С.А. Доленко

НИИЯФ МГУ

Лаборатория адаптивных методов обработки данных (ЛАМОД)

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РЕШЕНИЯ МНОГОПАРАМЕТРИЧЕСКИХ ОБРАТНЫХ ЗАДАЧ

“…она поспела как раз вовремя, чтобы заметить, как Белый Кролик скрылся в большой норе под колючей изгородью.

В ту же секунду Алиса не раздумывая ринулась за ним. А кой о чем подумать ей не мешало бы – ну хоть о том, как она выберется **обратно**!”

Льюис Кэрролл. Алиса в стране чудес

Научный семинар «Глубокое обучение в вычислительной физике»

13 мая 2021 г.

ОБРАТНЫЕ ЗАДАЧИ

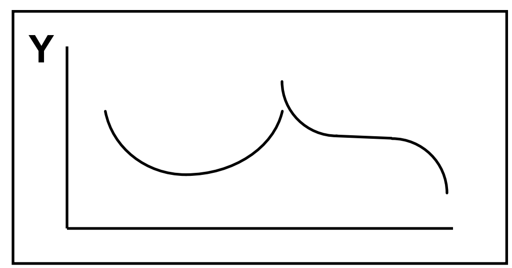
Общая постановка обратной задачи:

$X = (x_1, x_2, \ldots x_n)$ → **Исследуемый объект** $Y = F(X)$ → $Y = (y_1, y_2, \ldots y_m)$

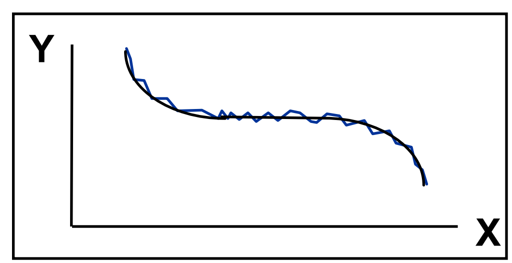
<u>Наблюдаемыми</u> являются значения Y; как правило, $m \gg n$

Задача: научиться восстанавливать значения X по значениям Y

Такая задача представляет собой задачу моделирования обратной функции $X = F^{-1}(Y)$



Проблемы:
- **единственность решения**
- **устойчивость решения**



Неединственность и неустойчивость:
- **теоретические**
- **практические**

КОРРЕКТНОСТЬ ОБРАТНЫХ ЗАДАЧ

Обратная задача называется **корректной по Адамару**, если:

1) $\forall Y \ \exists \ !X: Y=F(X)$ (решение единственно)

2) $\forall \delta \ \exists \varepsilon : \ (\Delta Y < \delta) \Leftrightarrow (\Delta X < \varepsilon)$ (решение устойчиво)

во всей области определения X.

Причины практической некорректности:

- наличие шумов в экспериментальных данных (неустойчивость)
- дискретность набора экспериментальных точек (неединственность)

Jacques Hadamard
(1865 – 1963)

КОРРЕКТНОСТЬ ОБРАТНЫХ ЗАДАЧ

Обратная задача называется корректной по Тихонову (в случае нарушения корректности по Адамару), если в пространстве {X} можно выделить более узкое множество решений {X'} такое, что:

1) *a priori* известно, что $\exists X \in \{X'\}$

(решение существует)

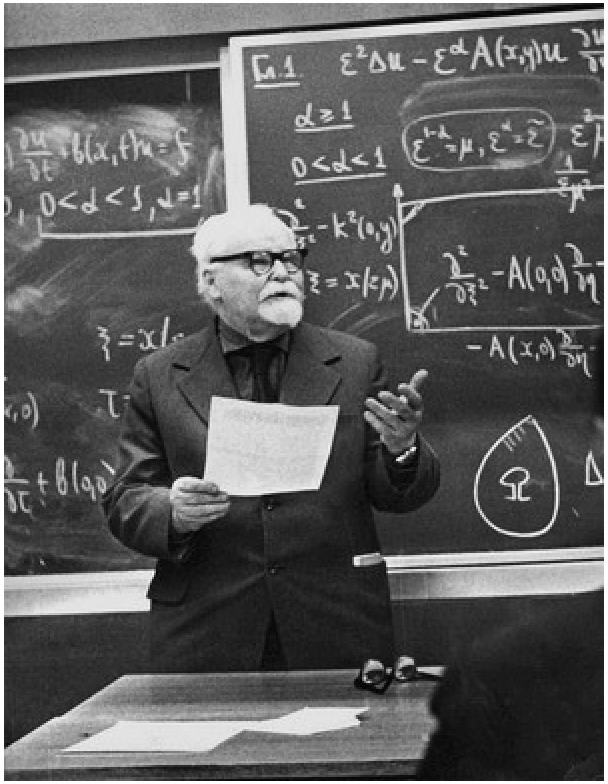
2) $\forall Y \; \exists \; !X \in \{X'\} : Y=F(X)$

(решение единственно)

3) $\forall \delta \; \exists \varepsilon : \; (\Delta Y < \delta) \Leftrightarrow (\Delta X < \varepsilon)$,

если $(X+\Delta X) \in \{X'\}$

(решение устойчиво).



