0.476 $\pm$ 0.052 $\bullet$	0.616$\pm$0.004
	200	0.469 $\pm$ 0.007 $\bullet$	0.563 $\pm$ 0.018 $\bullet$	0.558 $\pm$ 0.007 $\bullet$	0.551 $\pm$ 0.008 $\bullet$	0.486 $\pm$ 0.005 $\bullet$	0.459 $\pm$ 0.009 $\bullet$	0.596$\pm$0.004
	250	0.462 $\pm$ 0.010 $\bullet$	0.463 $\pm$ 0.013 $\bullet$	0.543 $\pm$ 0.010 $\bullet$	0.556 $\pm$ 0.008 $\bullet$	0.486 $\pm$ 0.007 $\bullet$	0.452 $\pm$ 0.004 $\bullet$	0.582$\pm$0.004
Corel16k	100	0.422 $\pm$ 0.055 $\bullet$	0.503 $\pm$ 0.009	0.465 $\pm$ 0.019 $\bullet$	0.427 $\pm$ 0.041 $\bullet$	0.338 $\pm$ 0.021 $\bullet$	0.397 $\pm$ 0.023 $\bullet$	0.516$\pm$0.037
	150	0.425 $\pm$ 0.055 $\bullet$	0.492 $\pm$ 0.008	0.460 $\pm$ 0.058 $\bullet$	0.427 $\pm$ 0.041 $\bullet$	0.338 $\pm$ 0.020 $\bullet$	0.393 $\pm$ 0.016 $\bullet$	0.514$\pm$0.035
	200	0.424 $\pm$ 0.054 $\bullet$	0.472 $\pm$ 0.012 $\bullet$	0.443 $\pm$ 0.016 $\bullet$	0.430 $\pm$ 0.032 $\bullet$	0.344 $\pm$ 0.012 $\bullet$	0.387 $\pm$ 0.025 $\bullet$	0.506$\pm$0.040
	250	0.427 $\pm$ 0.057 $\bullet$	0.557$\pm$0.011$\circ$	0.429 $\pm$ 0.006 $\bullet$	0.435 $\pm$ 0.045 $\bullet$	0.337 $\pm$ 0.015 $\bullet$	0.377 $\pm$ 0.033 $\bullet$	0.505$\pm$0.036
tmc2007	100	0.557 $\pm$ 0.071 $\bullet$	0.799 $\pm$ 0.003	0.791 $\pm$ 0.020	0.667 $\pm$ 0.067 $\bullet$	0.388 $\pm$ 0.023 $\bullet$	0.626 $\pm$ 0.077 $\bullet$	0.800$\pm$0.055
	150	0.557 $\pm$ 0.071 $\bullet$	0.792 $\pm$ 0.004	0.792 $\pm$ 0.007	0.671 $\pm$ 0.068 $\bullet$	0.385 $\pm$ 0.020 $\bullet$	0.624 $\pm$ 0.075 $\bullet$	0.798$\pm$0.056
	200	0.557 $\pm$ 0.071 $\bullet$	0.780 $\pm$ 0.004	0.774 $\pm$ 0.003	0.661 $\pm$ 0.064 $\bullet$	0.384 $\pm$ 0.019 $\bullet$	0.622 $\pm$ 0.074 $\bullet$	0.788$\pm$0.054
	250	0.556 $\pm$ 0.070 $\bullet$	0.774 $\pm$ 0.003	0.772 $\pm$ 0.003	0.673 $\pm$ 0.069 $\bullet$	0.386 $\pm$ 0.022 $\bullet$	0.624 $\pm$ 0.075 $\bullet$	0.786$\pm$0.051
rcv1subset1	100	0.277 $\pm$ 0.024 $\bullet$	0.589 $\pm$ 0.020 $\bullet$	0.582 $\pm$ 0.003 $\bullet$	0.483 $\pm$ 0.025 $\bullet$	0.388 $\pm$ 0.041 $\bullet$	0.559 $\pm$ 0.081 $\bullet$	0.774$\pm$0.068
	150	0.291 $\pm$ 0.018 $\bullet$	0.701 $\pm$ 0.014 $\bullet$	0.517 $\pm$ 0.053 $\bullet$	0.484 $\pm$ 0.067 $\bullet$	0.375 $\pm$ 0.041 $\bullet$	0.543 $\pm$ 0.068 $\bullet$	0.783$\pm$0.046
	200	0.294 $\pm$ 0.031 $\bullet$	0.537 $\pm$ 0.026 $\bullet$	0.458 $\pm$ 0.078 $\bullet$	0.440 $\pm$ 0.023 $\bullet$	0.393 $\pm$ 0.051 $\bullet$	0.568 $\pm$ 0.027 $\bullet$	0.785$\pm$0.038
	250	0.300 $\pm$ 0.022 $\bullet$	0.526 $\pm$ 0.017 $\bullet$	0.469 $\pm$ 0.003 $\bullet$	0.449 $\pm$ 0.056 $\bullet$	0.379 $\pm$ 0.043 $\bullet$	0.574 $\pm$ 0.030 $\bullet$	0.795$\pm$0.029
rcv1subset2	100	0.339 $\pm$ 0.021 $\bullet$	0.624 $\pm$ 0.033 $\bullet$	0.600 $\pm$ 0.035 $\bullet$	0.471 $\pm$ 0.023 $\bullet$	0.377 $\pm$ 0.016 $\bullet$	0.629 $\pm$ 0.033 $\bullet$	0.789$\pm$0.068
	150	0.332 $\pm$ 0.015 $\bullet$	0.718 $\pm$ 0.018 $\bullet$	0.571 $\pm$ 0.061 $\bullet$	0.502 $\pm$ 0.015 $\bullet$	0.394 $\pm$ 0.048 $\bullet$	0.626 $\pm$ 0.029 $\bullet$	0.777$\pm$0.076
	200	0.368 $\pm$ 0.056 $\bullet$	0.565 $\pm$ 0.017 $\bullet$	0.540 $\pm$ 0.061 $\bullet$	0.468 $\pm$ 0.093 $\bullet$	0.359 $\pm$ 0.025 $\bullet$	0.630 $\pm$ 0.058 $\bullet$	0.815$\pm$0.058
	250	0.346 $\pm$ 0.032 $\bullet$	0.568 $\pm$ 0.030 $\bullet$	0.515 $\pm$ 0.065 $\bullet$	0.470 $\pm$ 0.031 $\bullet$	0.370 $\pm$ 0.037 $\bullet$	0.583 $\pm$ 0.056 $\bullet$	0.818$\pm$0.055
rcv1subset3	100	0.339 $\pm$ 0.065 $\bullet$	0.611 $\pm$ 0.048 $\bullet$	0.565 $\pm$ 0.032 $\bullet$	0.425 $\pm$ 0.018 $\bullet$	0.365 $\pm$ 0.014 $\bullet$	0.620 $\pm$ 0.037 $\bullet$	0.769$\pm$0.103
	150	0.314 $\pm$ 0.033 $\bullet$	0.709 $\pm$ 0.016 $\bullet$	0.548 $\pm$ 0.054 $\bullet$	0.455 $\pm$ 0.096 $\bullet$	0.359 $\pm$ 0.027 $\bullet$	0.613 $\pm$ 0.034 $\bullet$	0.808$\pm$0.064
	200	0.353 $\pm$ 0.084 $\bullet$	0.559 $\pm$ 0.037 $\bullet$	0.551 $\pm$ 0.026 $\bullet$	0.464 $\pm$ 0.076 $\bullet$	0.354 $\pm$ 0.023 $\bullet$	0.579 $\pm$ 0.065 $\bullet$	0.782$\pm$0.059
	250	0.363 $\pm$ 0.086 $\bullet$	0.543 $\pm$ 0.048 $\bullet$	0.497 $\pm$ 0.043 $\bullet$	0.467 $\pm$ 0.058 $\bullet$	0.359 $\pm$ 0.012 $\bullet$	0.607 $\pm$ 0.065 $\bullet$	0.778$\pm$0.063
rcv1subset4	100	0.340 $\pm$ 0.038 $\bullet$	0.680 $\pm$ 0.025 $\bullet$	0.646 $\pm$ 0.007 $\bullet$	0.499 $\pm$ 0.086 $\bullet$	0.368 $\pm$ 0.030 $\bullet$	0.615 $\pm$ 0.056 $\bullet$	0.770$\pm$0.086
	150	0.359 $\pm$ 0.039 $\bullet$	0.752 $\pm$ 0.018 $\bullet$	0.579 $\pm$ 0.038 $\bullet$	0.474 $\pm$ 0.077 $\bullet$	0.372 $\pm$ 0.028 $\bullet$	0.613 $\pm$ 0.031 $\bullet$	0.770$\pm$0.077
	200	0.373 $\pm$ 0.057 $\bullet$	0.634 $\pm$ 0.016 $\bullet$	0.539 $\pm$ 0.069 $\bullet$	0.489 $\pm$ 0.043 $\bullet$	0.382 $\pm$ 0.014 $\bullet$	0.611 $\pm$ 0.031 $\bullet$	0.781$\pm$0.064
	250	0.387 $\pm$ 0.074 $\bullet$	0.617 $\pm$ 0.027 $\bullet$	0.615 $\pm$ 0.045 $\bullet$	0.489 $\pm$ 0.043 $\bullet$	0.382 $\pm$ 0.014 $\bullet$	0.615 $\pm$ 0.037 $\bullet$	0.772$\pm$0.072
rcv1subset5	100	0.384 $\pm$ 0.024 $\bullet$	0.644 $\pm$ 0.030 $\bullet$	0.579 $\pm$ 0.024 $\bullet$	0.494 $\pm$ 0.038 $\bullet$	0.369 $\pm$ 0.033 $\bullet$	0.584 $\pm$ 0.061 $\bullet$	0.777$\pm$0.022
	150	0.394 $\pm$ 0.026 $\bullet$	0.730 $\pm$ 0.013 $\bullet$	0.619 $\pm$ 0.054 $\bullet$	0.471 $\pm$ 0.051 $\bullet$	0.384 $\pm$ 0.054 $\bullet$	0.595 $\pm$ 0.044 $\bullet$	0.777$\pm$0.024
	200	0.391 $\pm$ 0.024 $\bullet$	0.604 $\pm$ 0.019 $\bullet$	0.531 $\pm$ 0.075 $\bullet$	0.540 $\pm$ 0.062 $\bullet$	0.398 $\pm$ 0.078 $\bullet$	0.571 $\pm$ 0.054 $\bullet$	0.782$\pm$0.033
	250	0.409 $\pm$ 0.037 $\bullet$	0.614 $\pm$ 0.016 $\bullet$	0.555 $\pm$ 0.054 $\bullet$	0.520 $\pm$ 0.079 $\bullet$	0.379 $\pm$ 0.055 $\bullet$	0.577 $\pm$ 0.068 $\bullet$	0.774$\pm$0.033

参数敏感性分析. 为评估 DeNo 对超参数 α 与 β 的敏感性, 本文在噪声比例 $\gamma = 200(\%)$ 的设置下, 对所有数据集进行了系统验证. 此处选取 6 个具有代表性的数据集进行展示, 结果如图 2 所示. 图中呈现了不同 α

表 6 方法在线性分类模型上的“胜/平/负”实验结果统计.

Table 6 The statistics of win/tie/lose in the linear classification model.

