

The 9th International Conference
on Deep Learning in Computational Physics
4 июля 2025 г.

Применение концепции переноса обучения для градиентного бустинга при решении обратных задач разведочной геофизики

Шалеев М.К. ¹, Исаев И.В. ², Оборнев И.Е. ², Доленко С.А. ²

(1) Физический факультет МГУ им. М.В. Ломоносова, Москва

(2) НИИ ядерной физики им. Д.В. Скобельцына, Москва

* Исследование выполнено за счёт гранта Российского научного фонда, проект № 24-11-00266,
<https://rscf.ru/project/24-11-00266/>.

Обратные задачи разведочной геофизики

Обратные задачи разведочной геофизики (ОЗ РГ) – задачи восстановления пространственного распределения свойств среды в толще Земли по геофизическим полям, измеренным на её поверхности.

Особенности ОЗ РГ :

- Некорректно поставлены по Адамару
- Имеют высокую размерность
- Имеют численное решение прямой задачи

Геофизические методы

В рамках данной работы изучаются обратные задачи гравиметрии (Г), магнитометрии (М) и магнитотеллурического зондирования (МТЗ), которые заключаются в определении распределения плотности, намагниченности и удельного электрического сопротивления соответственно.

Используются значения гравитационных и электромагнитных полей, измеренные на поверхности Земли

Методы решения ОЗ РГ

Традиционные методы решения – **методы оптимизации**, основанные на многократном решении прямой задачи с минимизацией невязок в пространстве наблюдаемых полей

- ❑ Высокая вычислительная стоимость
- ❑ Необходимо задать хорошее первое приближение
- ❑ Малая невязка в пространстве наблюдаемых величин **не гарантирует** малой невязки в пространстве определяемых параметров

Методы решения ОЗ РГ

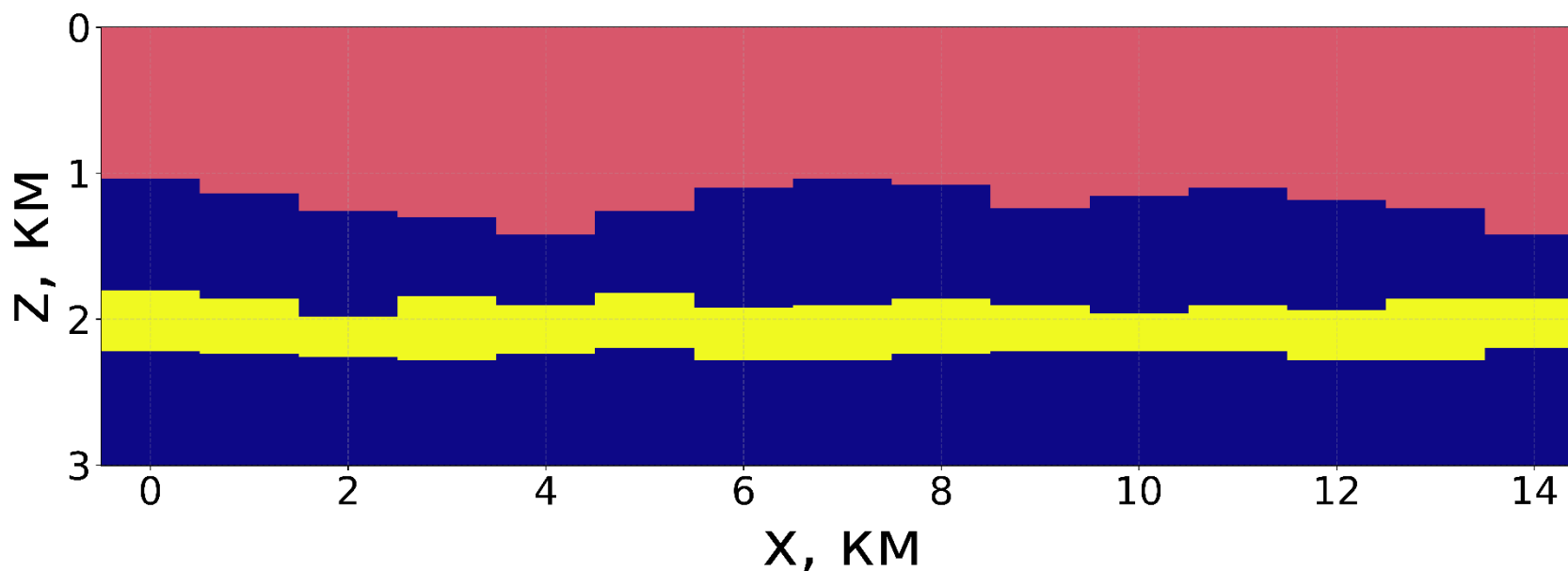
В исследовании использовались **методы машинного обучения**:

- ❑ Высокие вычислительные затраты смещаются от стадии применения вычислительной системы к стадии ее разработки, что увеличивает удобство практического использования
- ❑ Возможность одновременно использовать данные нескольких геофизических методов для улучшения качества решения

