

Low-rank self-play fine-tuning for small LLMs

Saturday, 17 May 2025 15:30 (12 minutes)

В работе исследуется проблема дообучения больших языковых моделей (LLM) в условиях ограниченных ресурсов. Под ограниченными ресурсами понимается видеопамять, человеческое участие и время обучения. В работе рассматриваются модели до 1.5B. Предлагается метод дообучения, основанный на внедрении адаптеров LoRA, малоранговых разложений матриц, в слои архитектуры трансформера, и использовании стратегии self-play - текущая итерация генерирует предсказания, а обучающаяся повышает качество с помощью разграничения настоящих предсказаний от сгенерированных. Метод может снизить количество обучаемых параметров в 10000, и память в три раза, также он не требует размеченных данных помимо используемых на этапе SFT.

Primary author: MUN, Pavel

Co-author: Mr GRABOVOY, Andrey (MIPT, Moscow Russia)

Presenter: MUN, Pavel

Session Classification: 17-Машинное обучение и нейросети

Track Classification: Машинное обучение и нейросети