Акад. Тихонов Андрей Николаевич
(1906 – 1993)

НЕКОРРЕКТНОСТЬ И ПЛОХАЯ ОБУСЛОВЛЕННОСТЬ

Задача называется **хорошо обусловленной**, если
малым погрешностям исходных данных
соответствуют малые погрешности решения

Задача называется **плохо обусловленной**, если
малым погрешностям исходных данных
соответствуют большие погрешности решения

Абсолютное число обусловленности μ_Δ: $\Delta y \le \mu_\Delta \cdot \Delta x$

Относительное число обусловленности: μ_δ: $\delta y \le \mu_\delta \cdot \delta x$

Δ – абсолютное изменение, δ – относительное изменение

Чем меньше число обусловленности, тем лучше

Корректная задача может быть плохо обусловленной!

Некорректная задача имеет число обусловленности, равное ∞

ПОСТАНОВКА ОБРАТНЫХ ЗАДАЧ С ТОЧКИ ЗРЕНИЯ МЕТОДОВ ОБРАБОТКИ ДАННЫХ

Одну и ту же обратную задачу (ОЗ) можно поставить несколькими разными способами :

1) Как задачу *оценки*

(оценка значений обратной функции $X = F^{-1}(Y)$)

Используются методы восстановления зависимостей, в том числе и методы машинного обучения (ММО)

2) Как задачу *оптимизации*

(подбор оптимальных параметров прямой функции $Y=F(X)$)

Используются методы (многомерной) оптимизации

3) Как задачу *классификации*

(если вектор параметров X принимает дискретные значения)

Используются методы решения задач классификации (распознавания образов)

МЕТОДИЧЕСКИЕ ПОДХОДЫ К РЕШЕНИЮ ОБРАТНЫХ ЗАДАЧ С ПОМОЩЬЮ ММО

НИИЯФ МГУ

1. Подход "от эксперимента"

Задан экспериментальный массив данных для обучения

$S_k = (X, Y)_k$, $k = 1 \ldots N$

Значения искомых параметров известны *a priori* или определены экспериментально альтернативными методами

а) Разделение массива на тренировочный, валидационный и тестовый наборы

б) Обучение метода оценке значений обратной функции $X(Y)$

2. Подход "от модели"

Требует наличия аналитической или вычислительной модели прямой функции $Y = F(X)$

а) Задание диапазона изменения X и "сетки" значений

б) Генерация необходимого количества данных с необходимой представительностью для всех наборов посредством решения прямой задачи

в) Обучение метода оценке значений обратной функции $X(Y)$

МЕТОДИЧЕСКИЕ ПОДХОДЫ К РЕШЕНИЮ ОБРАТНЫХ ЗАДАЧ С ПОМОЩЬЮ ММО

3. Промежуточный “квазимодельный” подход

Используется интерполяционная или экстраполяционная генеративная модель, основанная на “опорных” векторах наблюдаемых значений, получаемых экспериментально

Области применения методических подходов

Имеется ли адекватная аналитическая/численная модель объекта?	Да	Нет	Нет
Имеется ли достаточное количество экспериментальных данных?		Да	Нет
Рекомендуемый методический подход	“от модели”	“от эксперимента”	“квазимодельный”

НИИЯФ МГУ

Сравнение свойств методических подходов

Характеристика	"От эксперимента"	"Квазимодельный"	"От модели"
Представитель-ность наборов данных	Как правило, недостаточная	Адекватная	Адекватная
Наличие шумов в данных	Присутствуют всегда	Присутствуют всегда	Могут быть введены искусственно
Уровень шумов в данных	Определяется экспериментом	Определяется экспериментом	Задается при моделировании
Проведение экспериментов	Необходимо, масштабное	Необходимо, ограниченное	Желательно для проверки
Наличие модели решения ПЗ	Не требуется	Требуется "квазимодель"	Требуется
Адекватность решения объекту	Высокая	В меру адекватности "квазимодели"	В меру адекватности модели

СВОЙСТВА НЕЙРОННЫХ СЕТЕЙ КАК МЕТОДА РЕШЕНИЯ ОБРАТНЫХ ЗАДАЧ

- **Обучение на примерах**
- **Устойчивость к противоречивым данным**
- **Устойчивость к шумам**
- **Устойчивость к умеренному изменению побочных параметров объекта, не являющихся искомыми в процессе решения обратной задачи**