Введение

Структура разреза

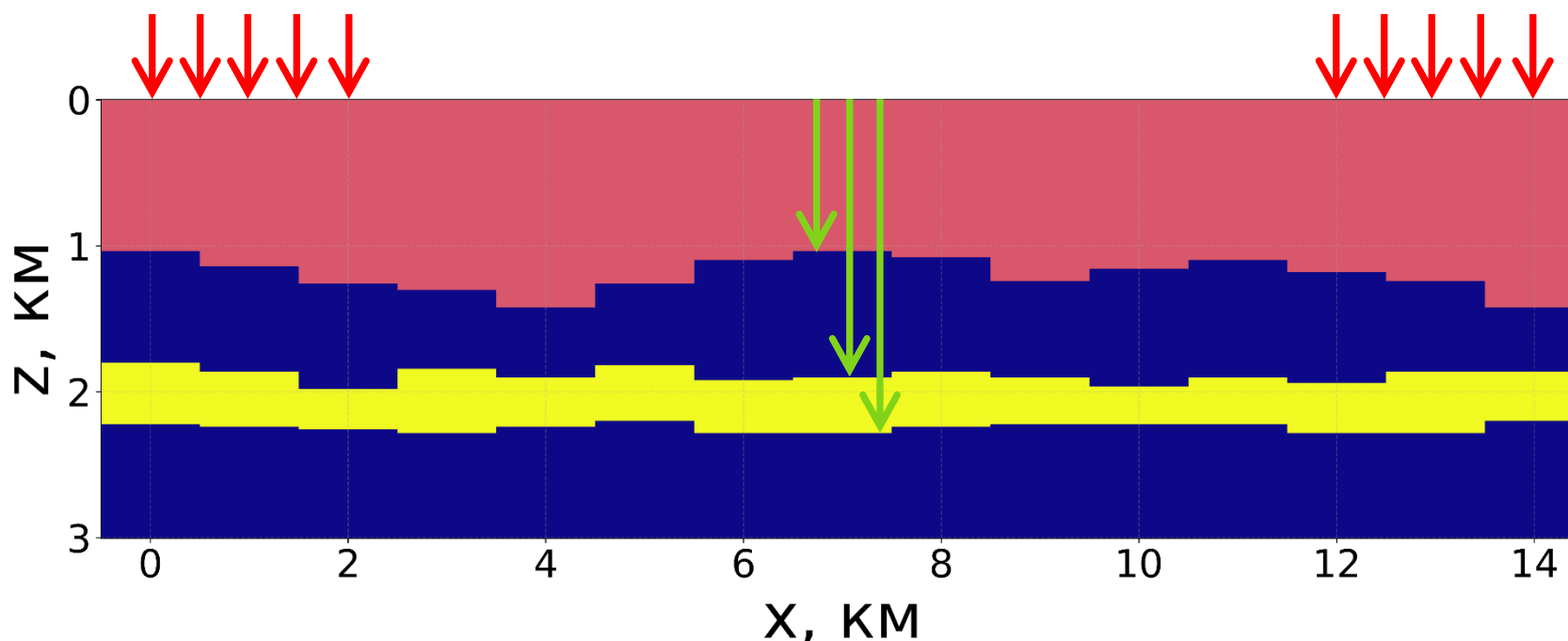
- 2D-модель
- 4 слоя
- По глубине - 3 км
- По ширине -15 км
- Разрез Норильского региона



Вычислительный эксперимент

Постановка

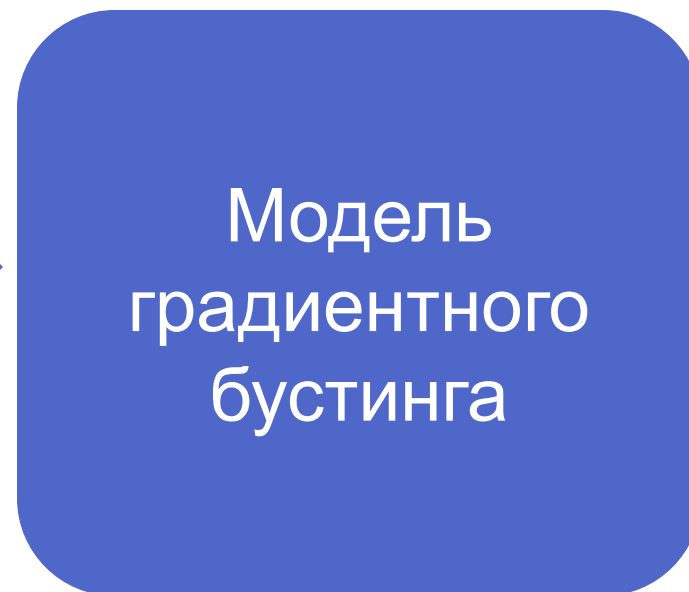
- Для совместного использования нескольких геофизических методов была выбрана **геометрическая постановка** задачи, в которой определялись **глубины границ слоёв**
- Была применена методика **автономного** определения параметров