DeNo vs.	fpml	PARD	PASE	GDF	$\bar{\mathcal{L}}_{\text{BCE}}$	CCMN	Total
One error	44/3/0	43/4/0	42/4/1	45/2/0	47/0/0	43/4/0	264/17/1
Coverage	45/2/0	38/9/0	39/7/1	47/0/0	47/0/0	47/0/0	263/18/1
Ranking loss	47/0/0	38/9/0	38/9/0	47/0/0	47/0/0	46/1/0	263/19/0
Average precision	47/0/0	38/7/2	40/5/2	47/0/0	47/0/0	46/0/1	265/12/5
Total	183/5/0	157/29/2	159/25/4	186/2/0	188/0/0	182/5/1	1055/66/7

图 2 (网络版彩图) 噪声比例 $\gamma = 200$ (%) 下不同超参数 α 、 β 的验证集表现. 颜色越深表示数值越大.**Figure 2** (Color online) Validation performance under different hyperparameter settings of α and β at a noise ratio of $\gamma = 200$ (%). The darker the color, the larger the value.

(候选内集中似然的权重) 与 β (排他约束的权重) 组合下模型在 average precision 指标上的表现. 整体结果显示, DeNo 在较宽范围的超参数配置下均保持稳定性能, 没有出现剧烈波动, 说明模型对超参数具有较好的鲁棒性. 在 α 、 β 取 2~10 的中等偏大区间时, 各数据集的性能普遍较优, 说明候选内可信标记强化与排他性约束在此范围内协同作用最为有效. 对于 bookmark 与 rcv1 系列等稀疏标记数据集, 较大的超参数会带来轻微性能下降, 而在较小或中等范围 (如 $\alpha \in [0.3, 2]$ 、 $\beta \in [0.8, 2]$) 更为稳定. 这表明在标记稀疏的场景中, 过强的正向或负向强化可能导致信息偏置, 适度的权重能带来更好的泛化效果. 基于图 2 的验证结果, 我们进一步整理了不同数据集的最优超参数组合, 如表 2 所示. 表中超参均由验证集性能选取, 能在各基准数据集上获得稳定且可靠的效果.

分布集中机制的去噪性能分析. 为验证本文提出的分布集中机制对于去噪的有效性, 本文展示了在不同噪声比例 γ 下, 某样本的集中式伪标记分布 (图中 “DeNo” 所对应的曲线) 与 sigmoid 输出的动态演化过程, 如图 3 所示. 这里, sigmoid 表示模型原始输出 z 经 sigmoid 函数处理的结果, 并为了便于比较, 对其输出向量进行了归

表 7 在噪声比例为 $\gamma = 200$ (%) 下采用线性分类模型的消融实验结果 (均值 $\pm$ 标准差). 最优结果加粗表示.

Table 7 The experimental results (mean $\pm$ std) of ablation studies with $\gamma=200$ (%) on the linear classification model. The best performance for each dataset is shown in boldface.														
L_{excl}	L_{intra}	L_{align}	$L_{\text{align w/o p}}$	scene	yeast	bookmark	delicious	Corell6k	tmu2007	rev1subset1	rev1subset2	rev1subset3	rev1subset4	rev1subset5
Ranking loss $\downarrow$														
✓				0.441 $\pm$ 0.019	0.217 $\pm$ 0.014	0.207 $\pm$ 0.004	0.353 $\pm$ 0.006	0.306 $\pm$ 0.038	0.195 $\pm$ 0.048	0.254 $\pm$ 0.019	0.270 $\pm$ 0.016	0.287 $\pm$ 0.054	0.243 $\pm$ 0.040	0.227 $\pm$ 0.050
	✓			0.089 $\pm$ 0.010	0.236 $\pm$ 0.014	0.094 $\pm$ 0.004	0.332 $\pm$ 0.004	0.249 $\pm$ 0.039	0.078 $\pm$ 0.024	0.094 $\pm$ 0.014	0.090 $\pm$ 0.030	0.097 $\pm$ 0.028	0.113 $\pm$ 0.017	0.116 $\pm$ 0.009
		✓		0.103 $\pm$ 0.011	0.239 $\pm$ 0.013	0.132 $\pm$ 0.006	0.329 $\pm$ 0.006	0.261 $\pm$ 0.021	0.082 $\pm$ 0.027	0.161 $\pm$ 0.024	0.140 $\pm$ 0.035	0.157 $\pm$ 0.029	0.122 $\pm$ 0.025	0.123 $\pm$ 0.014
✓			✓	0.126 $\pm$ 0.014	0.218 $\pm$ 0.014	0.300 $\pm$ 0.008	0.362 $\pm$ 0.005	0.312 $\pm$ 0.039	0.325 $\pm$ 0.026	0.246 $\pm$ 0.016	0.263 $\pm$ 0.019	0.260 $\pm$ 0.049	0.258 $\pm$ 0.067	0.262 $\pm$ 0.090
✓				0.100 $\pm$ 0.010	0.205 $\pm$ 0.015	0.094 $\pm$ 0.005	0.314 $\pm$ 0.006	0.249 $\pm$ 0.022	0.076 $\pm$ 0.021	0.094 $\pm$ 0.024	0.095 $\pm$ 0.032	0.100 $\pm$ 0.030	0.109 $\pm$ 0.024	0.105 $\pm$ 0.013
✓			✓	0.092 $\pm$ 0.011	0.234 $\pm$ 0.014	0.094 $\pm$ 0.005	0.307 $\pm$ 0.004	0.249 $\pm$ 0.027	0.079 $\pm$ 0.020	0.085 $\pm$ 0.024	0.086 $\pm$ 0.033	0.091 $\pm$ 0.030	0.110 $\pm$ 0.024	0.102 $\pm$ 0.013
✓				0.091 $\pm$ 0.011	0.235 $\pm$ 0.014	0.094 $\pm$ 0.005	0.312 $\pm$ 0.005	0.249 $\pm$ 0.028	0.072 $\pm$ 0.025	0.085 $\pm$ 0.024	0.088 $\pm$ 0.035	0.091 $\pm$ 0.030	0.107 $\pm$ 0.025	0.103 $\pm$ 0.014
✓		✓		0.083$\pm$0.011	0.204$\pm$0.014	0.094 $\pm$ 0.005	0.296$\pm$0.004	0.247$\pm$0.023	0.067$\pm$0.024	0.084$\pm$0.024	0.085$\pm$0.033	0.091 $\pm$ 0.031	0.101$\pm$0.023	0.100$\pm$0.013
One error $\downarrow$														
✓				0.765 $\pm$ 0.016	0.254 $\pm$ 0.025	0.495 $\pm$ 0.010	0.486 $\pm$ 0.013	0.684 $\pm$ 0.051	0.389 $\pm$ 0.107	0.658 $\pm$ 0.080	0.612 $\pm$ 0.042	0.678 $\pm$ 0.066	0.621 $\pm$ 0.210	0.552 $\pm$ 0.205
	✓			0.250 $\pm$ 0.033	0.416 $\pm$ 0.036	0.324 $\pm$ 0.007	0.459 $\pm$ 0.007	0.627 $\pm$ 0.073	0.238 $\pm$ 0.095	0.274 $\pm$ 0.055	0.207 $\pm$ 0.112	0.276 $\pm$ 0.116	0.263 $\pm$ 0.075	0.241 $\pm$ 0.075
		✓		0.300 $\pm$ 0.033	0.522 $\pm$ 0.027	0.446 $\pm$ 0.010	0.469 $\pm$ 0.017	0.643 $\pm$ 0.062	0.323 $\pm$ 0.086	0.754 $\pm$ 0.070	0.603 $\pm$ 0.070	0.621 $\pm$ 0.099	0.515 $\pm$ 0.093	0.551 $\pm$ 0.066
✓			✓	0.353 $\pm$ 0.018	0.250$\pm$0.026	0.657 $\pm$ 0.009	0.540 $\pm$ 0.010	0.700 $\pm$ 0.050	0.740 $\pm$ 0.042	0.639 $\pm$ 0.044	0.629 $\pm$ 0.086	0.646 $\pm$ 0.081	0.608 $\pm$ 0.212	0.563 $\pm$ 0.183
✓				0.250 $\pm$ 0.033	0.348 $\pm$ 0.022	0.324 $\pm$ 0.010	0.437 $\pm$ 0.010	0.629 $\pm$ 0.060	0.232 $\pm$ 0.082	0.269 $\pm$ 0.062	0.192 $\pm$ 0.065	0.270 $\pm$ 0.102	0.260 $\pm$ 0.083	0.242 $\pm$ 0.068
✓			✓	0.244 $\pm$ 0.032	0.465 $\pm$ 0.060	0.323 $\pm$ 0.010	0.432 $\pm$ 0.013	0.625 $\pm$ 0.060	0.236 $\pm$ 0.081	0.269 $\pm$ 0.063	0.189$\pm$0.066	0.272 $\pm$ 0.111	0.260 $\pm$ 0.090	0.246 $\pm$ 0.070
✓				0.246 $\pm$ 0.032	0.376 $\pm$ 0.033	0.323 $\pm$ 0.010	0.443 $\pm$ 0.006	0.624 $\pm$ 0.061	0.231 $\pm$ 0.081	0.269 $\pm$ 0.062	0.191 $\pm$ 0.066	0.270 $\pm$ 0.099	0.262 $\pm$ 0.079	0.237 $\pm$ 0.072
✓		✓		0.237$\pm$0.031	0.258 $\pm$ 0.029	0.314$\pm$0.010	0.427$\pm$0.012	0.617$\pm$0.061	0.222$\pm$0.081	0.259$\pm$0.062	0.190 $\pm$ 0.068	0.264$\pm$0.100	0.255$\pm$0.083	0.236$\pm$0.071
Coverage $\downarrow$														
✓				0.381 $\pm$ 0.016	0.533 $\pm$ 0.025	0.231 $\pm$ 0.005	0.619 $\pm$ 0.007	0.409 $\pm$ 0.037	0.339 $\pm$ 0.020	0.335 $\pm$ 0.026	0.363 $\pm$ 0.032	0.366 $\pm$ 0.030	0.310 $\pm$ 0.015	0.308 $\pm$ 0.066
	✓			0.091 $\pm$ 0.009	0.511 $\pm$ 0.024	0.119 $\pm$ 0.004	0.601 $\pm$ 0.005	0.356 $\pm$ 0.035	0.168 $\pm$ 0.009	0.168 $\pm$ 0.026	0.189 $\pm$ 0.050	0.180 $\pm$ 0.047	0.165 $\pm$ 0.037	0.198 $\pm$ 0.024
		✓		0.101 $\pm$ 0.010	0.496 $\pm$ 0.034	0.154 $\pm$ 0.005	0.588 $\pm$ 0.006	0.364 $\pm$ 0.028	0.172 $\pm$ 0.015	0.241 $\pm$ 0.020	0.218 $\pm$ 0.054	0.232 $\pm$ 0.042	0.182 $\pm$ 0.042	0.199 $\pm$ 0.021
✓			✓	0.120 $\pm$ 0.011	0.532 $\pm$ 0.023	0.318 $\pm$ 0.008	0.627 $\pm$ 0.005	0.413 $\pm$ 0.040	0.465 $\pm$ 0.010	0.322 $\pm$ 0.028	0.350 $\pm$ 0.016	0.342 $\pm$ 0.026	0.324 $\pm$ 0.036	0.339 $\pm$ 0.104
✓				0.091 $\pm$ 0.009	0.496 $\pm$ 0.028	0.118 $\pm$ 0.005	0.585 $\pm$ 0.010	0.356 $\pm$ 0.026	0.167 $\pm$ 0.013	0.157 $\pm$ 0.021	0.183 $\pm$ 0.050	0.171 $\pm$ 0.043	0.171 $\pm$ 0.043	0.187 $\pm$ 0.022
✓			✓	0.094 $\pm$ 0.010	0.504 $\pm$ 0.030	0.118 $\pm$ 0.005	0.582 $\pm$ 0.008	0.355 $\pm$ 0.030	0.162 $\pm$ 0.016	0.156 $\pm$ 0.020	0.174 $\pm$ 0.051	0.173 $\pm$ 0.043	0.173 $\pm$ 0.040	0.183 $\pm$ 0.025
✓				0.093 $\pm$ 0.010	0.530 $\pm$ 0.029	0.117 $\pm$ 0.005	0.585 $\pm$ 0.008	0.354 $\pm$ 0.034	0.166 $\pm$ 0.013	0.159 $\pm$ 0.020	0.176 $\pm$ 0.052	0.165 $\pm$ 0.044	0.169 $\pm$ 0.045	0.185 $\pm$ 0.027
✓		✓		0.085$\pm$0.009	0.491$\pm$0.028	0.117 $\pm$ 0.005	0.575$\pm$0.006	0.353$\pm$0.027	0.159$\pm$0.022	0.157 $\pm$ 0.021	0.174 $\pm$ 0.051	0.163$\pm$0.044	0.164$\pm$0.039	0.181$\pm$0.024
Average precision $\downarrow$														
✓				0.469 $\pm$ 0.013	0.708 $\pm$ 0.019	0.603 $\pm$ 0.007	0.544 $\pm$ 0.007	0.444 $\pm$ 0.039	0.582 $\pm$ 0.073	0.480 $\pm$ 0.032	0.496 $\pm$ 0.014	0.463 $\pm$ 0.051	0.505 $\pm$ 0.099	0.550 $\pm$ 0.126
	✓			0.837 $\pm$ 0.019	0.658 $\pm$ 0.011	0.754 $\pm$ 0.008	0.565 $\pm$ 0.004	0.504 $\pm$ 0.041	0.784 $\pm$ 0.057	0.782 $\pm$ 0.043	0.803 $\pm$ 0.061	0.767 $\pm$ 0.055	0.765 $\pm$ 0.070	0.764 $\pm$ 0.033
		✓		0.820 $\pm$ 0.019	0.644 $\pm$ 0.012	0.674 $\pm$ 0.006	0.563 $\pm$ 0.003	0.489 $\pm$ 0.053	0.730 $\pm$ 0.066	0.530 $\pm$ 0.022	0.613 $\pm$ 0.066	0.585 $\pm$ 0.067	0.654 $\pm$ 0.049	0.645 $\pm$ 0.034
✓			✓	0.787 $\pm$ 0.014	0.708 $\pm$ 0.020	0.472 $\pm$ 0.007	0.520 $\pm$ 0.006	0.434 $\pm$ 0.037	0.356 $\pm$ 0.031	0.488 $\pm$ 0.015	0.489 $\pm$ 0.042	0.489 $\pm$ 0.062	0.499 $\pm$ 0.120	0.525 $\pm$ 0.141
✓				0.847 $\pm$ 0.019	0.719 $\pm$ 0.016	0.764 $\pm$ 0.007	0.579 $\pm$ 0.005	0.505 $\pm$ 0.040	0.786 $\pm$ 0.055	0.785 $\pm$ 0.038	0.805 $\pm$ 0.056	0.780 $\pm$ 0.058	0.779 $\pm$ 0.064	0.775 $\pm$ 0.032
✓			✓	0.848 $\pm$ 0.020	0.657 $\pm$ 0.011	0.764 $\pm$ 0.008	0.585 $\pm$ 0.004	0.505 $\pm$ 0.040	0.784 $\pm$ 0.053	0.784 $\pm$ 0.039	0.805 $\pm$ 0.058	0.779 $\pm$ 0.060	0.779 $\pm$ 0.067	0.774 $\pm$ 0.033
✓				0.808 $\pm$ 0.019	0.667 $\pm$ 0.012	0.754 $\pm$ 0.008	0.576 $\pm$ 0.003	0.505 $\pm$ 0.041	0.786 $\pm$ 0.054	0.784 $\pm$ 0.039	0.802 $\pm$ 0.060	0.766 $\pm$ 0.060	0.770 $\pm$ 0.064	0.773 $\pm$ 0.035
✓		✓		0.856$\pm$0.019	0.721$\pm$0.014	0.764 $\pm$ 0.007	0.596$\pm$0.004	0.506$\pm$0.040	0.788$\pm$0.054	0.785 $\pm$ 0.038	0.815$\pm$0.058	0.782$\pm$0.059	0.781$\pm$0.064	0.782$\pm$0.032

