

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

Pavel Mun, Nikita Okhotnikov

6 марта 2025 г.

1 Abstract

Работа исследует проблему сложности дообучения языковых моделей (LLM). Под сложностью подразумевается количество затрачиваемых человеческих ресурсов. Обычно для повышения качества полученной модели используют размеченные данные, что влечет использование большого числа ресурсов. Рассматривается метод SPIN, метод повышения качества дообучения LLM с использованием стратегии self-play, который не требует наличия данных помимо, используемых на дообучении. Предлагается оценить границы применимости данного метода для малых моделей и оправданности применения метода в условиях ограниченности ресурсов. Для анализа качества метода будет использована группа датасетов, таких как MMLU, Winogrande.