Вычислительный эксперимент

Постановка

Значения
геофизических
полей



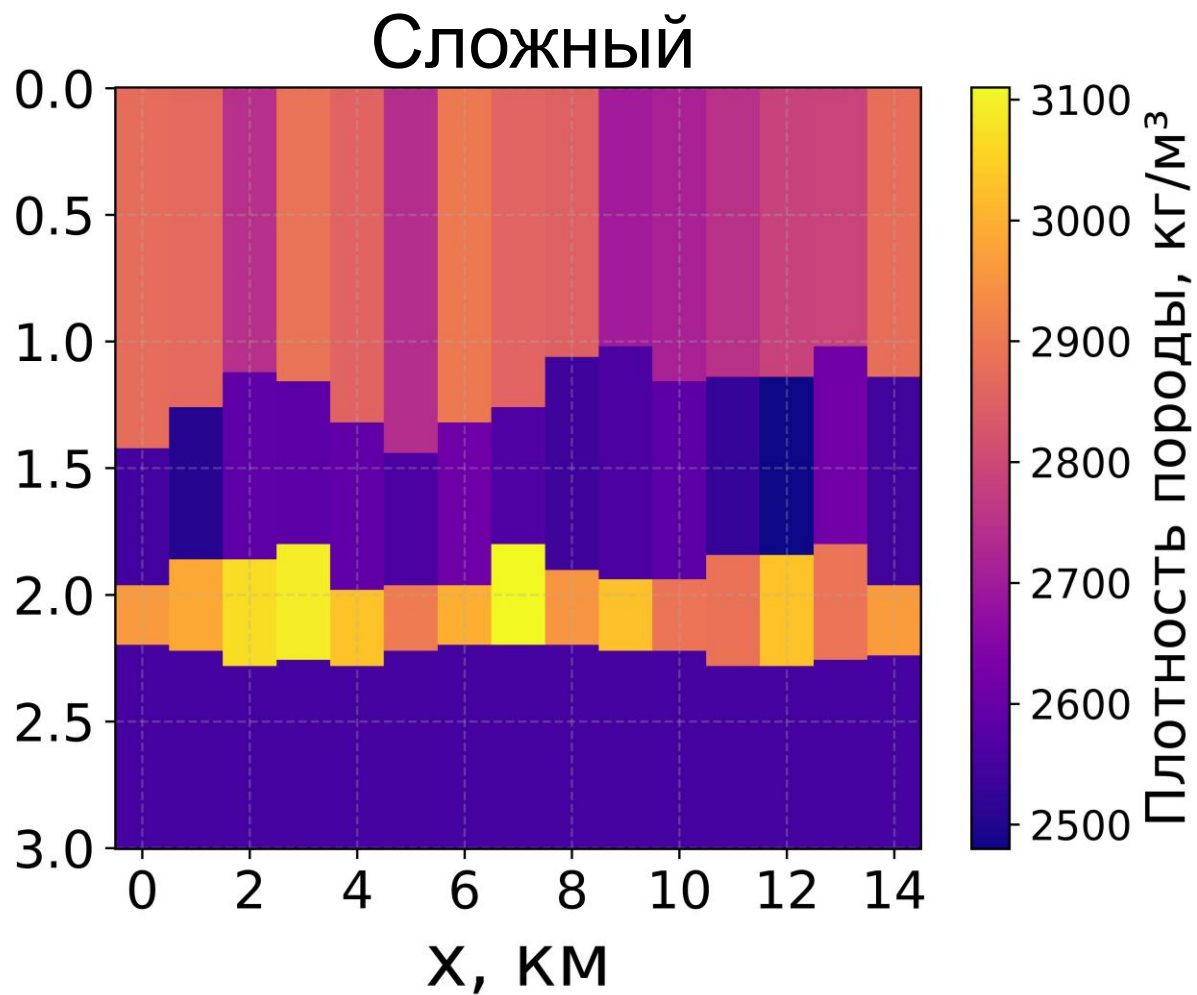
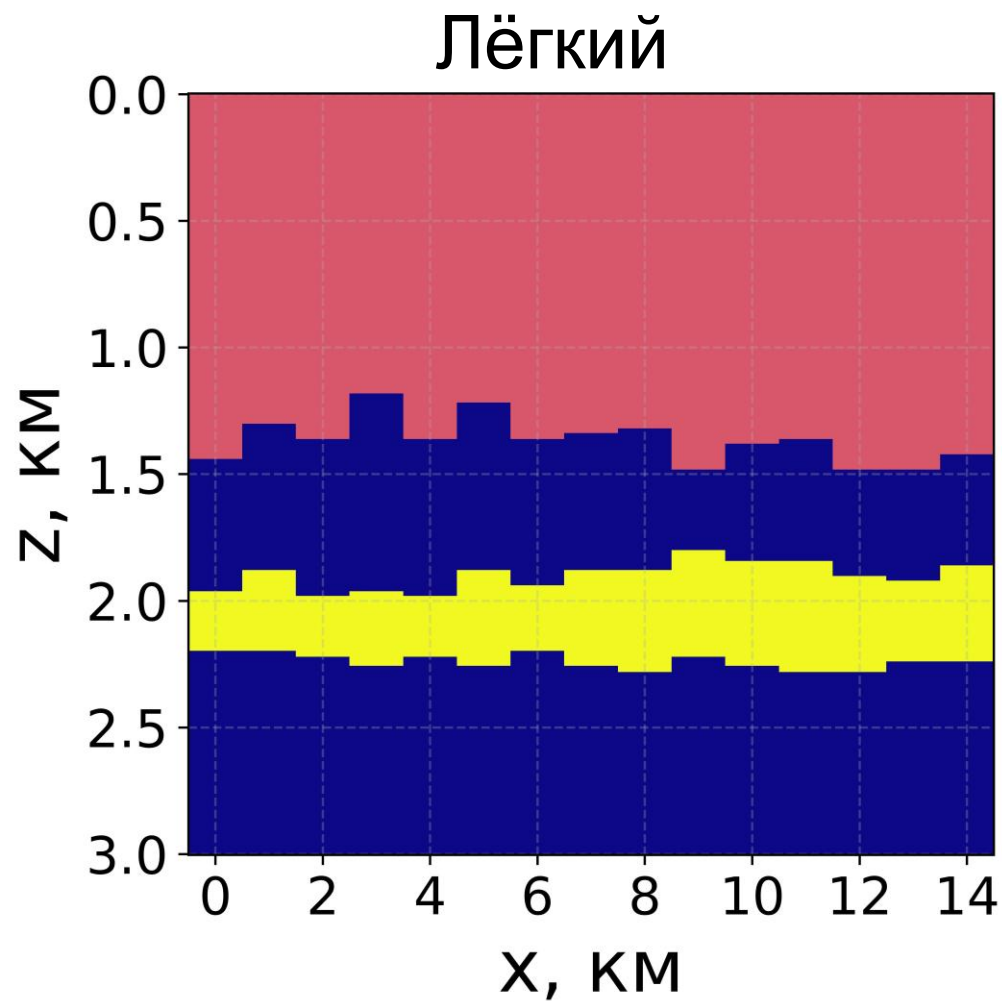
Глубина границы слоя

