

Ziplt



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Яндекс

Weight averaging



- Averaging Weights Leads to Wider Optima and Better Generalization, 2019 – усредняют веса во время обучения моделей
- Robust fine-tuning of zero-shot models, 2022 – объединяют fine-tuned модель и zero-shot

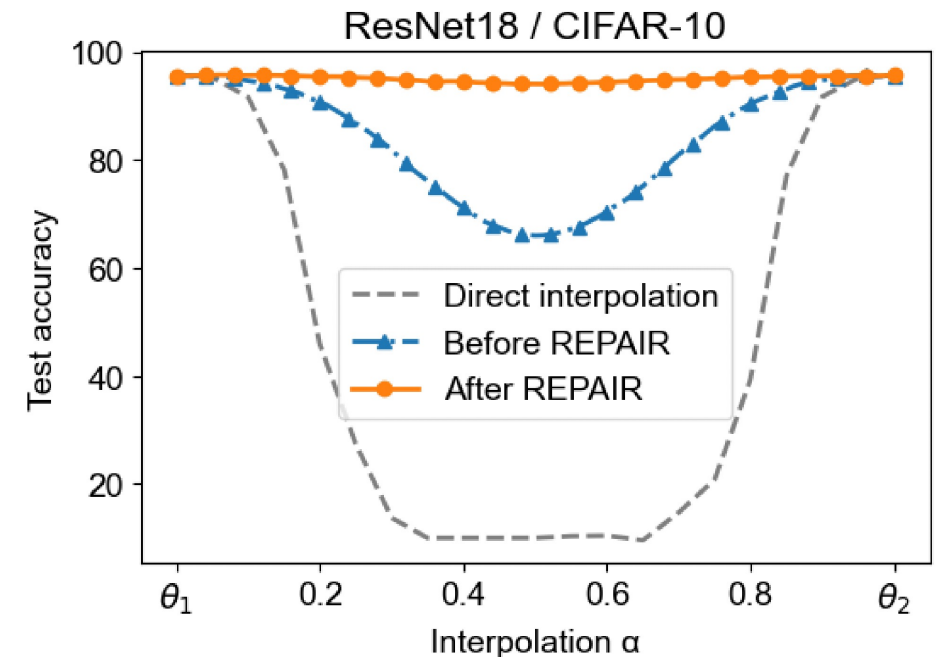
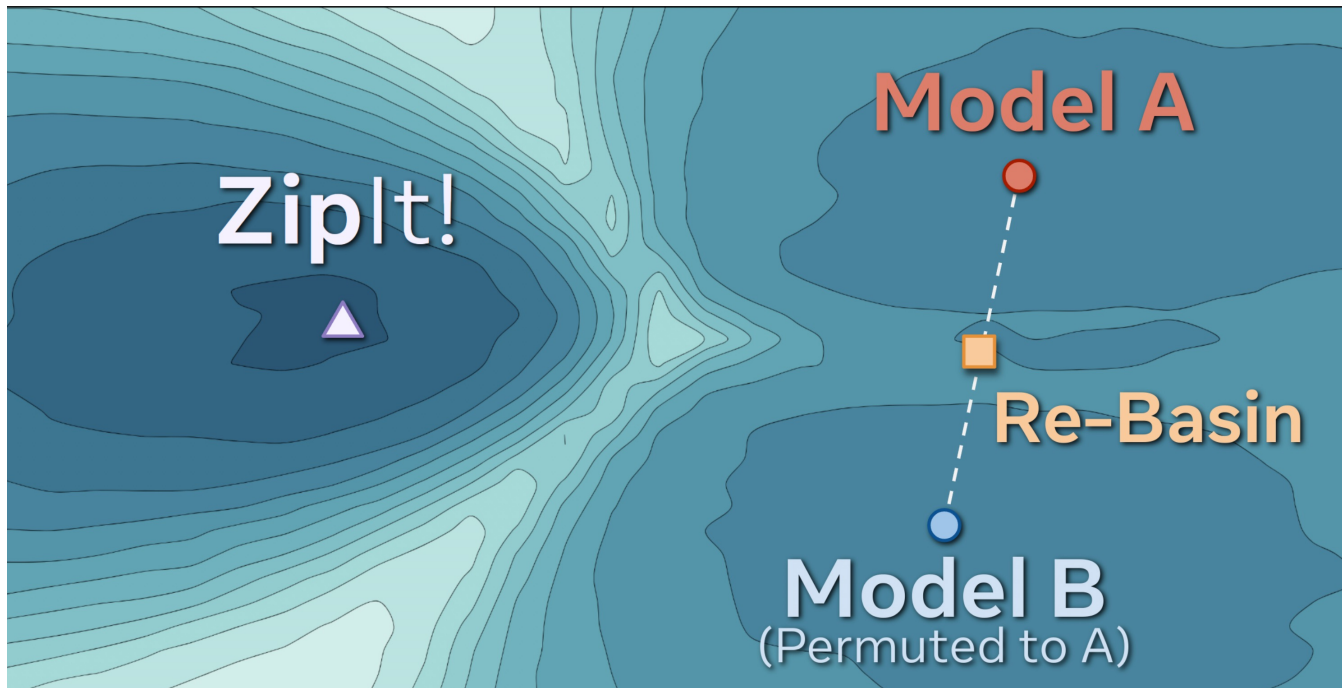
Связанные задачи



- Model soups – про объединение одинаковых и одинаково обученных моделей с разными гиперпараметрами в одну более качественную
- Distillation for Multi-task Learning – умеет решать разные задачи, но требует обучения финальной модели

Варианты без дообучения

- [Git Re-Basin, 2022](#) – объединяет модели, обученные на одинаковой задаче, использует конкатенацию слоев через матрицу отображения
- [Repair, 2023](#) – улучшает Re-Basin, решает такую же задачу



Связанные задачи



- [Model Fusion, 2023](#) – объединяет модели с помощью оптимизации через задачу оптимального транспорта
- [Re-basin via implicit sinkhorn differentiation, 2022](#) – находит оптимальную перестановку для Re-basin с помощью решения непрерывной задачи, есть эксперименты на Split Cifar