代表性工作详述

Revisiting Sharpness-Aware Minimization: A More Faithful and Effective Implementation [PDF]

ICLR 2026 · 唯一通讯作者 优化算法 | 锐度感知优化

问题 SAM 旨在最小化参数邻域内的最大损失以改善泛化，实际则先沿梯度上升方向扰动参数，再沿扰动后的梯度反方向更新原参数：这一更新方式仍缺乏直观的几何解释。分析 揭示单步扰动后的梯度能比原梯度更准确地以原参数为起点指向邻域最大损失点，该更新方式使原参数更直接地远离高损失区域。但这一方向估计往往仍有偏差，多步扰动还可能加剧偏差。方法 提出 XSAM，利用单步或多步梯度上升信息，构造经过原参数与当前已知最大损失点的二维平面，按指定半径在平面内显式搜索最大损失方向，沿其反方向更新参数。理论 在二阶近似、Hessian 正定等条件下，证明在同一足够大的邻域半径下，单步扰动后的梯度方向比原梯度方向对应更高损失，且 XSAM 的搜索平面内还存在对应损失更高的方向：更具体地解释了 SAM 改善泛化的机制，并说明了 XSAM 相较于 SAM 在邻域最大值方向估计上的优势。实验 在单步扰动设置下的所有测试任务中，XSAM 稳定优于 SAM。在多步扰动设置下，XSAM 稳定优于多步 SAM 及其变体，且性能随扰动步数增加呈提升趋势。曲率与损失曲面分析显示，XSAM 获得比 SAM 更平坦的解。

个人贡献： 主导方法设计，完成核心代码实现与理论分析，指导实验研究，主导论文撰写。

Self-Supervised Disentanglement via Cluster-Dependent Subspace Rotational Equivariance [PDF]

投稿中 · 唯一通讯作者 表示学习 | 自监督解耦

问题 自监督不变性训练拉近同一样本不同增强视图的表示，能形成语义簇，却缺乏对簇内表示的控制。已有方法在不变性训练中加入全局旋转等变性约束，以跨簇共享的旋转组织增强引起的表示变化。分析 揭示不变性目标在联合训练中会干扰等变性学习；全局旋转等变性约束不同语义簇的旋转变化空间共享朝向。方法 提出簇依赖子空间旋转等变框架 CRED：先学习不变性形成语义簇，再移除不变性目标，在各簇的局部坐标系中学习旋转等变性，并将不同变化分量分别表征在相互正交的二维子空间内。理论 证明该框架可精确解耦增强诱导变化，包括单个增强引起的多种变化；在施加防角度回绕约束后，簇内角度与非周期变化分量能进一步形成严格单调的一一对应。实验 经坐标变换后处理的表示在合成数据上达到近乎完美解耦（DCI 三项指标均 >=0.99）。真实图像的等变与解耦结构指标均优于所比较基线。

个人贡献： 主导方法设计，完成核心代码实现，指导理论分析与实验研究，主导论文撰写。

Residual Multi-Task Learner for Applied Ranking [PDF]

KDD 2024 · 唯一通讯作者 网络架构 | 多任务学习

问题 电商排序需联合预测点击、购买等行为。现有共享专家类方法缺少对任务关系的显式建模；而已有的跨任务注意力模块又开销较高，难以满足低时延要求。方法 提出 ResFlow，通过跨任务残差连接将前序任务的中间表示以加法传递给后续任务的对应层，使后者在此基础上学习增量。支持按任务依赖或数据稀疏性组织信息流，并扩展至回归任务。实验 在 Shopee Search 预排序线上 A/B 测试中，每用户订单数较原线上模型提升 1.29%，系统时延基本不变，已全量部署。离线多行为预测及回归任务中均取得性能提升。

个人贡献： 指导方法设计，指导实验研究，主导论文撰写。

代表性工作详述

Recurrent Temporal Revision Graph Networks [PDF]

NeurIPS 2023 · 唯一通讯作者 网络架构 | 时序图神经网络

问题 现有时序图模型常仅聚合近期邻居以控制开销，导致表示偏向短期行为，较早的邻居信息被遗漏。方法 提出 RTRGN，在聚合层用循环神经网络维护各节点状态，持续聚合历史邻居信息，并根据历史邻居后续的状态变化修订节点表示。进一步通过节点标记区分当前递归路径中已出现的节点与其他节点。理论 在相关映射单射等条件下，证明其在时序图同构判别与链接预测中的表达能力严格超越Temporal-1WL 类模型。实验 在 8 个数据集上，训练中已见与未见节点的链接预测性能均优于对比方法。两项时序图同构判别任务均达 100% 准确率，去除节点标记后均出现误判。

个人贡献： 主导方法设计，指导理论分析与实验研究，主导论文撰写。

Clustered Embedding Learning for Recommender Systems [PDF]

WWW 2023 · 唯一通讯作者 表示学习 | 推荐系统嵌入

问题 推荐系统的独立嵌入表存储开销大，在稀疏交互下泛化不足。已有聚类方案依赖预设结构，而既有哈希共享方案则面临随机映射损害预测质量或映射调整受限等问题。方法 提出 CEL，通过动态聚类使同簇用户或物品共同学习并共享嵌入：从单簇出发，每次选取一簇，对簇内各实体损失关于簇共享嵌入的梯度做主成分分析，按梯度在第一主成分上的投影将该簇二分；按预测损失动态调整实体的簇归属。在非负矩阵分解框架及满秩等条件下，证明聚类与嵌入解本质唯一、最优簇数唯一。CEL-Lite 通过限定各实体的交互缓存与候选簇采样规模，支持高效在线更新。实验 改善对交互稀疏物品的泛化。企业电商数据上，CEL-Lite 相较采用完整嵌入表的业务模型，将嵌入表缩至约1/2650，同时将 AUC 提高 0.61 个百分点。

