

# Улучшение Fine Tuning LLM с помощью Multi Token Prediction

*Sunday, 18 May 2025 15:30 (12 minutes)*

**Abstract:**

Мы расширяем результат статьи Gloeckle (2024), заключающийся в том, что LLM обученная с несколькими дополнительными головами (используя метод multi token prediction) получается лучшего качества. В нашей работе мы показываем, что используя multi token prediction, можно улучшить также качество Fine Tuning-а моделей. Также в данной работе улучшения появляются начиная с размера моделей 1B, в том время как в оригинале - начиная с 7B.

**Primary author:** MOSTOVYKH, Egor (МИПТ)

**Co-author:** Mr БОГДАНОВ, Кирилл (МФТИ)

**Presenter:** MOSTOVYKH, Egor (МИПТ)

**Session Classification:** 18-Машинное обучение и нейросети

**Track Classification:** Машинное обучение и нейросети