В работе использовались 3 набора данных разной сложности:

- **Лёгкий (Л)** - каждый слой характеризуется постоянными значениями плотности, намагниченности и удельного сопротивления, которые не изменяются внутри слоя и одинаковы для всех примеров.
- **Средний (Ср)** - физические свойства слоев остаются постоянными внутри каждого слоя, но варьируются от примера к примеру в пределах $\pm 10\%$ от значений, заданных в первой схеме.
- **Сложный (Сл)** - физические свойства изменяются как между примерами, так и внутри каждого слоя

Вычислительный эксперимент

Постановка



Вычислительный эксперимент

Цель работы

Проблематика: необходимость для каждой **новой геологической структуры** генерировать новый набор синтетических примеров и обучать новый комплект моделей градиентного бустинга.

Для снижения вычислительных затрат и повышения качества решения в условиях ограниченных ресурсов предлагается использовать подход, основанный на **переносе обучения**

Вычислительный эксперимент

Постановка

Модели **предварительно** обучаются на данных одного из наборов, а затем настраиваются на данных из другого набора.

Рассматриваются различные размеры наборов для финальной настройки модели

Набор данных

Характеристики

Исходное разбиение на поднаборы:

- | | | | |
|-----------------------|-----|-------|----------|
| • Тренировочный набор | 70% | 7 000 | примеров |
| • Валидационный набор | 20% | 2 000 | примеров |
| • Тестовый набор | 10% | 1 000 | примеров |

Размер тренировочного (валидационного) наборов **снижался**:

3500(1750), 1 750(500), 1000(250),
700(200), 350(100), 175(50), 70(20)

Градиентный бустинг

Параметры переноса обучения

В исследовании были изучены 4 подхода к переносу на градиентом бустинге:

- **Прямое дообучение:** обучаем исходную модель, далее продолжаем обучения с целевыми данными
- **Предварительное усечение:** перед дообучением исходная модель усекается до лучшей итерации по валидационному набору целевой выборки