ОТЛИЧИЯ НЕЙРОННЫХ СЕТЕЙ ОТ ДРУГИХ АЛГОРИТМОВ РЕШЕНИЯ ОБРАТНЫХ ЗАДАЧ

- Основаны на **аппроксимации** обратной функции
- Не требуют **априорного знания** о характере искомой зависимости (наличия точной модели)
- Требуют наличия **представительного** набора данных
- Обладают высокой практической **устойчивостью**
- Основываются в своей оценке на информации, получаемой из **всей** обучающей выборки
- Используют в качестве целевой функции в оптимизационной процедуре
  расстояние непосредственно
  в пространстве искомых параметров,
  а не в пространстве наблюдаемых величин
  (минимумы невязки в разных пространствах
  не соответствуют друг другу!)
- Проблема наличия **ошибок**

РЕКОМЕНДУЕМЫЕ ОБЛАСТИ ПРИМЕНЕНИЯ ММО КАК МЕТОДОВ РЕШЕНИЯ ОБРАТНЫХ ЗАДАЧ

Аппроксимационные методы (ММО) следует применять, когда:

- **Обратная задача не может быть решена аналитически или численно**
- **Данные противоречивы или характеризуются значительным уровнем шума**
- **Аналитическая или вычислительная модель решения прямой задачи отсутствует или неустойчива**
- **Априорная информация об объекте отсутствует, или её использование затруднено**

АППРОКСИМАЦИОННЫЕ МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ, ПРИМЕНЯЕМЫЕ ДЛЯ РЕШЕНИЯ ОЗ

НИИЯФ МГУ

Некоторые аппроксимационные методы машинного обучения, эффективные при решении обратных задач:

- Искусственные нейронные сети (далеко не всегда глубокие!) ИНС – универсальный аппроксиматор!
- Градиентный бустинг
- Алгоритм случайного леса
- Линейная регрессия в нелинейном базисе
- Метод частичных наименьших квадратов (метод проекций на латентные структуры, ПЛС)
- Метод группового учёта аргументов (МГУА)

ПОНИЖЕНИЕ ВХОДНОЙ РАЗМЕРНОСТИ ДАННЫХ ПРИ РЕШЕНИИ ОЗ ВЫСОКОЙ РАЗМЕРНОСТИ

Высокая размерность входных данных сильно снижает эффективность применения аппроксимационных методов.

Основные способы понижения входной размерности:

- **Отбор** существенных входных признаков, напр., с помощью:
  - Кросс-корреляционного и кросс-энтропийного анализа
  - Оценки дисперсии значений входного признака
  - Анализа весов нейронной сети
- **Преобразование** входного признакового пространства, напр.:
  - Анализ главных компонент, линейный и нелинейный
  - Агрегация признаков
  - Вейвлет-анализ

СПОСОБЫ ОПРЕДЕЛЕНИЯ ВЫХОДНЫХ ПАРАМЕТРОВ МНОГОПАРАМЕТРИЧЕСКИХ ОБРАТНЫХ ЗАДАЧ

- **Одновременное** определение (1 модель, N выходов)
- **Автономное** определение (N моделей с 1 выходом каждая)
- **Групповое** определение (объединение параметров в группы с одновременным определением внутри каждой группы)
- **Поэтапное** определение (на начальных этапах определяются те параметры, для которых задача решается хорошо;
  на последующих этапах их полученные значения подаются на вход вместе со значениями входных признаков.

Наиболее эффективным является **групповое** определение при наличии возможности правильного объединения параметров.

ПОВЫШЕНИЕ УСТОЙЧИВОСТИ МОДЕЛЕЙ ММО ПО ОТНОШЕНИЮ К ШУМАМ РАЗЛИЧНОЙ ПРИРОДЫ

НИИЯФ МГУ

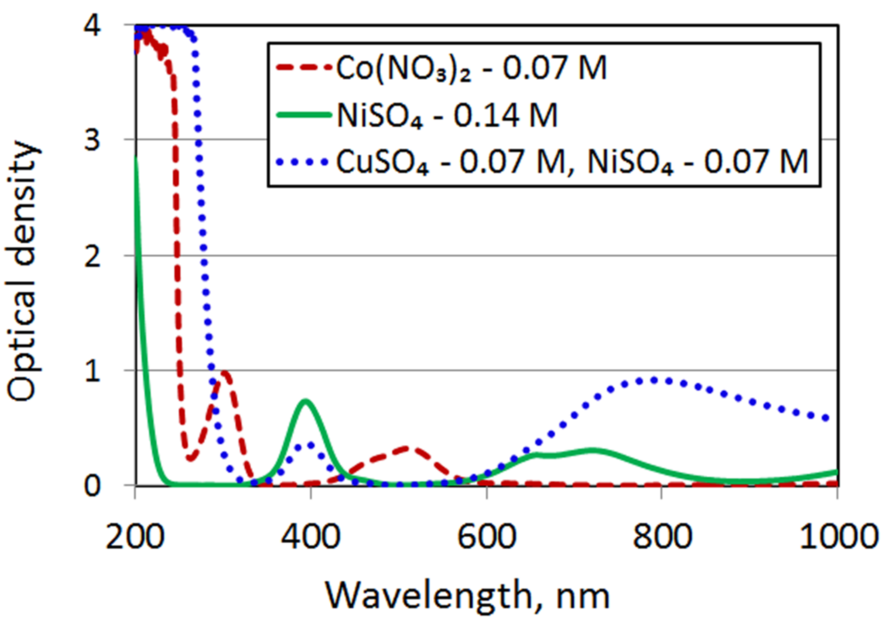
- Эффективный приём: **добавление шума** к обучающим примерам в процессе обучения модели
- Параметры **добавляемого** шума должны соответствовать аналогичным параметрам шума в **тестовом** наборе:
  - **Природа** шума (способ влияния на данные – напр. аддитивный, мультипликативный, «корневой»)
  - **Статистика** шума (распределение реализаций в пределах амплитуды (напр. равномерный, гауссов)
  - **Амплитуда** шума (максимальная или характерная величина)