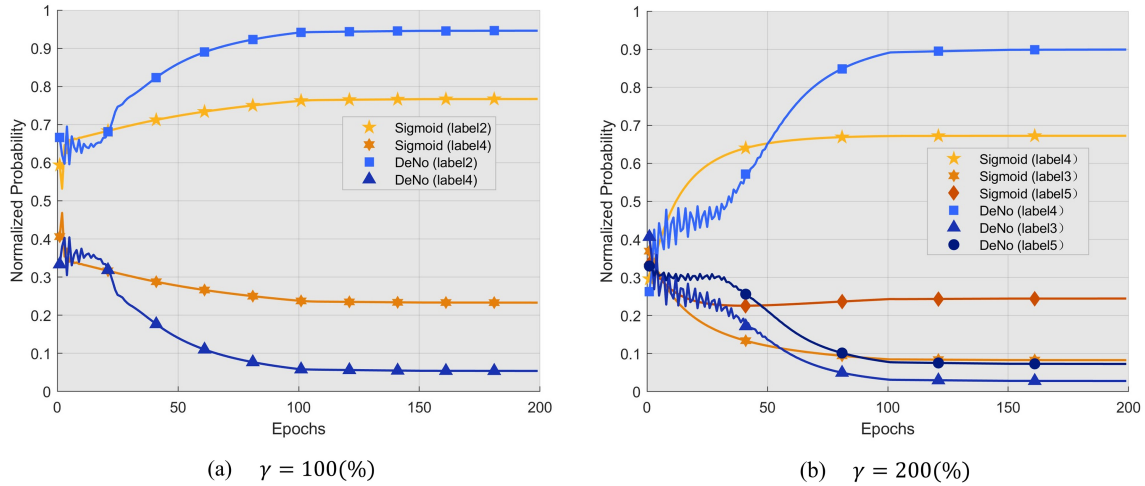


图 3 (网络版彩图) 不同噪声比例 γ 下各方法的噪声抑制表现 (线性分类模型). 图 (a) 对应的样本的真实标记集合为 $\{2\}$, 候选标记集合为 $\{2, 4\}$; 图 (b) 中样本的真实标记集合为 $\{3\}$, 候选标记集合为 $\{3, 4, 5\}$.

Figure 3 (Color online) Noise suppression performance of different methods under varying noise ratios γ (the linear classification model). For (a), the ground-truth label set is $\{2\}$ and the candidate label set is $\{2, 4\}$; for (b), the ground-truth label set is $\{3\}$ and the candidate label set is $\{3, 4, 5\}$.

表 8 在噪声比例 $\gamma = 200(\%)$ 、使用线性分类模型的运行时间 (s).

Table 8 The experimental results of ablation studies with noise rate $\gamma = 200(\%)$ on the linear classification model.

Dataset	PARD	PASE	GDF	$\bar{\mathcal{L}}_{\text{BCE}}$	CCMN	DeNo
YeastCC	90.86	209.09	0.03	43.14	0.00	53.15
YeastBP	145.35	341.05	0.01	74.34	0.00	93.27
Music_style	65.58	75.95	0.05	22.49	0.00	34.83
scene	29.57	116.88	24.69	21.12	21.28	20.62
yeast	20.16	64.27	26.14	21.35	21.39	18.53
bookmark	393.78	1879.79	99.80	100.61	114.42	123.08
delicious	151.04	718.51	43.61	43.27	49.26	54.70
Corel16k	113.43	643.65	39.15	33.47	34.50	48.60
tmc2007	489.04	1532.67	73.31	69.34	88.22	91.91
rcv1subset1	58.00	323.22	31.38	32.39	31.42	36.83
rcv1subset2	54.17	277.31	28.65	28.79	29.65	33.32
rcv1subset3	55.02	400.93	28.85	32.81	30.17	35.98
rcv1subset4	59.36	406.52	29.39	31.76	30.51	34.39
rcv1subset5	54.60	308.62	27.04	26.97	31.60	35.39

一化. 从图 3 可以看到, 分布集中机制在训练过程中能够逐步压制噪声标记并强化真实标记的概率. 具体而言, 在噪声比例 $\gamma = 100(\%)$ 的图 3(a) 中, 样本的真实标记为 2. 集中分布中标记 2 的概率快速上升并趋于稳定, 而噪声标记 4 的概率持续下降; 相比之下, sigmoid 输出在前期出现显著波动, 对噪声的抑制效果较弱, 难以形成可靠的偏好. 该结果表明, 集中分布能够更快且更稳定地聚焦真实标记, 同时有效削弱噪声干扰. 当噪声比例进一步提高至 $\gamma = 200(\%)$ (图 3(b)) 时, 样本的真实标记为 3, 候选标记集合为 $\{3, 4, 5\}$, 噪声比例显著提升. 在这一更具挑战性的场景下, 集中分布依然能够在约 50 轮迭代后收敛到以真实标记 3 为主的分布, 并将标记 4 和 5 的概率压缩到极低水平. 相比之下, sigmoid 输出在高噪声环境中更易受干扰, 对噪声标记长时间维持偏高响应, 难以形成稳定区分. 综上, 分布集中机制通过持续提升真实标记的相对权重并抑制噪声响应, 实现了渐进式的去噪

表 9 在噪声比例为 $\gamma = 100$ (%) 条件下采用 MLP 分类模型的实验结果 (均值 $\pm$ 标准差). 最优结果加粗表示, 其中符号 $\bullet/\circ$ 分别表示在成对 t 检验 (显著性水平 0.05) 下, DeNo 显著优于/劣于相应对比方法.

Table 9 The results (mean $\pm$ std) using MLP classification model with noise rate $\gamma = 100$ (%). The best performance for each dataset is shown in boldface, where $\bullet/\circ$ indicates whether DeNo is superior/inferior to baselines with pairwise t -test (at the 0.05 significance level).

