



復旦大學

数据科学与深度学习
青年学者论坛

What Makes Looped Transformers Perform Better Than Non-Recursive Ones (Provably)



滕佳烨

(上海财经大学)

时间: 2025年11月13日 14: 00-15: 00

地点: 光华楼东主楼 2001

Abstract:

While looped transformers (Looped-Attn) often outperform standard transformers (Single-Attn) on complex reasoning tasks, their theoretical advantage remains unclear. Guided by empirical observations of distinct sample- and Hessian-level dynamics, we interpret this phenomenon via loss landscape geometry, extending the River-Valley model to distinguish U-shaped (flat) and V-shaped (steep) valleys. Based on these observations, we conjecture that Looped-Attn's recursive architecture induces a River-V-Valley inductive bias; under this bias, our theoretical derivations guarantee improved loss convergence along the river via valley hopping and encourage learning of complex patterns, compared to the River-U-Valley bias of Single-Attn. Building on this insight, we propose SHIFT (Staged Hlerarchical Framework for Progressive Training), a staged training framework that accelerates Looped-Attn training while maintaining comparable performance.

个人简介:

滕佳烨, 上海财经大学统计与数据科学学院助理教授, 主要研究方向为理论机器学习, 包括泛化理论与共形预测。他博士毕业于清华大学交叉信息研究院, 曾赴普林斯顿大学访问。曾获清华大学优秀毕业生、优秀博士论文等奖项, 并入选2025年CCF理论计算机科学博士学位论文激励计划提名。他是人工智能研讨班 FAI-Seminar 的发起人。主页: www.tengjiaye.com.