КОМПЛЕКСИРОВАНИЕ МЕТОДОВ КАК СПОСОБ КОМПЕНСАЦИИ ОШИБОК РЕШЕНИЯ ОЗ

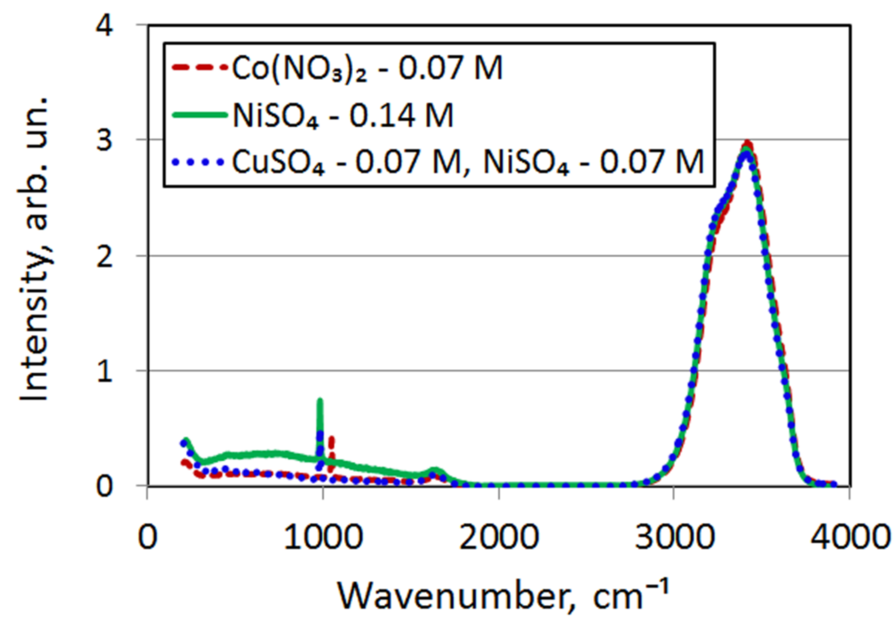
- Комплексирование **аппроксиматоров** (комитеты / ансамбли)
  - Основано на **нескоррелированности** ошибок
  - Максимальное **разнообразие** предикторов
  - **Раздельные или совместные** отбор или преобразование признаков
- Комплексирование **данных** (бэггинг)
  - **Независимость** всех поднаборов
  - Сохранение **представительности** всех поднаборов
- Комплексирование **физических методов**
  - Согласованное решение нескольких **различных** ОЗ с одними и теми же определяемыми параметрами
  - **Раздельное или совместное** решение ОЗ
  - **Раздельные или совместные** отбор или преобразование признаков

Пример 1: Обратные задачи спектроскопии

- **Объект исследования**: многокомпонентные водные растворы неорганических солей тяжёлых металлов $CuSO_4$, $NiSO_4$, $CoSO_4$, $Cu(NO_3)_2$, $Ni(NO_3)_2$, $Co(NO_3)_2$
- В каждом растворе от 0 до 6 солей, от 0 до 5 ионов
- Для каждого раствора измерялось также pH
- Спектры поглощения и спектры комбинационного рассеяния



