

Fine-Tuning can Distort Pretrained Features and Underperform Out-of-Distribution

Рецензия

Вклад работы

- Теоретическое и эмпирическое сравнение fine-tuning'а и linear probing'а на основе их ID/OOD результатов
- LP показывает лучшие результаты чем FT в задачах OOD, а LP-FT даёт нам лучшее из двух миров

Публикации и авторы

- Работа была опубликована 21 февраля 2022 года и представлена на конференции ICLR в апреле этого же года.
- Авторами работы: Ananya Kumar, Aditi Raghunathan, Robbie Jones, Tengyu Ma и Percy Liang

Предшествующие работы

- Rudelson & Vershynin (<https://arxiv.org/pdf/0802.3956.pdf>)
- Других работ, оказавших сильное влияние, нет

Сильные стороны

- Вклад: несмотря на то, что сама техника дообучения, которую используют авторы, не является новой, авторы предлагают сильный теоретический анализ её надёжности.
- Новизна: авторы отвечают на вопрос о том, почему разные техники дообучения сетей приводят к сильно разным результатам и как мы можем получить лучшие результаты fine-tuning'а, если будем учитывать его особенности.

Слабые стороны

- Странный выбор моделей для экспериментов
- Разные модели используются для разных ID & OOD пар