



# Сети Колмогорова-Арнольда

VS

## Персептрон:

# Решение обратной задачи георазведки



Г.А. Куприянов<sup>1,2</sup>, И.В. Исаев<sup>2,3</sup>, С.А. Доленко<sup>2</sup>

<sup>1</sup>Физический факультет Московского государственного университета имени М.В. Ломоносова, Москва, Россия

<sup>2</sup>НИИ ядерной физики им. Д.В. Скобельцына Московского государственного университета им. М.В. Ломоносова, Москва, Россия

<sup>3</sup>Институт радиотехники и электроники им. Котельникова Российской академии наук, Москва, Россия

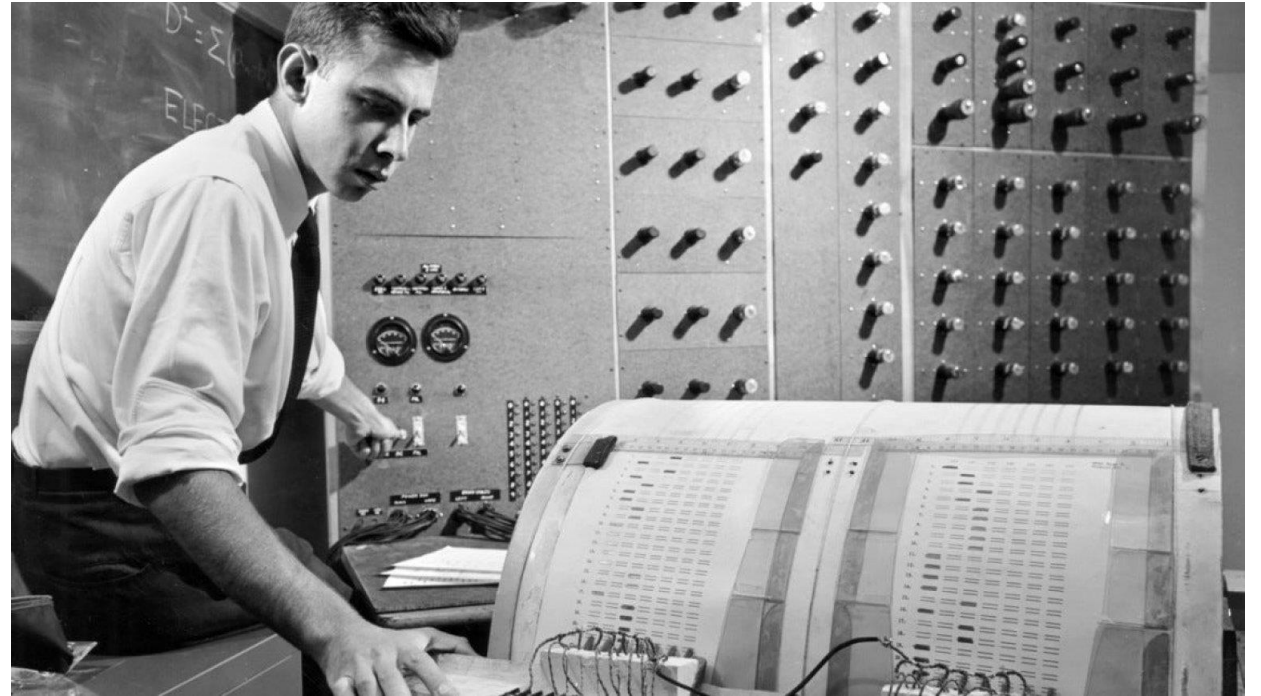
# План доклада

1. Сети Колмогорова-Арнольда (КАН)
2. Предметная область: задача разведочной геофизики
3. Сравнение результатов применения персептрона и КАН

# Первая искусственная нейронная сеть



Фрэнк Розенблат  
(1928-1971)

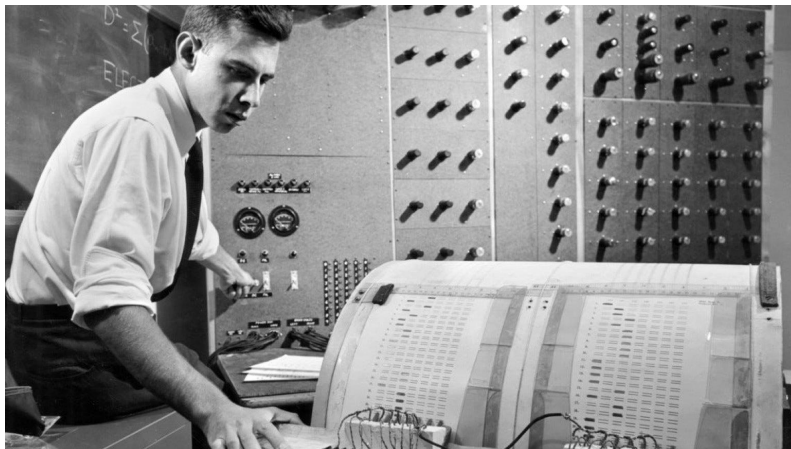


Первый персептрон «Марк-1»  
(1960)

# Первая искусственная нейронная сеть



Фрэнк  
Розенблат



«Марк-1»  
(1960)

$$f(\mathbf{x}) \approx \sum_{i=1}^{N(\epsilon)} a_i \sigma(\mathbf{w}_i \cdot \mathbf{x} + b_i)$$

Универсальная теорема  
аппроксимации персептроном  
(Цыбенко, 1989)

# 13-я проблема Гильберта

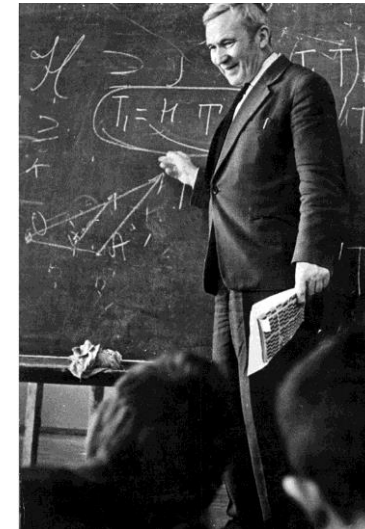
Можно ли представить функцию многих переменных в виде суперпозиции функций двух переменных?