Dataset	GDF	$\tilde{\mathcal{L}}_{\text{BCE}}$	CCMN	DeNo	GDF	$\tilde{\mathcal{L}}_{\text{BCE}}$	CCMN	DeNo
One error $\downarrow$				Coverage $\downarrow$				
scene	0.355 $\pm$ 0.026 $\bullet$	0.343 $\pm$ 0.035 $\bullet$	0.708 $\pm$ 0.063 $\bullet$	0.262$\pm$0.028	0.128 $\pm$ 0.011 $\bullet$	0.124 $\pm$ 0.007 $\bullet$	0.385 $\pm$ 0.005 $\bullet$	0.099$\pm$0.006
yeast	0.243 $\pm$ 0.012	0.244 $\pm$ 0.009	0.246 $\pm$ 0.014 $\bullet$	0.234$\pm$0.052	0.530 $\pm$ 0.016 $\bullet$	0.524 $\pm$ 0.011 $\bullet$	0.554 $\pm$ 0.057 $\bullet$	0.524$\pm$0.020
bookmark	0.526 $\pm$ 0.011 $\bullet$	0.485 $\pm$ 0.006 $\bullet$	0.711 $\pm$ 0.048 $\bullet$	0.385$\pm$0.008	0.230 $\pm$ 0.003 $\bullet$	0.214 $\pm$ 0.007 $\bullet$	0.410 $\pm$ 0.039 $\bullet$	0.167$\pm$0.007
delicious	0.422 $\pm$ 0.012	0.424 $\pm$ 0.020	0.833 $\pm$ 0.004 $\bullet$	0.418$\pm$0.014	0.595 $\pm$ 0.009 $\bullet$	0.594 $\pm$ 0.009 $\bullet$	0.397 $\pm$ 0.008 $\bullet$	0.504$\pm$0.011
Corel16k	0.755 $\pm$ 0.045 $\bullet$	0.713 $\pm$ 0.033	0.557 $\pm$ 0.010 $\bullet$	0.703$\pm$0.032	0.415 $\pm$ 0.056 $\bullet$	0.410 $\pm$ 0.056	0.701 $\pm$ 0.026 $\bullet$	0.402$\pm$0.031
tmc2007	0.382 $\pm$ 0.126 $\bullet$	0.373 $\pm$ 0.125 $\bullet$	0.620 $\pm$ 0.071 $\bullet$	0.327$\pm$0.018	0.260 $\pm$ 0.020	0.261 $\pm$ 0.015	0.393 $\pm$ 0.058 $\bullet$	0.260$\pm$0.011
rcv1subset1	0.742 $\pm$ 0.032 $\bullet$	0.832 $\pm$ 0.025 $\bullet$	0.514 $\pm$ 0.069 $\bullet$	0.459$\pm$0.022	0.442 $\pm$ 0.004 $\bullet$	0.446 $\pm$ 0.010 $\bullet$	0.414 $\pm$ 0.028 $\bullet$	0.179$\pm$0.016
rcv1subset2	0.792 $\pm$ 0.064 $\bullet$	0.808 $\pm$ 0.041 $\bullet$	0.554 $\pm$ 0.078 $\bullet$	0.412$\pm$0.037	0.431 $\pm$ 0.057 $\bullet$	0.475 $\pm$ 0.040 $\bullet$	0.384 $\pm$ 0.049 $\bullet$	0.212$\pm$0.025
rcv1subset3	0.724 $\pm$ 0.051 $\bullet$	0.810 $\pm$ 0.012 $\bullet$	0.544 $\pm$ 0.049 $\bullet$	0.410$\pm$0.033	0.436 $\pm$ 0.028 $\bullet$	0.473 $\pm$ 0.019 $\bullet$	0.358 $\pm$ 0.025 $\bullet$	0.233$\pm$0.012
rcv1subset4	0.780 $\pm$ 0.037 $\bullet$	0.813 $\pm$ 0.031 $\bullet$	0.560 $\pm$ 0.039 $\bullet$	0.419$\pm$0.045	0.452 $\pm$ 0.022 $\bullet$	0.468 $\pm$ 0.020 $\bullet$	0.353 $\pm$ 0.031 $\bullet$	0.240$\pm$0.027
rcv1subset5	0.773 $\pm$ 0.028 $\bullet$	0.845 $\pm$ 0.042 $\bullet$	0.480$\pm$0.120$\circ$	0.613 $\pm$ 0.058	0.382 $\pm$ 0.029 $\bullet$	0.407 $\pm$ 0.024 $\bullet$	0.428 $\pm$ 0.020 $\bullet$	0.348$\pm$0.025
Ranking loss $\downarrow$				Average precision $\uparrow$				
scene	0.135 $\pm$ 0.012 $\bullet$	0.130 $\pm$ 0.007 $\bullet$	0.443 $\pm$ 0.003 $\bullet$	0.100$\pm$0.008	0.781 $\pm$ 0.015 $\bullet$	0.788 $\pm$ 0.017 $\bullet$	0.488 $\pm$ 0.028 $\bullet$	0.837$\pm$0.016
yeast	0.218 $\pm$ 0.012 $\bullet$	0.211 $\pm$ 0.010	0.241 $\pm$ 0.026 $\bullet$	0.203$\pm$0.014	0.719 $\pm$ 0.011	0.724 $\pm$ 0.011	0.687 $\pm$ 0.016 $\bullet$	0.726$\pm$0.025
bookmark	0.209 $\pm$ 0.003 $\bullet$	0.192 $\pm$ 0.007 $\bullet$	0.318 $\pm$ 0.031 $\bullet$	0.143$\pm$0.007	0.589 $\pm$ 0.007 $\bullet$	0.620 $\pm$ 0.007 $\bullet$	0.420 $\pm$ 0.032 $\bullet$	0.696$\pm$0.008
delicious	0.296 $\pm$ 0.008	0.293$\pm$0.008	0.390 $\pm$ 0.006 $\bullet$	0.304 $\pm$ 0.006	0.586 $\pm$ 0.008	0.588$\pm$0.007	0.353 $\pm$ 0.002 $\bullet$	0.586 $\pm$ 0.006
Corel16k	0.322 $\pm$ 0.055	0.314 $\pm$ 0.050	0.414 $\pm$ 0.018 $\bullet$	0.310$\pm$0.021	0.405 $\pm$ 0.038	0.426 $\pm$ 0.036	0.487$\pm$0.010$\circ$	0.407 $\pm$ 0.005
tmc2007	0.150 $\pm$ 0.047	0.149 $\pm$ 0.045	0.303 $\pm$ 0.061 $\bullet$	0.147$\pm$0.007	0.640 $\pm$ 0.092 $\bullet$	0.644 $\pm$ 0.086 $\bullet$	0.483 $\pm$ 0.032 $\bullet$	0.667$\pm$0.005
rcv1subset1	0.373 $\pm$ 0.004 $\bullet$	0.381 $\pm$ 0.014 $\bullet$	0.281 $\pm$ 0.027 $\bullet$	0.107$\pm$0.011	0.377 $\pm$ 0.017 $\bullet$	0.331 $\pm$ 0.017 $\bullet$	0.518 $\pm$ 0.031 $\bullet$	0.682$\pm$0.018
rcv1subset2	0.333 $\pm$ 0.053 $\bullet$	0.380 $\pm$ 0.032 $\bullet$	0.269 $\pm$ 0.051 $\bullet$	0.124$\pm$0.016	0.382 $\pm$ 0.027 $\bullet$	0.351 $\pm$ 0.007 $\bullet$	0.514 $\pm$ 0.040 $\bullet$	0.697$\pm$0.026
rcv1subset3	0.344 $\pm$ 0.028 $\bullet$	0.388 $\pm$ 0.012 $\bullet$	0.253 $\pm$ 0.029 $\bullet$	0.145$\pm$0.013	0.403 $\pm$ 0.029 $\bullet$	0.340 $\pm$ 0.009 $\bullet$	0.519 $\pm$ 0.034 $\bullet$	0.676$\pm$0.017
rcv1subset4	0.377 $\pm$ 0.021 $\bullet$	0.396 $\pm$ 0.024 $\bullet$	0.291 $\pm$ 0.038 $\bullet$	0.168$\pm$0.023	0.362 $\pm$ 0.014 $\bullet$	0.335 $\pm$ 0.020 $\bullet$	0.507 $\pm$ 0.032 $\bullet$	0.660$\pm$0.035
rcv1subset5	0.331 $\pm$ 0.032 $\bullet$	0.352 $\pm$ 0.021 $\bullet$	0.272 $\pm$ 0.053	0.256$\pm$0.017	0.388 $\pm$ 0.028 $\bullet$	0.339 $\pm$ 0.028 $\bullet$	0.486 $\pm$ 0.080	0.502$\pm$0.036

过程, 为后续的似然损失和分布对齐损失提供了更可靠的监督信号.

复杂度分析. 从计算复杂度角度来看, DeNo 在每次迭代中的主要开销包括模型的前向与反向传播, 以及基于类别维度的概率计算与损失评估. 其中, softmax 归一化、候选集合内的重归一化操作和各损失项 (L_{intra} 、 L_{align} 和 L_{excl}) 的计算均在类别维度上进行, 其时间复杂度均为 $\mathcal{O}(K)$. 因此, DeNo 的整体时间复杂度与标准神经网络模型保持在同一量级, 不会引入额外的高阶计算开销. 在空间复杂度方面, 模型仅需维护与类别数 K 成线性关系的预测向量及伪标记分布, 额外存储开销较小. 此外, 本文进一步给出各方法在噪声比例 $\gamma = 200$ (%) 条件下, 使用线性分类模型的实际运行时间 (s). 考虑到不同噪声比例对模型运行时间影响较小, 因此仅报告该代表性设置下的结果. 实验结果基于 200 轮迭代训练, 具体如表 8 所示 (运行时间越短表示效率越高). 需要说明的是, 由于 fpml 方法基于 MATLAB 实现, 与本文基于线性模型的方法在实现环境上存在差异, 直接比较其运行时间不具备公平性, 故不作对比. 从整体来看, DeNo 在大多数数据集上的运行时间与 GDF 和 BCE 处于同一量级, 表明其未引入显著的额外计算开销. 这与前述复杂度分析一致, 即 DeNo 的主要计算过程在类别维度上呈线性增长, 具有良好的可扩展性. 具体而言, 在 scene、yeast 等中小规模数据集上, DeNo 表现出较高的计算效率; 在 bookmark、tmc2007 等大规模数据集上, 尽管运行时间相较部分方法略有增加, 但整体仍处于可接受范围内, 未出现明显的计算开销增长. 这主要源于 DeNo 在训练过程中引入了伪标记更新及多项损失计算, 但相关操作

均为线性复杂度,不涉及高阶计算开销.与 PARD、PASE 等方法相比,DeNo 在所有数据集上均表现出更低的运行时间,进一步验证了其在计算效率方面的优势.

DeNo 在 MLP 模型上的表现.为验证所提出 DeNo 框架在非线性模型中的适用性,本文进一步采用 MLP 作为分类模型进行实验.表 9 给出了在噪声比例为 $\gamma = 100$ (%) 时,不同方法在四个评价指标上的实验结果.可以观察到,在采用 MLP 模型后,DeNo 在大多数数据集和评价指标上仍取得最优或具有竞争力的性能,整体趋势与线性模型下保持一致.特别是在 rcv1 系列及 bookmark 等高维或复杂数据集上,DeNo 在 one error、ranking loss、average precision 等指标上均表现出明显优势.上述结果表明,DeNo 框架并不依赖于线性模型假设,在更具表达能力的非线性模型下同样能够有效发挥其去噪作用,体现出良好的模型结构适应性与泛化能力.