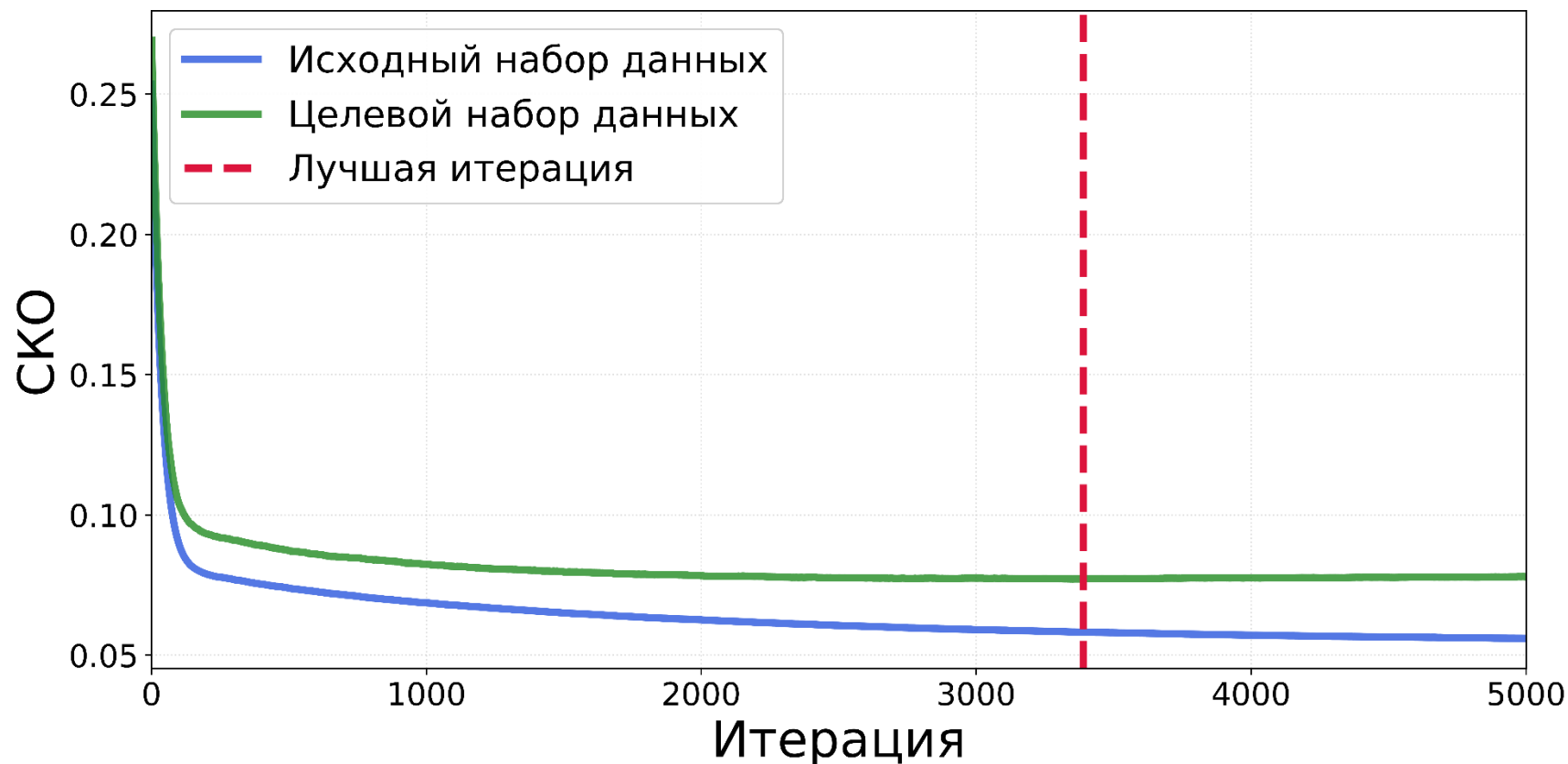
Параметры переноса обучения

- **Признак из предсказания:** в входные признаки целевой модели добавляем ответ исходной модели. Целевую модель обучаем с нуля
- **Взвешенная комбинация:** обучаем обе модели с нуля на своих наборах данных. Финальным предсказанием является взвешенная сумма с коэффициентом α

Градиентный бустинг

Предварительное усечение

Исходная модель усекает до лучшей итерации по валидационному набору целевой выборки. Оставляя только полезное «ядро» модели



Градиентный бустинг

Взвешенная комбинация

Обе модели (исходная и целевая) обучаются с нуля.

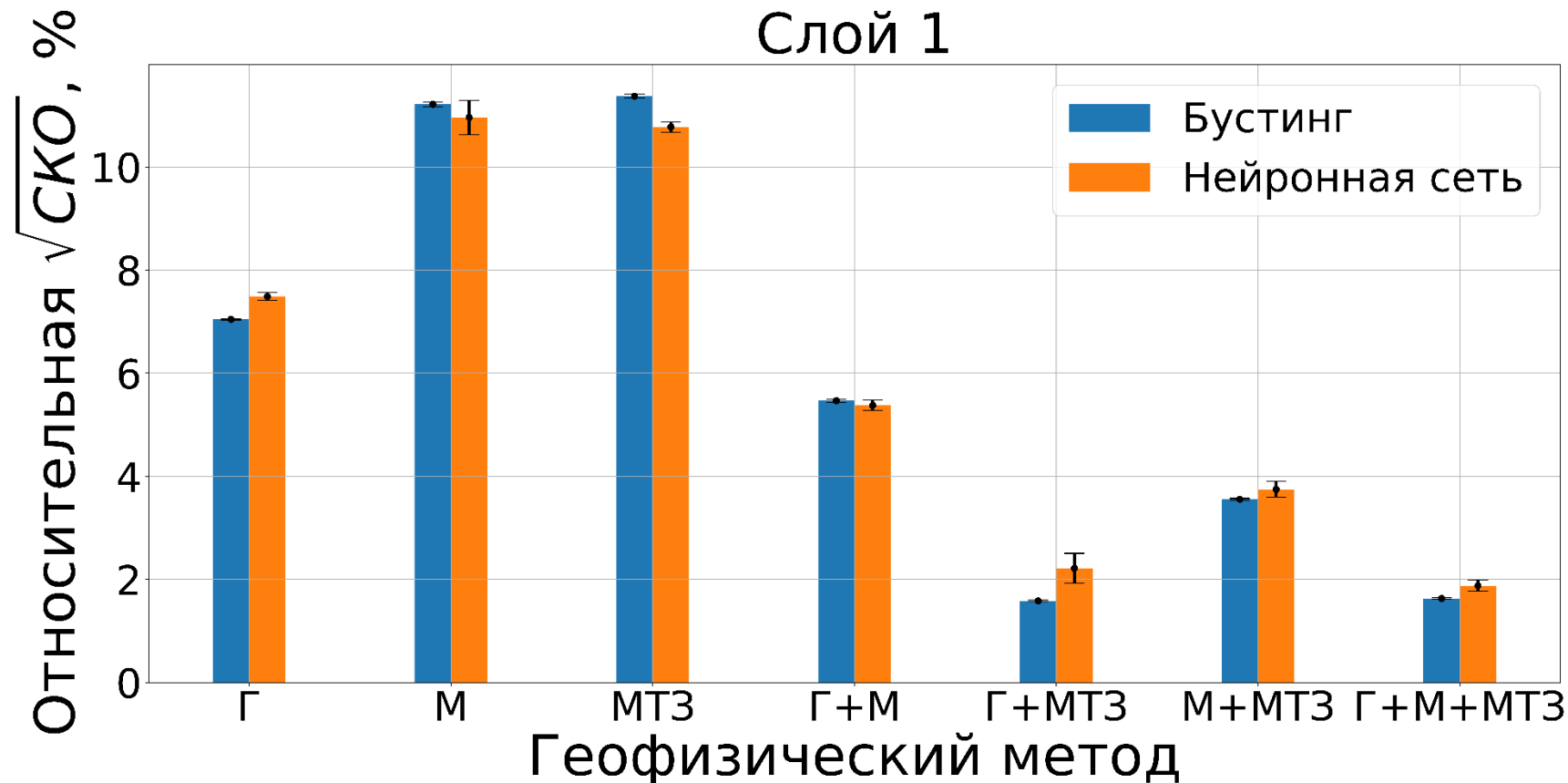
В качестве ответа берём взвешенную сумму от ответов моделей:

