精选快讯　软件快讯　2026年11月21日 09:38

腾讯姚顺雨领衔混元团队深夜AMA，回应模型规模与训练挑战

要点

- 11月21日凌晨，腾讯首席AI科学家姚顺雨带领混元团队在海外社交媒体社区进行了一场持续两个多小时的文字公开问答。
- 在模型规模与性能方面，混元团队指出Hy3和Hy4已证明较小模型能以更低成本超越较大模型。
- 关于训练过程，团队成员坦言从Hy3到Hy4的过渡非常艰难，由于缺乏详细消融实验，训练初期存在大量不确定性。

11月21日凌晨，腾讯首席AI科学家姚顺雨带领混元团队在海外社交媒体社区进行了一场持续两个多小时的文字公开问答。该环节面向全球开发者，团队就模型技术路线、训练挑战及未来规划进行了详细回应。

在模型规模与性能方面，混元团队指出Hy3和Hy4已证明较小模型能以更低成本超越较大模型。尽管Hy4仍有改进空间，团队正投入大量精力进行大规模强化学习，并计划训练更大规模模型。在架构选择上，团队在MLA和GQA之间反复权衡，最终选定MLA，因DSA的引入降低了MLA解码过程中的计算强度，改变了平衡状态。

关于训练过程，团队成员坦言从Hy3到Hy4的过渡非常艰难，由于缺乏详细消融实验，训练初期存在大量不确定性。然而训练后期出乎意料地顺利，被形容为“全押后获胜”。在后训练环节，团队认为最难修正的行为是“过度思考”，同时强调将继续改善用户体验和模型安全。

报道来源：mrjjxw

来源网站：https://www.inews.today/