5 结论

在 PML 任务中,候选标记集合同时包含真实标记与噪声标记,这使得现有依赖候选集合内部预测概率建模或特定假设的消歧方法难以在复杂数据环境中保持稳定性能.为缓解这一问题,本文提出一种基于排他约束的去噪偏标记学习框架 DeNo.该框架从三个方面提升模型的消歧能力与抗噪性能.(1)通过分布集中机制在候选集合内部构建归一化的集中伪标记分布,实现对高可信标记的逐轮强化,从而增强真实标记与噪声标记的可分性并形成稳定的渐进式去噪过程.(2)借助平方损失在训练全过程中保持模型输出与更新分布的持续对齐,避免误差累积并提供平滑可靠的优化信号.(3)利用非候选标记所蕴含的排他性信息抑制无关标记激活,进一步挖掘可靠负信号,减少噪声对整体学习过程的干扰.在多个数据集上的实验表明,DeNo 在消除候选集合中的噪声干扰、提升标记判别能力及整体泛化性能方面均取得了显著优势,验证了所提出框架的有效性.

致谢 感谢东南大学大数据计算中心共享服务平台为本文工作提供算力支持.

参考文献

- 1 Boutell M R, Luo J B, Shen X, et al. Learning multi-label scene classification. *Pattern Recognition*, 2004, 37: 1757–1771
- 2 Chen Z M, Cui Q, Zhao B, et al. SST: spatial and semantic transformers for multi-label image recognition. *IEEE Trans Image Process*, 2022, 31: 2570–2583
- 3 Jin B, Muller B, Zhai C, et al. Multi-label literature classification based on the gene ontology graph. *BMC BioInform*, 2008, 9: 525
- 4 Ashburner M, Ball C A, Blake J A, et al. Gene Ontology: tool for the unification of biology. *Nat Genet*, 2000, 25: 25–29
- 5 Zhang Y F, Zhang M L. Generalization analysis for label-specific representation learning. In: *Proceedings of Advances in Neural Information Processing Systems*, 2024. 37: 104904–104933
- 6 Zhang M L, Wu L. Lift: multi-label learning with label-specific features. *IEEE Trans Pattern Anal Mach Intell*, 2015, 37: 107–120
- 7 Zhang M L, Zhou Z H. A review on multi-label learning algorithms. *IEEE Trans Knowl Data Eng*, 2014, 26: 1819–1837
- 8 Zhang Y F, Zhang M L. Generalization analysis for multi-label learning. In: *Proceedings of the 41st International Conference on Machine Learning*, 2024. 60220–60243
- 9 Gao Y, Zhu J Y, Xu M, et al. Multi-label learning with multiple complementary labels. *IEEE Trans Pattern Anal Mach Intell*, 2025, 47: 8013–8024
- 10 Gao Y, Xu M, Zhang M L. Complementary to multiple labels: a correlation-aware correction approach. *IEEE Trans Pattern Anal Mach Intell*, 2024, 46: 9179–9191
- 11 Yu G X, Chen X, Domeniconi C, et al. Feature-induced partial multi-label learning. In: *Proceedings of the 2018 IEEE International Conference on Data Mining*, 2018. 1398–1403
- 12 Xie M K, Huang S J. Partial multi-label learning. In: *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, 2018. 4302–4309
- 13 Hang J Y, Zhang M L. Partial multi-label learning with probabilistic graphical disambiguation. In: *Proceedings of Advances in Neural Information Processing Systems*, 2023. 36: 1339–1351
- 14 Yang F C, Jia Y H, Liu H, et al. Noisy label removal for partial multi-label learning. In: *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2024. 3724–3735

- 15 Tang Y, Gao Y, Luo Y G, et al. Unlearning from weakly supervised learning. In: Proceedings of the 33rd International Joint Conference on Artificial Intelligence, 2024. 5000–5008
- 16 Zhou Z H. A brief introduction to weakly supervised learning. *Natl Sci Rev*, 2018, 5: 44–53
- 17 Mahajan D, Girshick R B, Ramanathan V, et al. Exploring the limits of weakly supervised pretraining. In: Proceedings of the 15th European Conference on Computer Vision, 2018. 185–201
- 18 Liu W W, Wang H B, Shen X B, et al. The Emerging trends of multi-label learning. *IEEE Trans Pattern Anal Mach Intell*, 2022, 44: 7955–7974
- 19 Lv J Q, Liu B, Feng L, et al. On the robustness of average losses for partial-label learning. *IEEE Trans Pattern Anal Mach Intell*, 2023, 46: 2569–2583
- 20 Zhang M L, Fang J P. Partial multi-label learning via credible label elicitation. *IEEE Trans Pattern Anal Mach Intell*, 2021, 43: 3587–3599
- 21 Li F, Shi S F, Wang H Z. Partial multi-label learning via specific label disambiguation. *Knowl-Based Syst*, 2022, 250: 109093
- 22 Wang H B, Liu W W, Zhao Y, et al. Discriminative and correlative partial multi-label learning. In: Proceedings of the 28th International Joint Conference on Artificial Intelligence, 2019. 3691–3697
- 23 Chen Y, Wu Y N, Han N, et al. Partial multi-label learning based on near-far neighborhood label enhancement and nonlinear guidance. In: Proceedings of the 32nd ACM International Conference on Multimedia, 2024. 3722–3731
- 24 Cao N, Zhang T, Jin H. Partial multi-label optimal margin distribution machine. In: Proceedings of the 30th International Joint Conference on Artificial Intelligence, 2021. 2198–2204
- 25 Sun L J, Feng S H, Wang T, et al. Partial multi-label learning by low-rank and sparse decomposition. In: Proceedings of the 33rd AAAI Conference on Artificial Intelligence, 2019. 5016–5023
- 26 Sun L J, Feng S H, Liu J, et al. Global-local label correlation for partial multi-label learning. *IEEE Trans Multim*, 2022, 24: 581–593
- 27 Xie M K, Huang S J. Partial multi-label learning with noisy label identification. *IEEE Trans Pattern Anal Mach Intell*, 2022, 44: 3676–3687
- 28 Zhang M L, Li Y K, Liu X Y, et al. Binary relevance for multi-label learning: an overview. *Front Comput Sci*, 2018, 12: 191–202
- 29 Jia B B, Zhang M L. Multi-dimensional classification via selective feature augmentation. *Mach Intell Res*, 2022, 19: 38–51
- 30 Fürnkranz J, Hüllermeier E, Loza Mencía E, et al. Multilabel classification via calibrated label ranking. *Mach Learn*, 2008, 73: 133–153
- 31 Li Y C, Song Y L, Luo J B. Improving pairwise ranking for multi-label image classification. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017. 1837–1845
- 32 Burkhardt S, Kramer S. Online multi-label dependency topic models for text classification. *Mach Learn*, 2018, 107: 859–886
- 33 Singh I P, Ghorbel E, Oyedotun O, et al. Multi-label image classification using adaptive graph convolutional networks: from a single domain to multiple domains. *Comput Vis Image Underst*, 2024, 247: 104062
- 34 Gao Y, Xu M, Zhang M L. Unbiased risk estimator to multi-labeled complementary label learning. In: Proceedings of the 32nd International Joint Conference on Artificial Intelligence, 2023. 3732–3740
- 35 Gerych W, Hartvigsen T, Buquicchio L, et al. Recurrent Bayesian classifier chains for exact multi-label classification. In: Proceedings of Advances in Neural Information Processing Systems, 2021. 34: 15981–15992
- 36 Lyu G Y, Feng S H, Li Y D. Partial multi-label learning via probabilistic graph matching mechanism. In: Proceedings of the 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2020. 105–113
- 37 Xu N, Liu Y P, Zhang Y, et al. Progressive enhancement of label distributions for partial multilabel learning. *IEEE Trans Neural Netw Learn Syst*, 2021, 34: 4856–4867
- 38 Gong X W, Yuan D, Bao W. Partial multi-label learning via large margin nearest neighbour embeddings. In: Proceedings of the 36th AAAI Conference on Artificial Intelligence, 2022. 6729–6736
- 39 He S, Deng K, Li L, et al. Discriminatively relabel for partial multi-label learning. In: Proceedings of the 2019 IEEE International Conference on Data Mining, 2019. 280–288
- 40 Lyu G Y, Feng S H, Jin Y, et al. Prior knowledge regularized self-representation model for partial multilabel learning. *IEEE Trans Cybern*, 2023, 53: 1618–1628
- 41 Hang J Y, Zhang M L. Partial multi-label learning via label-specific feature corrections. *Sci China Inf Sci*, 2025, 68: 132104
- 42 Laine S, Aila T. Temporal ensembling for semi-supervised learning. In: Proceedings of the 5th International Conference on Learning Representations, 2017
- 43 Tarvainen A, Valpola H. Mean teachers are better role models: weight-averaged consistency targets improve semi-supervised deep learning results. In: Proceedings of Advances in Neural Information Processing Systems, 2017. 30: 1195–1204
- 44 Zhang B W, Wang Y D, Hou W X, et al. Flexmatch: boosting semi-supervised learning with curriculum pseudo labeling. In: Proceedings of Advances in Neural Information Processing Systems, 2021. 34: 18408–18419

- 45 Paszke A, Gross S, Massa F, et al. PyTorch: an imperative style, high-performance deep learning library. In: Proceedings of Advances in Neural Information Processing Systems, 2019. 32: 8024–8035
- 46 Xie M K, Huang S J. CCMN: a general framework for learning with class-conditional multi-label noise. IEEE Trans Pattern Anal Mach Intell, 2023, 45: 154–166

附录 A 整体实验结果补充

表 A1 和 A2 分别汇报了使用线性分类模型在 11 个数据集上, 各方法在 one error 和 coverage 指标下的表现. 整体来看, DeNo 在多数实验设置中均取得最优结果, 充分体现其对候选标记集合中噪声标记的鲁棒性. 从表 A1 可见, DeNo 在绝大部分数据集

表 A1 关于 one error 评价指标在线性分类模型上的实验结果 (均值 $\pm$ 标准差). 最优结果加粗表示, 其中符号 $\bullet/\circ$ 分别表示在成对 t 检验 (显著性水平 0.05) 下, DeNo 显著优于/劣于相应对比方法.

Table A1 Experimental results of one error (mean $\pm$ std) on the linear model. The best performance for each dataset is shown in boldface, where $\bullet/\circ$ indicates whether DeNo is superior/inferior to baselines with pairwise t -test (at the 0.05 significance level).