$\alpha \times \text{исходная} + (1 - \alpha) \times \text{целевая}$, где α подбирается для каждого случая отдельно с помощью симплекс-метода оптимизации

Результаты

Сравнение с нейронными сетями

Градиентный бустинг демонстрирует точность, конкурирующую с нейронными сетями, при этом значительно превосходя их по скорости обучения

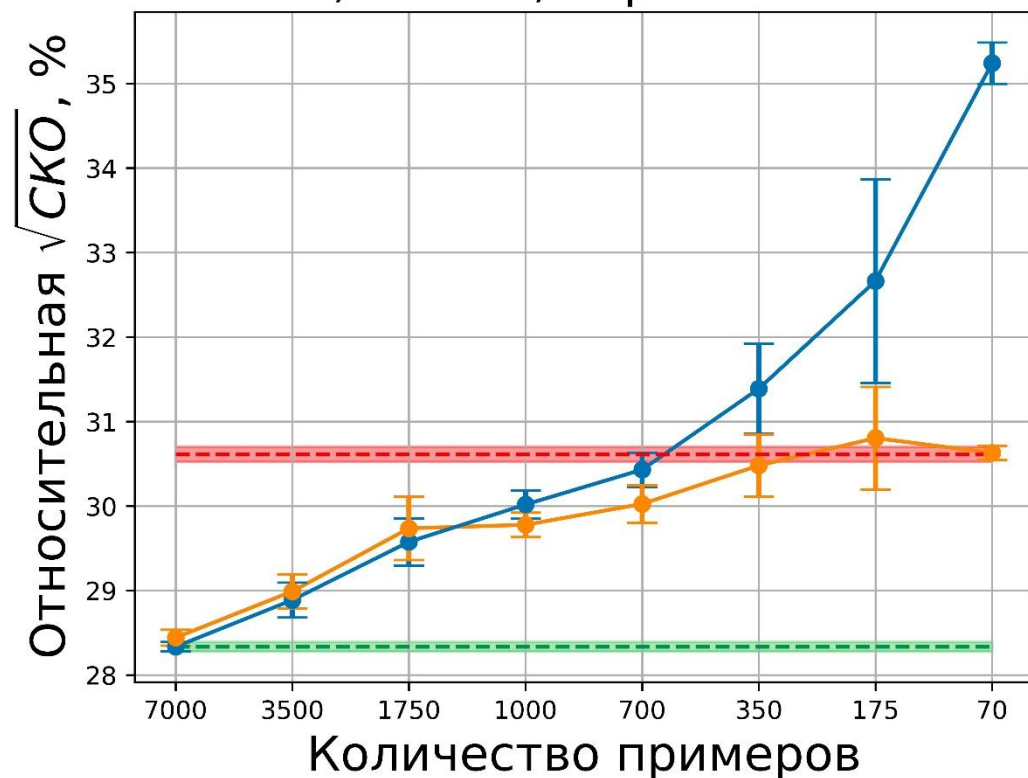


Стратегии с **прямым дообучением** и **усечением** показали сопоставимые результаты. Обрезка исходной модели до лучшей итерации не привела к значительному улучшению качества. Стратегии продемонстрировали высокую устойчивость к переобучению, что выгодно отличает их от нейронных сетей

