

Ensemble Aware Optimization In Federated Learning

Белинский Тимофей Зайнуллин Амир

МФТИ

МФТИ

Феоктистов Дмитрий

ВМК МГУ, Yandex Research ML Residency

14 апреля 2025 г.

План

1. Постановка задачи
2. Алгоритм
3. Текущий прогресс
4. Некоторые результаты
5. Гетерогенность
6. Дальнейшее исследование

Постановка задачи

- ▶ Пусть имеется k агентов
- ▶ Данные у каждого конфиденциальные
- ▶ Хочется построить ансамбль, при этом минимально обмениваясь данными
- ▶ Имеем дело со следующей оптимизационной задачей:

$$\min_{\theta_1 \dots \theta_k} \mathcal{L} \left(\frac{1}{k} \sum_{i=1}^k f(\theta_i) \right) + \mu \frac{1}{k} \sum_{i=1}^k \mathcal{L}(f(\theta_i))$$

Алгоритм

1. Каждый агент делит собственные данные на 2 части: для *обучающего* цикла и для *коммуникации*
2. Все части для коммуникации объединяются в целую выборку
3. Агенты обучают модели исключительно на своих данных
4. Иногда делятся логитами на *коммуникационной* выборки и делают шаг обучения, как ансамбль
5. На инференсе агенты делятся зашумленными весами обученных моделей

Текущий прогресс

- ▶ Реализован пайплайн умного ансамбля на распределенном этапе
- ▶ Проведены эксперименты по оценки эффективности метода
- ▶ Проведены эксперименты с различными модификациями задачи

Некоторые результаты

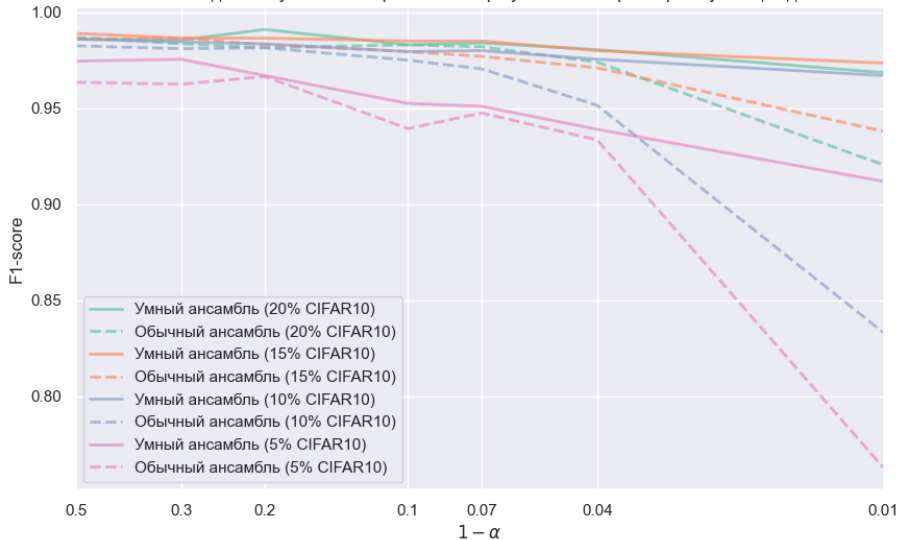
- ▶ В условиях гомогенности данных метод дает прирост в качестве, но небольшой
- ▶ Результат эксперимента становится слабо заметен, если модель слишком производительная
- ▶ В условиях гетерогенности эффект виден более выразительно

Гетерогенность

- ▶ Что если данные распределены не равномерно между агентами
- ▶ Пусть 2 агента решают задачу бинарной классификации
- ▶ Пусть у 1-го агента доля объектов первого класса в обучающей выборке α , и, соответственно $1 - \alpha$ – второго
- ▶ Если агенты поделятся данными для коммуникационной выборки, то она будет гомогенной

Гетерогенность

Качество моделей в условии гетерогенности при уменьшении размера обучающих данных



Дальнейшее исследование

- ▶ Подробнее изучить постановку с гетерогенностью (*тюнить параметры коммуникации, реализовать мультиклассовую постановку*)
- ▶ Рассмотреть коммуникации не по схеме "каждый с каждым"
- ▶ Попробовать провести обучение в других условиях (*поменять датасет, модель*)

Спасибо за внимание!