

Рецензия

Статья исследует актуальную тему обучения больших языковых моделей для конкретных задач. Подход заключается в применении низкоранговых обновлений весов. Данный подход позволяет снизить расходы на обучение и использование модели, сокращая затраты времени, памяти и дискового пространства. Это делает статью очень актуальной

Плюсы

1. В отличие от ранних подходов, метод из статьи не добавляет никаких задержек и не накладывает дополнительных ограничений
2. Проведен обширный эмпирический анализ эффективности по сравнению с другими подходами. Сравниваются несколько популярных подходов на нескольких датасетах и всех популярных моделях
3. Проведен обширный эмпирический анализ эффективности различных параметров метода. Сравнивается, какие матрицы нужно учить с помощью низкорангового обновления и какой ранг нужно брать

Минусы

1. Авторы сами не очень понимают, почему это так хорошо работает
2. Нет данных о времени обучения и графиков обучения

Качество статьи

Статья написана понятным языком с обширными комментариями и пояснениями

Воспроизводимость

Авторы дают все необходимые объяснения подхода, описывают датасеты и приводят оценки скорости обучения. Авторы использовали одну Tesla V100 для экспериментов, поэтому должно быть возможно воспроизвести их в домашних условиях. Но по статье непонятно, сколько времени это займет

Оценка статьи

Статья актуальна и очень хорошо написана, я считаю что она заслуживает оценки 9. В своей оценке я уверен на 4