Dataset	$\gamma\%$	fpml	PARD	PASE	GDF	$\tilde{\mathcal{L}}_{\text{BCE}}$	CCMN	DeNo
scene	100	0.794 $\pm$ 0.030 $\bullet$	0.276 $\pm$ 0.028 $\bullet$	0.285 $\pm$ 0.032 $\bullet$	0.359 $\pm$ 0.004 $\bullet$	0.417 $\pm$ 0.032 $\bullet$	0.711 $\pm$ 0.031 $\bullet$	0.238$\pm$0.031
	150	0.796 $\pm$ 0.016 $\bullet$	0.287 $\pm$ 0.035 $\bullet$	0.298 $\pm$ 0.017 $\bullet$	0.362 $\pm$ 0.020 $\bullet$	0.422 $\pm$ 0.015 $\bullet$	0.711 $\pm$ 0.024 $\bullet$	0.238$\pm$0.031
	200	0.794 $\pm$ 0.030 $\bullet$	0.277 $\pm$ 0.031 $\bullet$	0.285 $\pm$ 0.033 $\bullet$	0.359 $\pm$ 0.036 $\bullet$	0.418 $\pm$ 0.032 $\bullet$	0.711 $\pm$ 0.031 $\bullet$	0.238$\pm$0.031
	250	0.806 $\pm$ 0.031 $\bullet$	0.327 $\pm$ 0.017 $\bullet$	0.349 $\pm$ 0.049 $\bullet$	0.359 $\pm$ 0.030 $\bullet$	0.434 $\pm$ 0.023 $\bullet$	0.708 $\pm$ 0.019 $\bullet$	0.265$\pm$0.022
yeast	100	0.250 $\pm$ 0.026	0.244 $\pm$ 0.031	0.247 $\pm$ 0.020 $\bullet$	0.253 $\pm$ 0.025 $\bullet$	0.371 $\pm$ 0.021 $\bullet$	0.249 $\pm$ 0.025	0.229$\pm$0.021
	150	0.251 $\pm$ 0.026	0.248$\pm$0.031	0.249 $\pm$ 0.029	0.256 $\pm$ 0.026 $\bullet$	0.333 $\pm$ 0.029 $\bullet$	0.250 $\pm$ 0.025 $\bullet$	0.250 $\pm$ 0.032
	200	0.266 $\pm$ 0.021	0.266 $\pm$ 0.040 $\bullet$	0.249$\pm$0.020	0.253 $\pm$ 0.025	0.371 $\pm$ 0.021 $\bullet$	0.251 $\pm$ 0.027	0.258 $\pm$ 0.029
	250	0.608 $\pm$ 0.035 $\bullet$	0.342 $\pm$ 0.139 $\bullet$	0.317$\pm$0.136$\circ$	0.566 $\pm$ 0.056	0.634 $\pm$ 0.035 $\bullet$	0.310 $\pm$ 0.072 $\bullet$	0.560 $\pm$ 0.026
bookmark	100	0.826 $\pm$ 0.008 $\bullet$	0.356 $\pm$ 0.008 $\bullet$	0.340 $\pm$ 0.020 $\bullet$	0.480 $\pm$ 0.007 $\bullet$	0.610 $\pm$ 0.015 $\bullet$	0.515 $\pm$ 0.019 $\bullet$	0.314$\pm$0.010
	150	0.826 $\pm$ 0.008 $\bullet$	0.361 $\pm$ 0.008 $\bullet$	0.341 $\pm$ 0.006 $\bullet$	0.484 $\pm$ 0.008 $\bullet$	0.616 $\pm$ 0.012 $\bullet$	0.576 $\pm$ 0.033 $\bullet$	0.313$\pm$0.010
	200	0.826 $\pm$ 0.008 $\bullet$	0.357 $\pm$ 0.008 $\bullet$	0.340 $\pm$ 0.003 $\bullet$	0.480 $\pm$ 0.007 $\bullet$	0.610 $\pm$ 0.015 $\bullet$	0.515 $\pm$ 0.019 $\bullet$	0.314$\pm$0.010
	250	0.826 $\pm$ 0.008 $\bullet$	0.380 $\pm$ 0.011 $\bullet$	0.369 $\pm$ 0.006 $\bullet$	0.481 $\pm$ 0.006 $\bullet$	0.592 $\pm$ 0.012 $\bullet$	0.603 $\pm$ 0.022 $\bullet$	0.317$\pm$0.009
delicious	100	0.587 $\pm$ 0.037 $\bullet$	0.385 $\pm$ 0.013 $\bullet$	0.397 $\pm$ 0.003 $\bullet$	0.417 $\pm$ 0.012 $\bullet$	0.506 $\pm$ 0.012 $\bullet$	0.596 $\pm$ 0.017 $\bullet$	0.368$\pm$0.009
	150	0.615 $\pm$ 0.037 $\bullet$	0.403 $\pm$ 0.013 $\bullet$	0.430 $\pm$ 0.014 $\bullet$	0.455 $\pm$ 0.009 $\bullet$	0.578 $\pm$ 0.007 $\bullet$	0.595 $\pm$ 0.024 $\bullet$	0.401$\pm$0.007
	200	0.661 $\pm$ 0.030 $\bullet$	0.460 $\pm$ 0.034 $\bullet$	0.481 $\pm$ 0.018 $\bullet$	0.474 $\pm$ 0.014 $\bullet$	0.590 $\pm$ 0.013 $\bullet$	0.624 $\pm$ 0.013 $\bullet$	0.427$\pm$0.012
	250	0.673 $\pm$ 0.031 $\bullet$	0.648 $\pm$ 0.021 $\bullet$	0.509 $\pm$ 0.019 $\bullet$	0.462 $\pm$ 0.011 $\bullet$	0.583 $\pm$ 0.016 $\bullet$	0.630 $\pm$ 0.013 $\bullet$	0.451$\pm$0.011
Corel16k	100	0.714 $\pm$ 0.063 $\bullet$	0.612 $\pm$ 0.014	0.679 $\pm$ 0.020 $\bullet$	0.704 $\pm$ 0.054 $\bullet$	0.796 $\pm$ 0.025 $\bullet$	0.736 $\pm$ 0.038 $\bullet$	0.604$\pm$0.058
	150	0.714 $\pm$ 0.064 $\bullet$	0.619 $\pm$ 0.012 $\bullet$	0.680 $\pm$ 0.015 $\bullet$	0.704 $\pm$ 0.054 $\bullet$	0.796 $\pm$ 0.025 $\bullet$	0.733 $\pm$ 0.029 $\bullet$	0.605$\pm$0.054
	200	0.714 $\pm$ 0.063 $\bullet$	0.637 $\pm$ 0.017 $\bullet$	0.708 $\pm$ 0.028 $\bullet$	0.704 $\pm$ 0.037 $\bullet$	0.798 $\pm$ 0.019 $\bullet$	0.729 $\pm$ 0.047 $\bullet$	0.617$\pm$0.061
	250	0.714 $\pm$ 0.063 $\bullet$	0.464 $\pm$ 0.012 $\bullet$	0.714 $\pm$ 0.011 $\bullet$	0.700 $\pm$ 0.052 $\bullet$	0.805 $\pm$ 0.015 $\bullet$	0.750 $\pm$ 0.052 $\bullet$	0.618$\pm$0.054
tmc2007	100	0.434 $\pm$ 0.098 $\bullet$	0.225 $\pm$ 0.007	0.219 $\pm$ 0.019 $\bullet$	0.330 $\pm$ 0.098 $\bullet$	0.709 $\pm$ 0.033 $\bullet$	0.367 $\pm$ 0.108 $\bullet$	0.214$\pm$0.082
	150	0.434 $\pm$ 0.098 $\bullet$	0.230 $\pm$ 0.009 $\bullet$	0.227 $\pm$ 0.011 $\bullet$	0.330 $\pm$ 0.096 $\bullet$	0.717 $\pm$ 0.027 $\bullet$	0.376 $\pm$ 0.107 $\bullet$	0.216$\pm$0.084
	200	0.434 $\pm$ 0.098 $\bullet$	0.235 $\pm$ 0.009 $\bullet$	0.242 $\pm$ 0.005 $\bullet$	0.335 $\pm$ 0.093 $\bullet$	0.716 $\pm$ 0.028 $\bullet$	0.371 $\pm$ 0.106 $\bullet$	0.222$\pm$0.081
	250	0.434 $\pm$ 0.098 $\bullet$	0.239 $\pm$ 0.008 $\bullet$	0.239 $\pm$ 0.007 $\bullet$	0.325 $\pm$ 0.097 $\bullet$	0.714 $\pm$ 0.031 $\bullet$	0.376 $\pm$ 0.107 $\bullet$	0.219$\pm$0.081
rcv1subset1	100	0.960 $\pm$ 0.039 $\bullet$	0.506 $\pm$ 0.014 $\bullet$	0.570 $\pm$ 0.007 $\bullet$	0.649 $\pm$ 0.039 $\bullet$	0.757 $\pm$ 0.056 $\bullet$	0.523 $\pm$ 0.122 $\bullet$	0.273$\pm$0.111
	150	0.948 $\pm$ 0.038 $\bullet$	0.435 $\pm$ 0.020 $\bullet$	0.627 $\pm$ 0.044 $\bullet$	0.636 $\pm$ 0.098 $\bullet$	0.769 $\pm$ 0.055 $\bullet$	0.512 $\pm$ 0.064 $\bullet$	0.261$\pm$0.085
	200	0.936 $\pm$ 0.046 $\bullet$	0.568 $\pm$ 0.031 $\bullet$	0.707 $\pm$ 0.115 $\bullet$	0.713 $\pm$ 0.054 $\bullet$	0.733 $\pm$ 0.073 $\bullet$	0.503 $\pm$ 0.045 $\bullet$	0.259$\pm$0.062
	250	0.917 $\pm$ 0.035 $\bullet$	0.572 $\pm$ 0.027 $\bullet$	0.687 $\pm$ 0.075 $\bullet$	0.697 $\pm$ 0.079 $\bullet$	0.753 $\pm$ 0.056 $\bullet$	0.482 $\pm$ 0.051 $\bullet$	0.220$\pm$0.046
rcv1subset2	100	0.919 $\pm$ 0.035 $\bullet$	0.471 $\pm$ 0.022 $\bullet$	0.532 $\pm$ 0.036 $\bullet$	0.696 $\pm$ 0.049 $\bullet$	0.766 $\pm$ 0.017 $\bullet$	0.396 $\pm$ 0.103	0.292$\pm$0.136
	150	0.923 $\pm$ 0.029 $\bullet$	0.416 $\pm$ 0.028 $\bullet$	0.568 $\pm$ 0.088 $\bullet$	0.631 $\pm$ 0.052 $\bullet$	0.747 $\pm$ 0.048 $\bullet$	0.416 $\pm$ 0.075 $\bullet$	0.278$\pm$0.112
	200	0.847 $\pm$ 0.103 $\bullet$	0.524 $\pm$ 0.018 $\bullet$	0.587 $\pm$ 0.067 $\bullet$	0.691 $\pm$ 0.103 $\bullet$	0.784 $\pm$ 0.024 $\bullet$	0.418 $\pm$ 0.127 $\bullet$	0.190$\pm$0.068
	250	0.882 $\pm$ 0.053 $\bullet$	0.536 $\pm$ 0.051 $\bullet$	0.619 $\pm$ 0.084 $\bullet$	0.698 $\pm$ 0.039 $\bullet$	0.765 $\pm$ 0.029 $\bullet$	0.481 $\pm$ 0.094 $\bullet$	0.178$\pm$0.050
rcv1subset3	100	0.883 $\pm$ 0.100 $\bullet$	0.482 $\pm$ 0.042 $\bullet$	0.581 $\pm$ 0.101 $\bullet$	0.741 $\pm$ 0.026 $\bullet$	0.788 $\pm$ 0.028 $\bullet$	0.408 $\pm$ 0.064	0.299$\pm$0.174
	150	0.940 $\pm$ 0.053 $\bullet$	0.427 $\pm$ 0.032 $\bullet$	0.591 $\pm$ 0.083 $\bullet$	0.686 $\pm$ 0.094 $\bullet$	0.787 $\pm$ 0.022 $\bullet$	0.442 $\pm$ 0.074 $\bullet$	0.231$\pm$0.107
	200	0.845 $\pm$ 0.137 $\bullet$	0.550 $\pm$ 0.041 $\bullet$	0.582 $\pm$ 0.040 $\bullet$	0.712 $\pm$ 0.142 $\bullet$	0.801 $\pm$ 0.030 $\bullet$	0.488 $\pm$ 0.126 $\bullet$	0.264$\pm$0.100
	250	0.831 $\pm$ 0.143 $\bullet$	0.550 $\pm$ 0.043 $\bullet$	0.637 $\pm$ 0.047 $\bullet$	0.692 $\pm$ 0.041 $\bullet$	0.789 $\pm$ 0.033 $\bullet$	0.442 $\pm$ 0.109 $\bullet$	0.242$\pm$0.070
rcv1subset4	100	0.901 $\pm$ 0.053 $\bullet$	0.401 $\pm$ 0.021 $\bullet$	0.447 $\pm$ 0.032 $\bullet$	0.618 $\pm$ 0.102 $\bullet$	0.784 $\pm$ 0.028 $\bullet$	0.434 $\pm$ 0.036 $\bullet$	0.293$\pm$0.119
	150	0.880 $\pm$ 0.063 $\bullet$	0.351 $\pm$ 0.025 $\bullet$	0.548 $\pm$ 0.051 $\bullet$	0.642 $\pm$ 0.066 $\bullet$	0.783 $\pm$ 0.028 $\bullet$	0.451 $\pm$ 0.063 $\bullet$	0.295$\pm$0.109
	200	0.846 $\pm$ 0.091 $\bullet$	0.438 $\pm$ 0.019 $\bullet$	0.577 $\pm$ 0.099 $\bullet$	0.667 $\pm$ 0.051 $\bullet$	0.775 $\pm$ 0.011 $\bullet$	0.456 $\pm$ 0.082 $\bullet$	0.255$\pm$0.083
	250	0.812 $\pm$ 0.115 $\bullet$	0.467 $\pm$ 0.029 $\bullet$	0.492 $\pm$ 0.058 $\bullet$	0.667 $\pm$ 0.051 $\bullet$	0.775 $\pm$ 0.011 $\bullet$	0.453 $\pm$ 0.090 $\bullet$	0.283$\pm$0.082
rcv1subset5	100	0.828 $\pm$ 0.036 $\bullet$	0.438 $\pm$ 0.025 $\bullet$	0.542 $\pm$ 0.095 $\bullet$	0.656 $\pm$ 0.065 $\bullet$	0.770 $\pm$ 0.053 $\bullet$	0.444 $\pm$ 0.096 $\bullet$	0.291$\pm$0.053
	150	0.815 $\pm$ 0.054 $\bullet$	0.382 $\pm$ 0.028 $\bullet$	0.496 $\pm$ 0.059 $\bullet$	0.699 $\pm$ 0.080 $\bullet$	0.751 $\pm$ 0.009 $\bullet$	0.444 $\pm$ 0.078 $\bullet$	0.276$\pm$0.065
	200	0.811 $\pm$ 0.005 $\bullet$	0.482 $\pm$ 0.241 $\bullet$	0.626 $\pm$ 0.095 $\bullet$	0.552 $\pm$ 0.115 $\bullet$	0.733 $\pm$ 0.111 $\bullet$	0.496 $\pm$ 0.092 $\bullet$	0.236$\pm$0.071
	250	0.783 $\pm$ 0.078 $\bullet$	0.480 $\pm$ 0.038 $\bullet$	0.562 $\pm$ 0.057 $\bullet$	0.583 $\pm$ 0.137 $\bullet$	0.756 $\pm$ 0.085 $\bullet$	0.488 $\pm$ 0.091 $\bullet$	0.251$\pm$0.077

