Hey there,

Here is your disruptive ML/AI digest for the day. Here's what you folks might find interesting today.

TOP PICKS

Theory and Generalization / Probabilistic and Statistical Methods

Minimax Optimality of Score-Entropy Discrete Diffusion (Cho, Wu)

This paper establishes the *fundamental statistical limits* of concrete score estimation in discrete diffusion models, moving beyond the prior literature's focus on sampling efficiency under assumed small error. The authors prove a minimax lower bound under the score-entropy loss for uniform and masking discrete diffusions, and construct an MLE-based thresholding estimator that matches it up to polylogarithmic factors. The key insight is that neighboring density ratios -- which control the gap -- are naturally bounded under both diffusion variants, yielding nearly tight minimax bounds for the aggregated score estimation error. This matters because it closes a gap between SEDD's strong empirical performance and our theoretical understanding of *why* the score estimation itself is statistically tractable, with implications for sample complexity measured in KL divergence between target and generated distributions.

Novel Architectures and Objectives

Lightweight Adaptive ReduNet via Hyperspherical Manifold Learning (Huang, Yan, Dai, Tang)

ReduNet is a white-box architecture that derives layer parameters explicitly from the *maximal coding rate reduction* (MCR^2) principle, bypassing backpropagation entirely in favor of a forward layer-wise construction. This paper addresses its critical bottleneck -- the large number of unfolded layers needed for the MCR^2 objective to stabilize -- by introducing hyperspherical manifold learning and adaptive step sizes into the update rule. The result is an order-of-magnitude reduction in required layers, with the system using approximately 1/29 of the parameter storage of the original unfolded module at comparable accuracy. This is relevant because it strengthens the case for mathematically interpretable, non-backprop architectures as practical alternatives, rather than theoretical curiosities. The hyperspherical constraint also connects the representation geometry directly to the rate reduction objective in a principled way.

Probabilistic and Statistical Methods / Transferable Methods from Adjacent Fields

TracingFlow: A Simulation-Free Trajectory Inference Framework Based on Second-Order Dynamics (Sun, Wu, Huang, Zhou)

TracingFlow generalizes Flow Matching to *second-order dynamics* by regressing acceleration fields rather than velocity fields, solving what the authors call the Dynamical Optimal Acceleration Transport (DOAT) problem. This is a genuine conceptual departure from the first-order, memoryless assumption underlying most OT-based trajectory inference methods, and it matters because many real-world processes -- cell differentiation, regulatory networks -- exhibit momentum and time-delayed responses that first-order models structurally cannot capture. The framework is simulation-free, provides exact solutions to the DOAT objective, and integrates lineage tracing priors for biological plausibility. The demonstration that second-order formulation captures high-curvature transitions that first-order methods over-smooth is a concrete, mechanistic argument rather than a benchmark-driven claim.

Theory and Generalization

When Clean Data Hurts: Learning with Monotone Corruptions Beyond Binary Classification
(Asilis, Dughmi, Pabbaraju)

This paper extends the recently introduced model of *monotone adversarial corruptions* -- where correctly labeled but distributionally unrelated examples are injected into an i.i.d. training set -- from binary classification to multiclass and partial binary concept classes. The primary result is striking: a learnable multiclass problem of DS dimension only 2 becomes *altogether unlearnable* under a monotone adversary that inserts a linear number of corrupted points. This is a qualitative jump from the binary case, where corruption only degrades error rates by a logarithmic factor. The authors complement this impossibility with a tight threshold: every class remains learnable when adaptive additions are o(n). This matters because it reveals that the robustness of optimal learners to distributional shift is far more fragile in multiclass settings than binary theory suggests, and the gap between "degraded" and "broken" learning hinges on the interaction between adversary budget and problem structure.

Back tomorrow with more!

P.S. In the most recent submission window: cs.LG: 88 papers stat.ML: 10 papers