个人贡献： 指导方法设计，指导理论分析与实验研究，主导论文撰写。

AdaShift: Decorrelation and Convergence of Adaptive Learning Rate Methods [PDF]

ICLR 2019 · 共同第一作者 优化算法 | 自适应学习率

问题 Adam 根据梯度的二阶矩自适应调整步长，但在某些凸优化问题上无法收敛到最优解。AMSGrad 通过保留二阶矩估计的历史最大值解决这一问题，却可能持续压低有效步长，拖慢训练。分析 揭示当前梯度与二阶矩估计的相关性会导致等效累计步长偏置：幅度较大的梯度往往对应较小的累计步长，较小的梯度则相反，从而使整体更新方向偏移甚至反转。方法 提出 AdaShift，将梯度延后用于二阶矩估计，在梯度时序独立等条件下消除上述相关性，并证明不同梯度的等效累计步长具有相同期望。实验 在 Adam 的随机在线优化反例中，AdaShift 能正确收敛，且在相同基础学习率下比 AMSGrad 收敛更快。在所测试的一系列实际任务中，AdaShift 的训练与测试表现均不低于 Adam，部分场景表现更优。

个人贡献： 主导方法设计，完成核心代码实现，指导理论分析与实验研究，主导论文撰写。

代表性工作详述

Lipschitz Generative Adversarial Nets [PDF]

ICML 2019 · 唯一第一作者、共同通讯作者 　　生成模型 | 生成对抗网络

问题 针对真实与生成分布支撑不重叠时 JS 散度保持常数的问题，WGAN 引入 Wasserstein 距离改善训练。分析 指出GAN 稳定训练的关键在于判别器给生成器的梯度能否可靠传递真实分布的信息，训练困难不能简单归因于距离度量的不连续性。特别地，即使使用连续性良好的 Wasserstein 距离，若采用其放松 Lipschitz 约束的等价表达，所得梯度也可能缺乏有效信息。理论 证明在不约束判别函数的 GAN 中，最优判别函数在各点的取值仅由真实与生成分布在该点的密度决定，故不能保证其梯度提供有效的分布信息；解释由此产生的局部贪心更新为何可能导致模式崩塌。方法 提出以 Lipschitz 约束为核心的 LGANs 目标函数族，证明其最优判别函数在生成样本处的非零梯度径直指向具体真实样本。指出常规梯度惩罚会对非最大梯度也施加惩罚，与 Lipschitz 常数的正则化目标不一致；提出 MaxGP，仅惩罚采样点中的最大梯度范数。实验 在合成数据与真实图像上验证了判别器梯度径直指向具体真实样本的性质。部分设置下，常规梯度惩罚导致训练发散，而改用 MaxGP 后可收敛。在图像生成实验中，统一采用 MaxGP 时，所测试的 LGANs 非线性目标均取得比 WGAN 更低的 FID。非线性目标下的判别器输出也较 WGAN 更稳定。

个人贡献： 完成方法设计与代码实现，主导理论分析，完成实验研究，主导论文撰写。

Activation Maximization Generative Adversarial Nets [PDF]

ICLR 2018 · 唯一第一作者 　　生成模型 | 生成对抗网络

问题 已有 GAN 利用类别标签改善生成质量，但部分方法训练生成器时，以增大生成样本被判为各真实类别的概率总和为目标。分析 通过类别感知梯度分析，揭示这一目标会导致指向多个类别的梯度叠加，合成方向可能不指向任何类别。方法 提出AM-GAN，将真伪判别与类别识别统一为多分类对抗任务，并以明确的目标类别引导生成器训练。提出动态标签机制，将判别器当前预测概率最高的真实类别作为生成样本的目标标签，无需预设标签，保持无条件生成，并缓解预设标签训练中易出现的类内模式崩塌。从激活最大化视角解释 GAN 的训练机制，并通过目标函数分解，系统揭示AM-GAN 与多种既有方法的内在联系。指标 指出基于 ImageNet 分类器的 Inception Score 在 CIFAR-10 等非 ImageNet 数据集上主要反映样本多样性，难以可靠衡量单样本质量。提出 AM Score，使用目标数据集上预训练的分类器，以类别预测分布的确定性刻画单样本质量，以真实与生成样本的平均类别预测分布之间的匹配程度刻画类别多样性。实验 AM-GAN 的生成质量和多样性比基线更好。为对照模型的辅助分类器引入对抗训练，缓解了类内模式崩塌并改善生成质量。在AM-GAN 及改进后的对照模型中，动态标签版本的 Inception Score 均高于各自的预设标签版本。

个人贡献： 完成方法设计、代码实现、理论分析及实验研究，主导论文撰写。

Sparse-as-Possible SVBRDF Acquisition [PDF]

ACM TOG 2016 · 唯一第一作者 　　计算机图形学 | 表面材质建模

问题 已有表面反射重建方法将各位置的反射属性表示为共享基材质的加权混合，并通过核范数最小化近似控制基材质数量。这类方法可能高估基材质总数，且未显式约束每点混合的基材质数，因而需要较多观测。方法 提出双层稀疏结构：全局采用少量自适应基材质，且各表面点混合其中更少的基材质。利用预先扫描得到的物体几何和实测自然环境光照，联合优化基材质及空间混合权重，重建表面材质。实验 在一项合成材质对比实验中，仅用 6 个视图即可达到前人方法使用 62 个视图时的重建质量。真实物体实验中，用少量照片重建物体材质，部分物体仅需单幅照片；重建材质在新光照下的渲染结果与实拍外观相符。

个人贡献： 主导方法设计，完成代码实现与实验研究，协助论文撰写。