表 A2 关于 coverage 评价指标在线性分类模型上的实验结果 (均值 $\pm$ 标准差). 最优结果加粗表示, 其中符号 $\bullet/\circ$ 分别表示在成对 t 检验 (显著性水平 0.05) 下, DeNo 显著优于/劣于相应对比方法.

Table A2 Experimental results of coverage (mean $\pm$ std) on the linear model. The best performance for each dataset is shown in boldface, where $\bullet/\circ$ indicates whether DeNo is superior/inferior to baselines with pairwise t -test (at the 0.05 significance level).

Dataset	$\gamma\%$	fpml	PARD	PASE	GDF	$\bar{\mathcal{L}}_{\text{BCE}}$	CCMN	DeNo
scene	100	0.408 $\pm$ 0.032 $\bullet$	0.098 $\pm$ 0.009 $\bullet$	0.112 $\pm$ 0.032 $\bullet$	0.132 $\pm$ 0.012 $\bullet$	0.162 $\pm$ 0.012 $\bullet$	0.358 $\pm$ 0.029 $\bullet$	0.085$\pm$0.009
	150	0.410 $\pm$ 0.026 $\bullet$	0.100 $\pm$ 0.013 $\bullet$	0.110 $\pm$ 0.014 $\bullet$	0.134 $\pm$ 0.008 $\bullet$	0.158 $\pm$ 0.010 $\bullet$	0.352 $\pm$ 0.031 $\bullet$	0.085$\pm$0.009
	200	0.408 $\pm$ 0.032 $\bullet$	0.098 $\pm$ 0.010 $\bullet$	0.112 $\pm$ 0.032 $\bullet$	0.132 $\pm$ 0.012 $\bullet$	0.162 $\pm$ 0.012 $\bullet$	0.358 $\pm$ 0.029 $\bullet$	0.085$\pm$0.009
	250	0.420 $\pm$ 0.025 $\bullet$	0.118 $\pm$ 0.016 $\bullet$	0.138 $\pm$ 0.013 $\bullet$	0.136 $\pm$ 0.010 $\bullet$	0.172 $\pm$ 0.007 $\bullet$	0.376 $\pm$ 0.014 $\bullet$	0.091$\pm$0.007
yeast	100	0.488 $\pm$ 0.013 $\bullet$	0.628 $\pm$ 0.055 $\bullet$	0.491 $\pm$ 0.010 $\bullet$	0.547 $\pm$ 0.022 $\bullet$	0.611 $\pm$ 0.029 $\bullet$	0.523 $\pm$ 0.022 $\bullet$	0.465$\pm$0.019
	150	0.504 $\pm$ 0.017	0.660 $\pm$ 0.112 $\bullet$	0.498 $\pm$ 0.012	0.530 $\pm$ 0.017 $\bullet$	0.604 $\pm$ 0.020 $\bullet$	0.524 $\pm$ 0.019 $\bullet$	0.483$\pm$0.025
	200	0.509 $\pm$ 0.019	0.689 $\pm$ 0.083 $\bullet$	0.495 $\pm$ 0.007	0.547 $\pm$ 0.022 $\bullet$	0.611 $\pm$ 0.029 $\bullet$	0.546 $\pm$ 0.029 $\bullet$	0.491$\pm$0.028
	250	0.632 $\pm$ 0.037 $\bullet$	0.706 $\pm$ 0.050 $\bullet$	0.578 $\pm$ 0.136	0.595 $\pm$ 0.033 $\bullet$	0.730 $\pm$ 0.018 $\bullet$	0.610 $\pm$ 0.039 $\bullet$	0.567$\pm$0.029
bookmark	100	0.411 $\pm$ 0.009 $\bullet$	0.149 $\pm$ 0.006 $\bullet$	0.136 $\pm$ 0.005 $\bullet$	0.218 $\pm$ 0.005 $\bullet$	0.293 $\pm$ 0.009 $\bullet$	0.213 $\pm$ 0.004 $\bullet$	0.117$\pm$0.005
	150	0.405 $\pm$ 0.009 $\bullet$	0.156 $\pm$ 0.005 $\bullet$	0.139 $\pm$ 0.005 $\bullet$	0.219 $\pm$ 0.005 $\bullet$	0.295 $\pm$ 0.008 $\bullet$	0.235 $\pm$ 0.020 $\bullet$	0.118$\pm$0.005
	200	0.409 $\pm$ 0.008 $\bullet$	0.150 $\pm$ 0.005 $\bullet$	0.136 $\pm$ 0.005 $\bullet$	0.218 $\pm$ 0.004 $\bullet$	0.293 $\pm$ 0.009 $\bullet$	0.213 $\pm$ 0.004 $\bullet$	0.117$\pm$0.005
	250	0.400 $\pm$ 0.009 $\bullet$	0.175 $\pm$ 0.006 $\bullet$	0.156 $\pm$ 0.009 $\bullet$	0.219 $\pm$ 0.004 $\bullet$	0.287 $\pm$ 0.005 $\bullet$	0.269 $\pm$ 0.013 $\bullet$	0.121$\pm$0.005
delicious	100	0.661 $\pm$ 0.006 $\bullet$	0.565 $\pm$ 0.007 $\bullet$	0.575 $\pm$ 0.007 $\bullet$	0.608 $\pm$ 0.009 $\bullet$	0.649 $\pm$ 0.008 $\bullet$	0.697 $\pm$ 0.007 $\bullet$	0.556$\pm$0.008
	150	0.651 $\pm$ 0.008 $\bullet$	0.606 $\pm$ 0.030 $\bullet$	0.601 $\pm$ 0.002 $\bullet$	0.613 $\pm$ 0.006 $\bullet$	0.663 $\pm$ 0.007 $\bullet$	0.696 $\pm$ 0.012 $\bullet$	0.567$\pm$0.007
	200	0.652 $\pm$ 0.009 $\bullet$	0.632 $\pm$ 0.024 $\bullet$	0.615 $\pm$ 0.009 $\bullet$	0.623 $\pm$ 0.006 $\bullet$	0.674 $\pm$ 0.004 $\bullet$	0.707 $\pm$ 0.015 $\bullet$	0.575$\pm$0.006
	250	0.662 $\pm$ 0.009 $\bullet$	0.397 $\pm$ 0.014 $\bullet$	0.637 $\pm$ 0.015 $\bullet$	0.632 $\pm$ 0.007 $\bullet$	0.684 $\pm$ 0.005 $\bullet$	0.718 $\pm$ 0.011 $\bullet$	0.582$\pm$0.007
Corel16k	100	0.409 $\pm$ 0.053 $\bullet$	0.353 $\pm$ 0.010	0.364 $\pm$ 0.018 $\bullet$	0.413 $\pm$ 0.032 $\bullet$	0.504 $\pm$ 0.019 $\bullet$	0.451 $\pm$ 0.025 $\bullet$	0.336$\pm$0.025
	150	0.407 $\pm$ 0.051 $\bullet$	0.375 $\pm$ 0.013 $\bullet$	0.379 $\pm$ 0.009 $\bullet$	0.412 $\pm$ 0.032 $\bullet$	0.504 $\pm$ 0.019 $\bullet$	0.462 $\pm$ 0.036 $\bullet$	0.343$\pm$0.027
	200	0.407 $\pm$ 0.049 $\bullet$	0.391 $\pm$ 0.012 $\bullet$	0.392 $\pm$ 0.006 $\bullet$	0.417 $\pm$ 0.038 $\bullet$	0.495 $\pm$ 0.017 $\bullet$	0.476 $\pm$ 0.026 $\bullet$	0.353$\pm$0.027
	250	0.406 $\pm$ 0.052 $\bullet$	0.641 $\pm$ 0.026 $\bullet$	0.391 $\pm$ 0.006 $\bullet$	0.410 $\pm$ 0.051 $\bullet$	0.500 $\pm$ 0.028 $\bullet$	0.474 $\pm$ 0.031 $\bullet$	.353$\pm$0.026
tmc2007	100	0.316 $\pm$ 0.031 $\bullet$	0.146 $\pm$ 0.005	0.150 $\pm$ 0.002	0.268 $\pm$ 0.014 $\bullet$	0.425 $\pm$ 0.016 $\bullet$	0.247 $\pm$ 0.024 $\bullet$	0.143$\pm$0.014
	150	0.315 $\pm$ 0.030 $\bullet$	0.156 $\pm$ 0.003 $\bullet$	0.158 $\pm$ 0.008	0.259 $\pm$ 0.015 $\bullet$	0.424 $\pm$ 0.013 $\bullet$	0.252 $\pm$ 0.020 $\bullet$	0.145$\pm$0.012
	200	0.317 $\pm$ 0.030 $\bullet$	0.167 $\pm$ 0.005	0.176 $\pm$ 0.008 $\bullet$	0.265 $\pm$ 0.011 $\bullet$	0.429 $\pm$ 0.017 $\bullet$	0.272 $\pm$ 0.024 $\bullet$	0.159$\pm$0.022
	250	0.316 $\pm$ 0.029 $\bullet$	0.176 $\pm$ 0.005 $\bullet$	0.184 $\pm$ 0.007 $\bullet$	0.259 $\pm$ 0.015 $\bullet$	0.427 $\pm$ 0.013 $\bullet$	0.252 $\pm$ 0.020 $\bullet$	0.160$\pm$0.008
rcv1subset1	100	0.454 $\pm$ 0.014 $\bullet$	0.311 $\pm$ 0.052 $\bullet$	0.234 $\pm$ 0.035 $\bullet$	0.322 $\pm$ 0.026 $\bullet$	0.432 $\pm$ 0.023 $\bullet$	0.305 $\pm$ 0.055 $\bullet$	0.164$\pm$0.023
	150	0.465 $\pm$ 0.028 $\bullet$	0.160 $\pm$ 0.010	0.305 $\pm$ 0.022 $\bullet$	0.341 $\pm$ 0.049 $\bullet$	0.425 $\pm$ 0.025 $\bullet$	0.347 $\pm$ 0.061 $\bullet$	0.156$\pm$0.016
	200	0.449 $\pm$ 0.018 $\bullet$	0.348 $\pm$ 0.043 $\bullet$	0.348 $\pm$ 0.058 $\bullet$	0.358 $\pm$ 0.046 $\bullet$	0.424 $\pm$ 0.024 $\bullet$	0.313 $\pm$ 0.024 $\bullet$	0.157$\pm$0.021