Поглощение

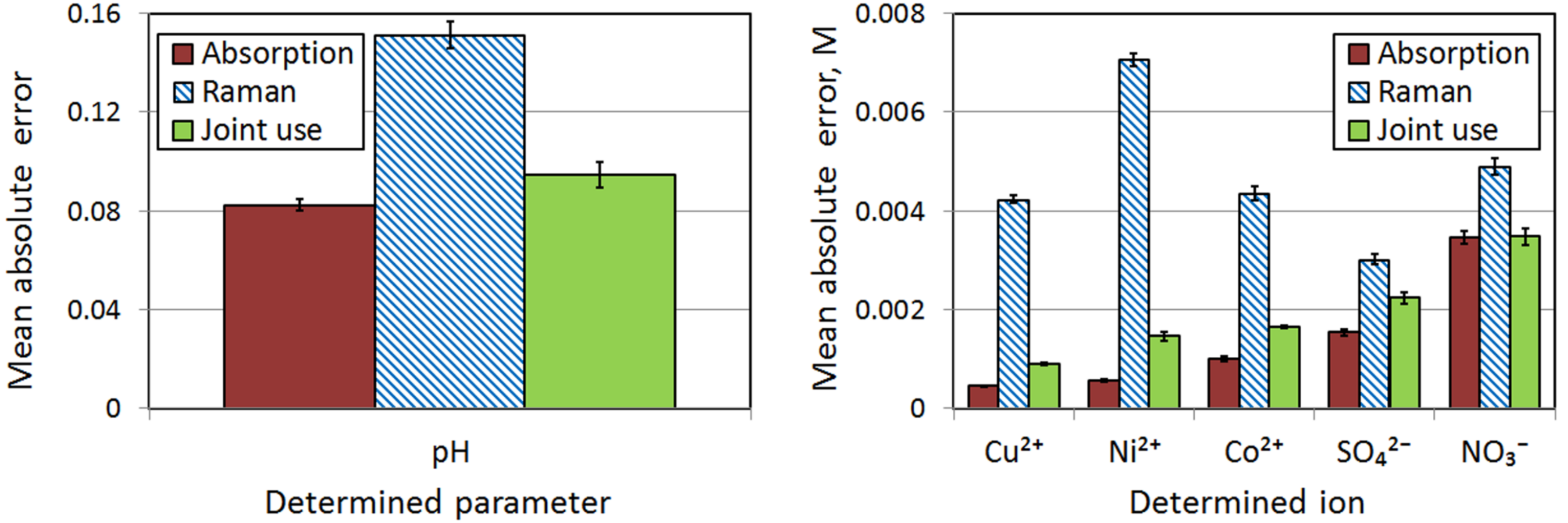


Комбинационное рассеяние

Пример 1: Методика работы с данными

- **3806 примеров** (растворов) = 2656 трн + 750 влд + 400 тст
- **2048 каналов** в спектре комбинационного рассеяния (КР)
- **811 каналов** в спектре поглощения
- Входные признаки: только поглощение, только КР, оба
- **pH** использовался как входной либо как выходной признак
- Автономное определение
  (отдельная сеть для каждого определяемого параметра)
- Многослойный персептрон с 1…3 скрытыми слоями («мелкий»)
- Однородный комитет из 5 идентичных сетей с разной инициализацией весов
- Исходный набор, отбор, преобразование признаков

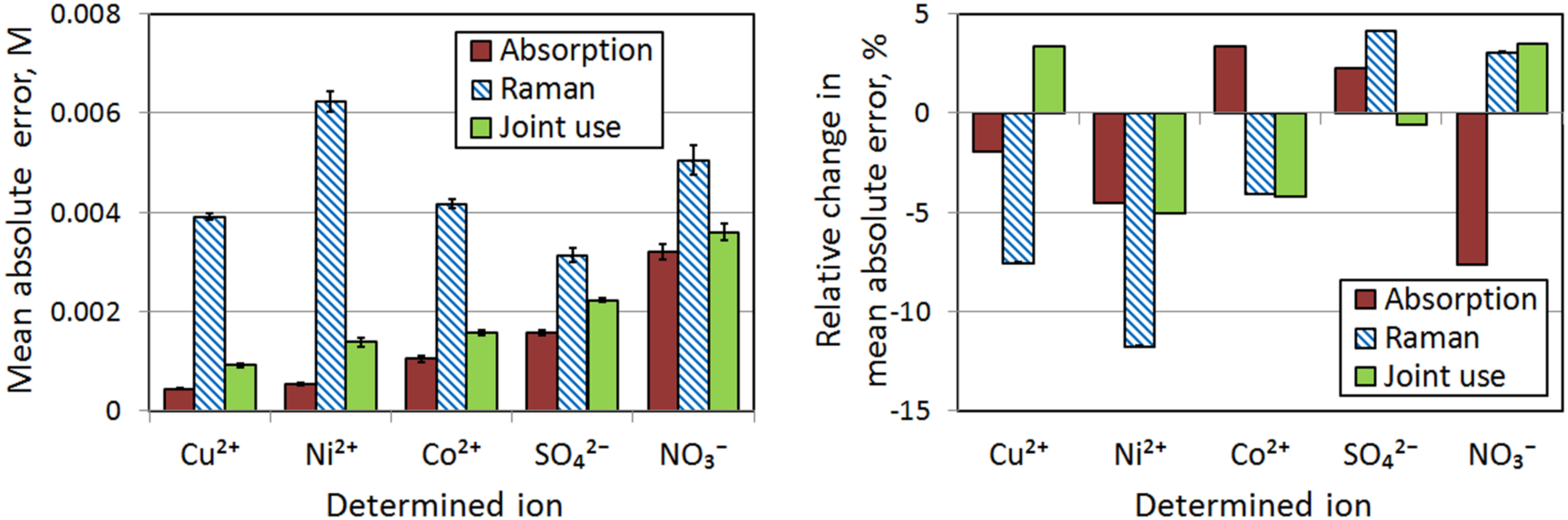
Пример 1: Базовый эксперимент



Базовый эксперимент (полные наборы входных признаков)