Результаты

Прямое дообучение

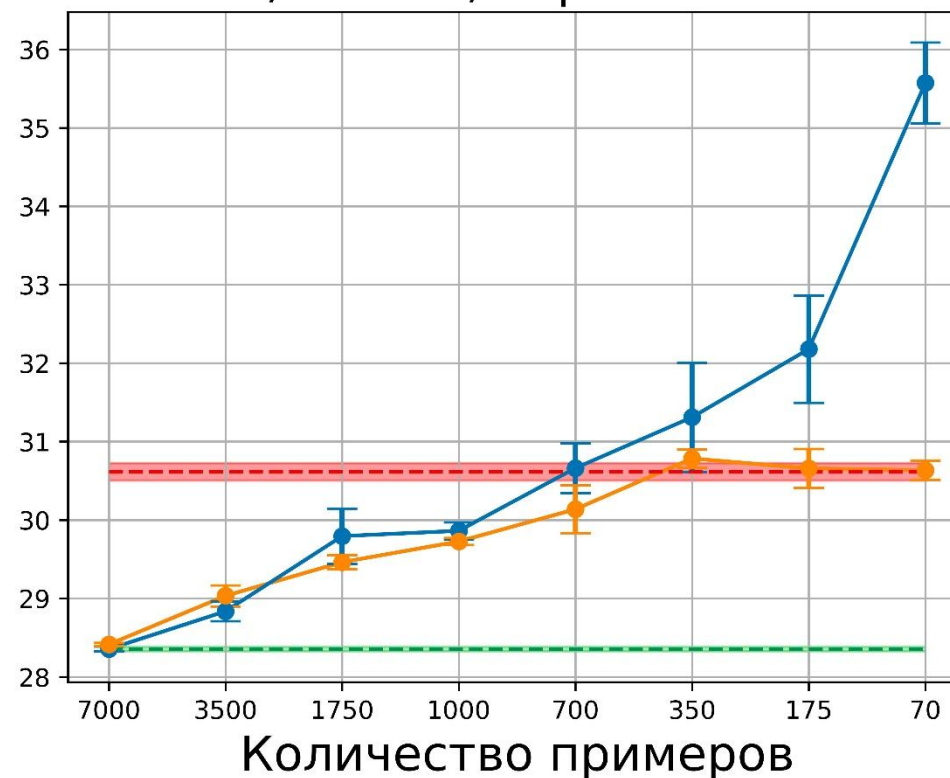
М+МТЗ, Слой 3, перенос с Сл на Л



--- до переноса обучения
--- базовое решение

Предварительное усечение

М+МТЗ, Слой 3, перенос с Сл на Л



—●— перенос обучения
—●— без переноса обучения

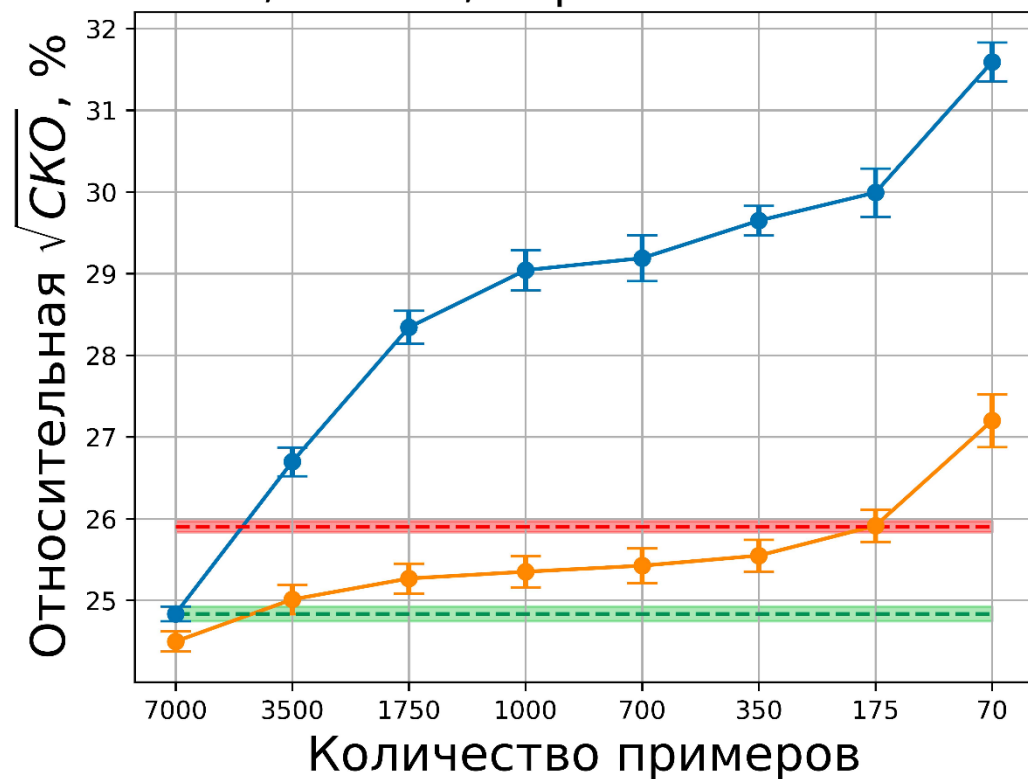
Результаты

- ❑ Обучение целевой модели **с нуля на расширенных признаках** продемонстрировало склонность к переобучению, особенно при малом объеме целевых данных.
- ❑ **Взвешенная комбинация** показала сопоставимую с базовыми стратегиями точность и стандартную устойчивость к переобучению.
- ❑ Стратегия **Признак из предсказания** требует контроля переобучения из-за обучения новой модели, в то время как **Взвешенная комбинация** обеспечивает предсказуемо хороший результат.