	250	0.455 $\pm$ 0.017 $\bullet$	0.380 $\pm$ 0.046 $\bullet$	0.340 $\pm$ 0.035 $\bullet$	0.350 $\pm$ 0.035 $\bullet$	0.427 $\pm$ 0.024 $\bullet$	0.315 $\pm$ 0.033 $\bullet$	0.162$\pm$0.019
rcv1subset2	100	0.414 $\pm$ 0.056 $\bullet$	0.283 $\pm$ 0.058 $\bullet$	0.255 $\pm$ 0.019 $\bullet$	0.325 $\pm$ 0.017 $\bullet$	0.429 $\pm$ 0.041 $\bullet$	0.282 $\pm$ 0.056 $\bullet$	0.155$\pm$0.032
	150	0.421 $\pm$ 0.057 $\bullet$	0.163 $\pm$ 0.016	0.288 $\pm$ 0.053 $\bullet$	0.322 $\pm$ 0.027 $\bullet$	0.418 $\pm$ 0.073 $\bullet$	0.274 $\pm$ 0.043 $\bullet$	0.167$\pm$0.039
	200	0.410 $\pm$ 0.057 $\bullet$	0.347 $\pm$ 0.060 $\bullet$	0.307 $\pm$ 0.033 $\bullet$	0.352 $\pm$ 0.099 $\bullet$	0.449 $\pm$ 0.060 $\bullet$	0.270 $\pm$ 0.021 $\bullet$	0.174$\pm$0.051
	250	0.414 $\pm$ 0.056 $\bullet$	0.329 $\pm$ 0.037 $\bullet$	0.333 $\pm$ 0.057 $\bullet$	0.341 $\pm$ 0.052 $\bullet$	0.441 $\pm$ 0.065 $\bullet$	0.297 $\pm$ 0.026 $\bullet$	0.175$\pm$0.056
rcv1subset3	100	0.419 $\pm$ 0.058 $\bullet$	0.282 $\pm$ 0.058 $\bullet$	0.281 $\pm$ 0.046 $\bullet$	0.369 $\pm$ 0.041 $\bullet$	0.427 $\pm$ 0.026 $\bullet$	0.287 $\pm$ 0.073 $\bullet$	0.157$\pm$0.058
	150	0.422 $\pm$ 0.051 $\bullet$	0.164 $\pm$ 0.013	0.302 $\pm$ 0.039 $\bullet$	0.356 $\pm$ 0.114 $\bullet$	0.440 $\pm$ 0.069 $\bullet$	0.273 $\pm$ 0.046 $\bullet$	0.147$\pm$0.041
	200	0.418 $\pm$ 0.070 $\bullet$	0.314 $\pm$ 0.063 $\bullet$	0.290 $\pm$ 0.025 $\bullet$	0.339 $\pm$ 0.053 $\bullet$	0.436 $\pm$ 0.049 $\bullet$	0.281 $\pm$ 0.030 $\bullet$	0.163$\pm$0.044
	250	0.412 $\pm$ 0.050 $\bullet$	0.351 $\pm$ 0.074 $\bullet$	0.331 $\pm$ 0.044 $\bullet$	0.336 $\pm$ 0.083 $\bullet$	0.441 $\pm$ 0.047 $\bullet$	0.273 $\pm$ 0.041 $\bullet$	0.180$\pm$0.062
rcv1subset4	100	0.373 $\pm$ 0.063 $\bullet$	0.228 $\pm$ 0.041 $\bullet$	0.222 $\pm$ 0.021 $\bullet$	0.330 $\pm$ 0.083 $\bullet$	0.413 $\pm$ 0.055 $\bullet$	0.280 $\pm$ 0.103 $\bullet$	0.157$\pm$0.057
	150	0.366 $\pm$ 0.058 $\bullet$	0.141 $\pm$ 0.010	0.268 $\pm$ 0.030 $\bullet$	0.345 $\pm$ 0.101 $\bullet$	0.407 $\pm$ 0.044 $\bullet$	0.262 $\pm$ 0.064 $\bullet$	0.153 $\pm$ 0.042
	200	0.370 $\pm$ 0.059 $\bullet$	0.281 $\pm$ 0.038 $\bullet$	0.291 $\pm$ 0.030 $\bullet$	0.308 $\pm$ 0.049 $\bullet$	0.394 $\pm$ 0.037 $\bullet$	0.262 $\pm$ 0.047 $\bullet$	0.164$\pm$0.039
	250	0.366 $\pm$ 0.069 $\bullet$	0.284 $\pm$ 0.043 $\bullet$	0.246 $\pm$ 0.028 $\bullet$	0.309 $\pm$ 0.049 $\bullet$	0.394 $\pm$ 0.037 $\bullet$	0.251 $\pm$ 0.034 $\bullet$	0.151$\pm$0.052
rcv1subset5	100	0.393 $\pm$ 0.015 $\bullet$	0.273 $\pm$ 0.051 $\bullet$	0.266 $\pm$ 0.054 $\bullet$	0.321 $\pm$ 0.017 $\bullet$	0.434 $\pm$ 0.016 $\bullet$	0.337 $\pm$ 0.045 $\bullet$	0.165$\pm$0.026
	150	0.392 $\pm$ 0.015 $\bullet$	0.161 $\pm$ 0.009	0.247 $\pm$ 0.041 $\bullet$	0.325 $\pm$ 0.014 $\bullet$	0.424 $\pm$ 0.012 $\bullet$	0.310 $\pm$ 0.019 $\bullet$	0.166 $\pm$ 0.024
	200	0.386 $\pm$ 0.021 $\bullet$	0.300 $\pm$ 0.044 $\bullet$	0.300 $\pm$ 0.041 $\bullet$	0.337 $\pm$ 0.024 $\bullet$	0.408 $\pm$ 0.030 $\bullet$	0.307 $\pm$ 0.062 $\bullet$	0.181$\pm$0.024
	250	0.373 $\pm$ 0.020 $\bullet$	0.281 $\pm$ 0.024 $\bullet$	0.308 $\pm$ 0.027 $\bullet$	0.334 $\pm$ 0.041 $\bullet$	0.431 $\pm$ 0.019 $\bullet$	0.299 $\pm$ 0.075 $\bullet$	0.187$\pm$0.022

和噪声比例下均达到最佳 one error, 其中在 rcv1subset1~5、delicious、bookmark 等数据集上的优势尤为明显, 相比次优方法平均降低 10%~30%. 这说明所引入的排他性约束能有效抑制非候选标记的误激活. 表 A2 显示, 随着噪声比例从 100% 提升至 250%, fpml、GDF、 $\bar{\mathcal{L}}_{\text{BCE}}$ 等方法的 coverage 值均出现显著上升, 而 DeNo 的增幅最小并始终保持最优. 该结果验证了集中式伪标记分布的渐进式去噪效果, 以及平方损失在排序优化过程中带来的平滑和稳定性. 综上, DeNo 在 one error 和 coverage 上的整体领先表明, 其排序结果更接近真实标记分布, 并在复杂噪声环境下展现出更强的排序鲁棒性和泛化能力.

A denoising framework with exclusivity constraints for partial multi-label learning

Yi GAO^{1,2}, Jia-Qi LI¹ & Min-Ling ZHANG^{1,2*}

1. School of Computer Science and Engineering, Southeast University, Nanjing 211189, China

2. Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education, Nanjing 211189, China

* Corresponding author. E-mail: zhangml@seu.edu.cn

Abstract In partial multi-label learning (PML), each training sample is associated with a set of candidate labels, which contains both relevant and noisy labels. Disambiguation is a key strategy to address the PML problem. Existing approaches mainly rely on modeling prediction probabilities within candidate label sets or on specific assumptions (such as smoothness or sparsity), which are often difficult to satisfy in real-world scenarios, and they often ignore the exclusive information contained in non-candidate labels. To effectively suppress noise and enhance the model's ability to identify true labels, we propose a denoising PML framework with exclusive constraints, called DeNo. The framework constructs a centralized pseudo-label distribution within the candidate set, progressively amplifies high-confidence labels across iterations, and achieves gradual denoising. Meanwhile, a square loss is employed during training to maintain continuous alignment between model outputs and the updated distribution, thereby preventing error accumulation. In addition, DeNo leverages the exclusive information contained in non-candidate labels to suppress irrelevant label activations, extract reliable negative signals, and further reduce interference from irrelevant information. Experiments on 11 multi-label datasets demonstrate that the proposed method outperforms baseline approaches in both suppressing noise and improving model's performance, validating its effectiveness. The code can be available at <https://github.com/gaoyi439/DeNo>.

Keywords multi-label learning, partial multi-label learning, denoising, square loss, exclusive information