При совместном использовании на вход подавались
оба массива входных признаков параллельно

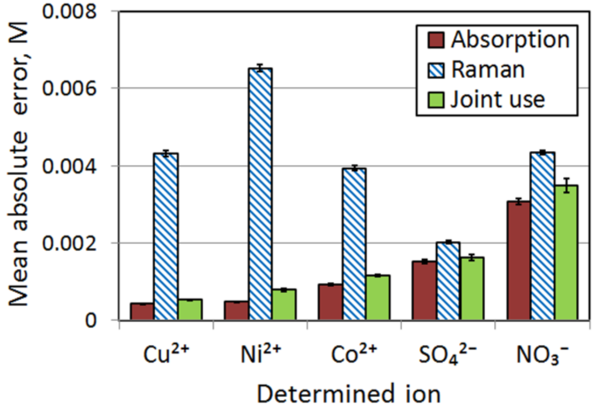
Пример 1: Использование pH

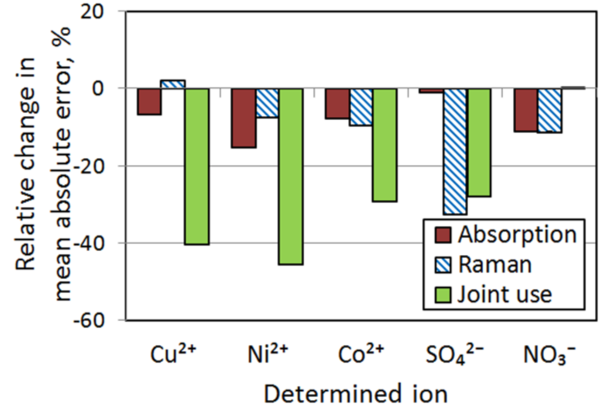


Результаты включения pH в число входных признаков

Справа – относительное изменение в % средней абсолютной ошибки

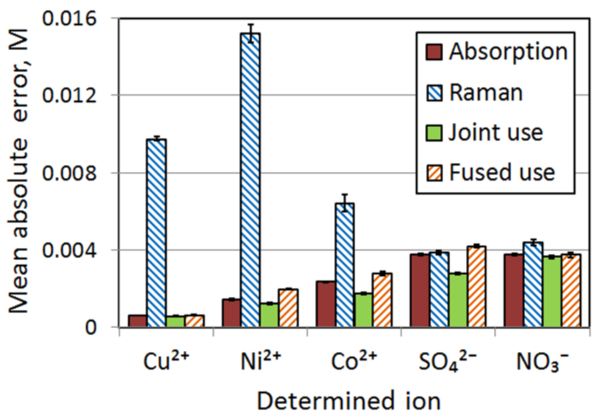
Пример 1: Отбор и преобразование признаков

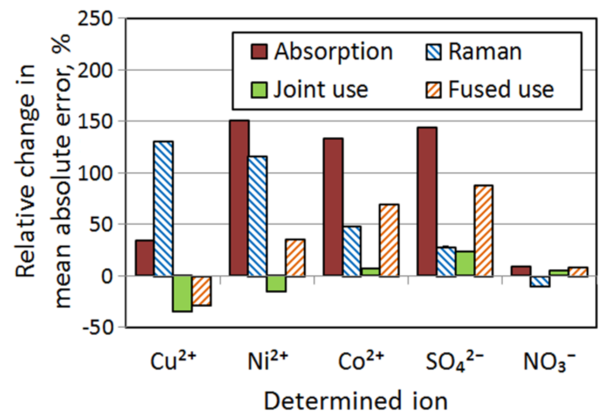




Отбор признаков (анализ весов нейронной сети)

100 признаков для поглощения, 250 признаков для КР, 350 для обоих





Преобразование признаков (анализ главных компонент)

99% описанной дисперсии. 6 признаков для поглощения,
5 признаков для КР, 11 для обоих, 7 для совместного