Результаты

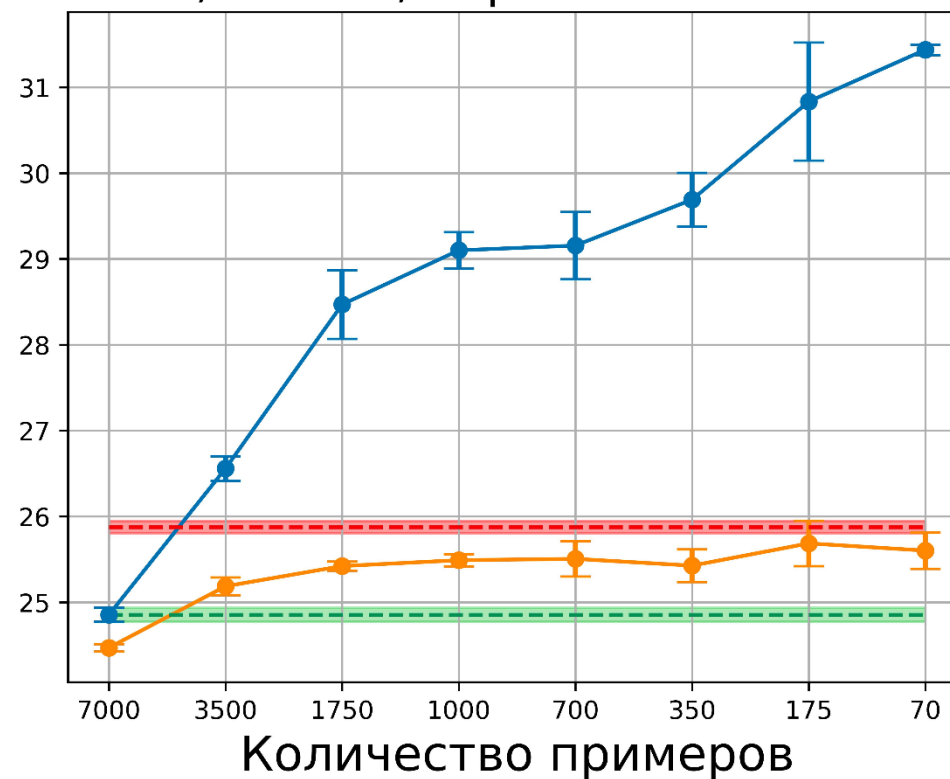
Признак из предсказания

М, Слой 2, перенос с Л на Сл



Взвешенная комбинация

М, Слой 2, перенос с Л на Сл



● перенос обучения
● без переноса обучения

- ❑ **Градиентный бустинг** конкурирует с нейронными сетями по точности решения ОЗ РГ, при этом превосходя их по скорости обучения
- ❑ Применение стратегий переноса обучения к ГБ позволяет **повысить качество** решения задач в условиях ограниченных данных
- ❑ Стратегия **Признак из предсказания** продемонстрировала склонность к переобучению при малом объеме данных
- ❑ Остальные стратегии демонстрируют высокую устойчивость к переобучению, являясь надежным выбором

Спасибо за внимание!

Вычислительный эксперимент

Постановка

№ слоя	Описание	Физические свойства			Пространственные свойства		
		Плотность σ , кг/м ³	Намагниченность μ , А/м	Удельное эл. сопр-ние ρ , Ом·м	Верхняя граница, min-max, км	Нижняя граница, min-max, км	Мощность слоя, min-max, км
1	Базальтовая толща	2 800 2 520 – 3 080	3 2.7 – 3.3	2 000 1 800 – 2 200	0	1 – 1.48	1 – 1.48
2	Терригенно-карбонатные отложения	2 550 2 295 – 2 805	0.5 0.45 – 0.55	100 90 – 110	1 – 1.48	1.8 – 1.98	0.32 – 0.98
3	Габбро-долериты массивные медно-никелево-платиновые руды	3 000 2 700 – 3 300	0.9 0.81 – 0.99	1 000 900 – 1 100	1.8 – 1.98	2.2 – 2.28	0.22 – 0.48
4	Терригенно-карбонатные отложения	2 550	0.5	100	2.2 – 2.28	—	—

Набор данных

Характеристики

❑ Размерность данных

- По выходу

✓ *45 параметров = 3 слоя * 15 значений глубин границ слоев*

- По входу

✓ Гравиметрия (G):

*31 признак = 1 компонента поля * 31 точка измерения*

✓ Магнитометрия (M):

*31 признак = 1 компонента поля * 31 точка измерения*

✓ МТЗ (MT):

*62 признака = 2 компоненты поля * 1 частота *
31 точка измерения

❑ Количество примеров

✓ *10 000 примеров*