Пример 2: Обратные задачи геологоразведки

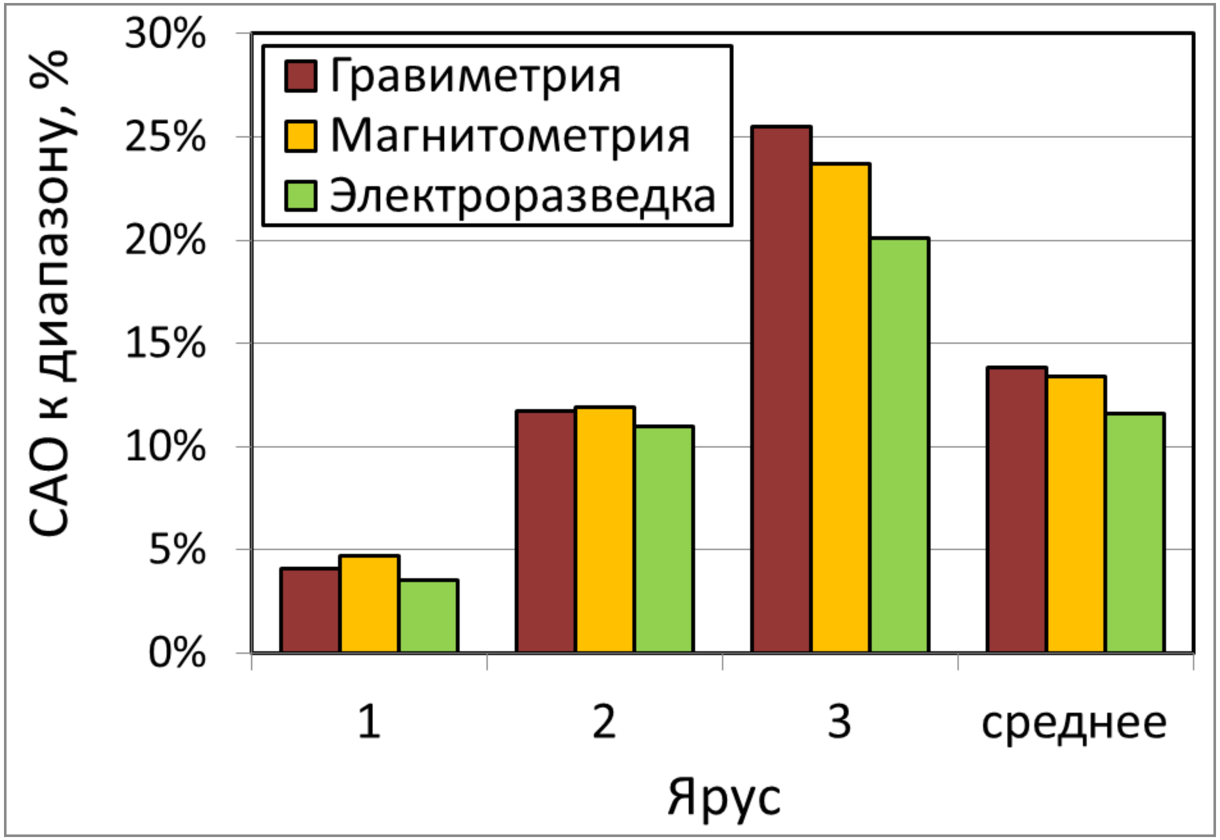
- Объект исследования: определение структурных границ, разделяющих геологические слои с постоянными значениями параметра:
  - Плотности в задаче гравиметрии
  - Намагниченности в магнитометрии
  - Удельного сопротивления в электроразведке
- Измеряются гравитационное, магнитное, электромагнитное поля
- 4-слойная синтезированная 2D модель
- Протяженность профиля 15 км
- Шаг измерения 0.5 км – итого 31 точка измерений по профилю
- Определяются 45 параметров

Пример 2: Методика работы с данными

- **10000 примеров** в каждой БД = 7000 трн + 2000 влд + 1000 тст
- **31** входной признак для гравиметрии и магнитометрии
- **62** входных признака для электроразведки
- **45** выходных признаков (мощности слоёв)
- Автономное определение
  (отдельная сеть для каждого определяемого параметра)
- Многослойный персептрон с 3 скрытыми слоями («мелкий»)
- Исходный набор признаков
- Комплексирование методов со слиянием признаков
- Мера ошибки – средняя абсолютная ошибка к диапазону D_s

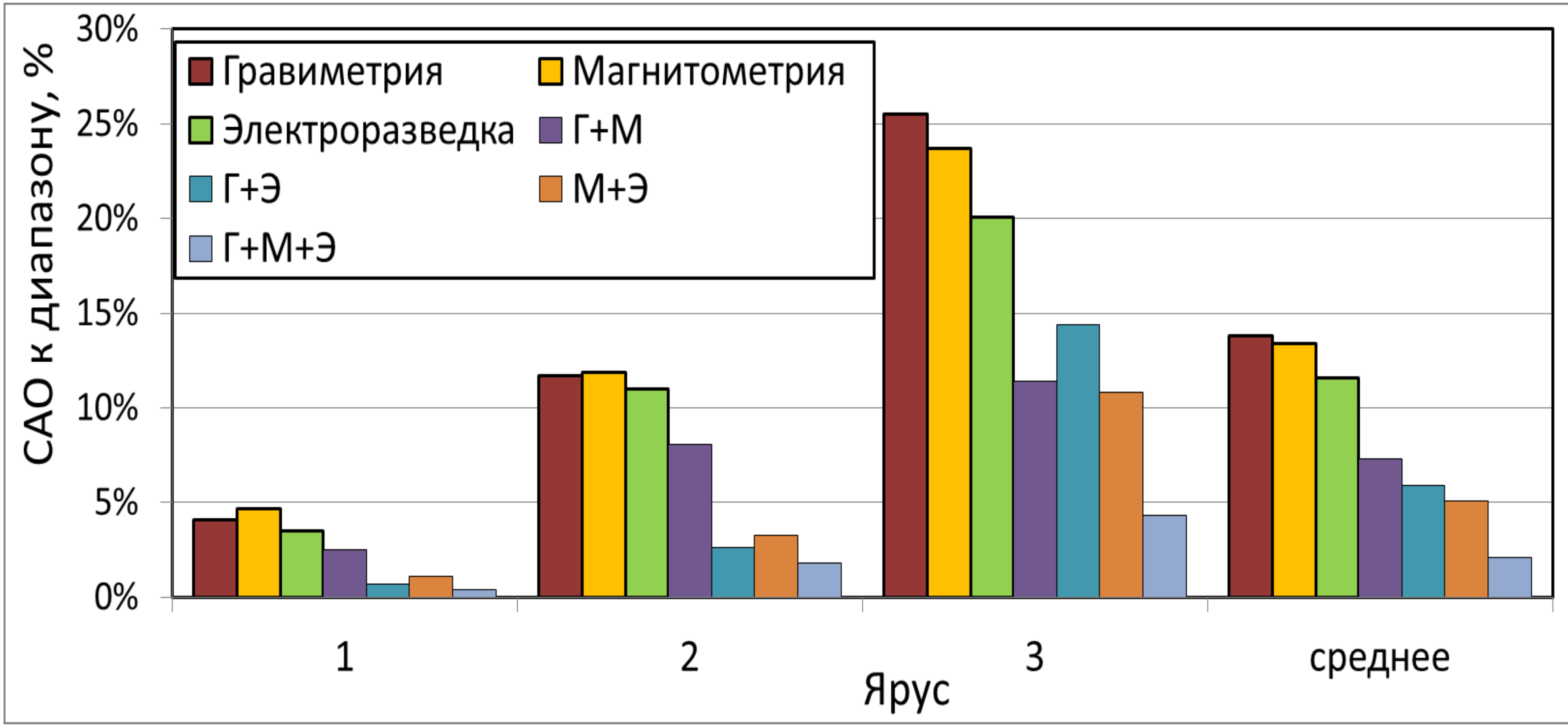
$$\| s - s' \|_{l_1}^{N} = \sum_{n=1}^{N} \frac{1}{D_s} \frac{1}{N} | s_n - s'_n | = \frac{1}{D_s N} \sum_{n=1}^{N} | s_n - s'_n |$$

Пример 2: Базовый эксперимент



Базовый эксперимент (полные наборы входных признаков)

Пример 2: Комплексирование



Результаты комплексирования физических методов

Выводы по комплексированию

- Если физические методы **сильно отличаются** по уровню ошибки, комплексирование **не работает**
- Если физические методы дают **близкий уровень** ошибки, комплексирование **работает**
- Дальнейшие направления работ
  - Использование на обеих задачах комплексирования аппроксиматоров
  - Использование на обеих задачах комплексирования данных

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РЕШЕНИЯ МНОГОПАРАМЕТРИЧЕСКИХ ОЗ: ИТОГИ

Рассмотрены различные методические подходы к решению ОЗ:

- От модели
- От эксперимента
- Квазимодельный

Продемонстрирована эффективность следующих методик:

- Понижение размерности входных данных
- Способы определения выходных параметров
    - Одновременное
    - Автономное
    - Групповое
    - Поэтапное
- Тренировка на данных с добавлением шума
- Комплексирование
    - Аппроксиматоров
    - Данных
    - Физических методов

С.А. Доленко

НИИЯФ МГУ

Лаборатория адаптивных методов обработки данных (ЛАМОД)

СПАСИБО ЗА ВНИМАНИЕ!

“- Просыпайся, дорогая, - сказала сестра. Что-то ты очень разоспалась!
- Ой, а какой я забавный сон видела! – сказала Алиса. “

Льюис Кэрролл. Алиса в стране чудес

Научный семинар «Глубокое обучение в вычислительной физике»

13 мая 2021 г.