# 13-я проблема Гильберта

Можно ли представить функцию многих переменных в виде суперпозиции функций двух переменных?

Можно, достаточно даже функций одной переменных:

$$f(\mathbf{x}) = \sum_{q=1}^{2n+1} \Phi_q \left( \sum_{p=1}^n \phi_{q,p}(x_p) \right)$$



Теорема Колмогорова-Арнольда

# Персептрон и КАН

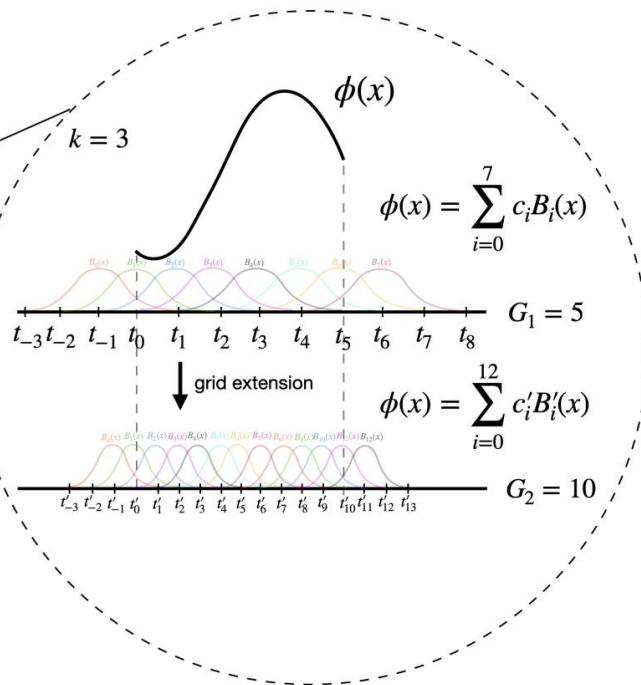
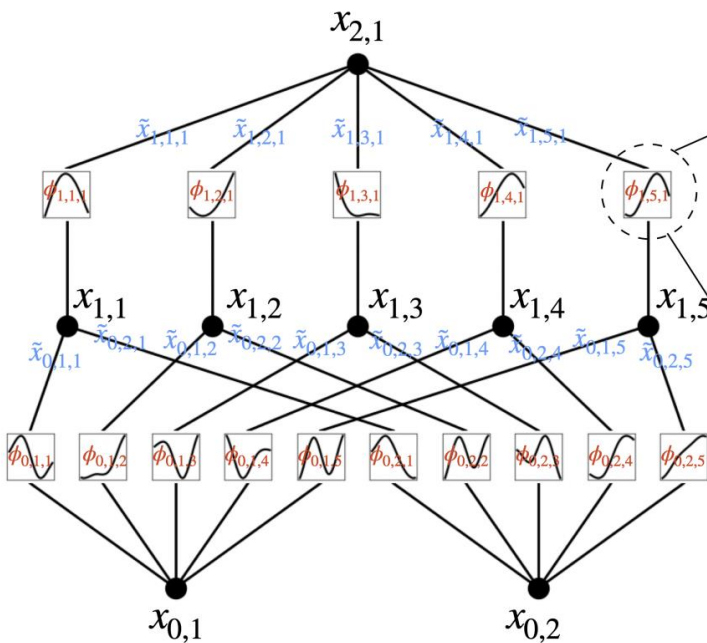
Model	<b>Multi-Layer Perceptron (MLP)</b>	<b>Kolmogorov-Arnold Network (KAN)</b>
Theorem	<b>Universal Approximation Theorem</b>	<b>Kolmogorov-Arnold Representation Theorem</b>
Formula (Shallow)	$f(\mathbf{x}) \approx \sum_{i=1}^{N(e)} a_i \sigma(\mathbf{w}_i \cdot \mathbf{x} + b_i)$	$f(\mathbf{x}) = \sum_{q=1}^{2n+1} \Phi_q \left( \sum_{p=1}^n \phi_{q,p}(x_p) \right)$
Model (Shallow)	<p><b>(a)</b></p> <p><i>fixed activation functions on nodes</i></p> <p><i>learnable weights on edges</i></p>	<p><b>(b)</b></p> <p><i>learnable activation functions on edges</i></p> <p><i>sum operation on nodes</i></p>

KAN: Kolmogorov-Arnold Networks, Cristian J. Vaca-Rubio, Luis Blanco, Roberto Pereira, Màrius Caus, <https://arxiv.org/abs/2404.19756>



# Активационные функции в KAN

- Параметризуем функции активаций В-сплайнами
- Обучение = подбор коэффициентов при В-сплайнах



## Преимущества В-сплайнов:

1. Параметризация вектором чисел
2. Локальность изменения
3. Возможность адаптивно изменять точность сплайна



# Детали реализации

- На самом деле активация выглядит следующим образом (b(x) - аналог residuals connections)

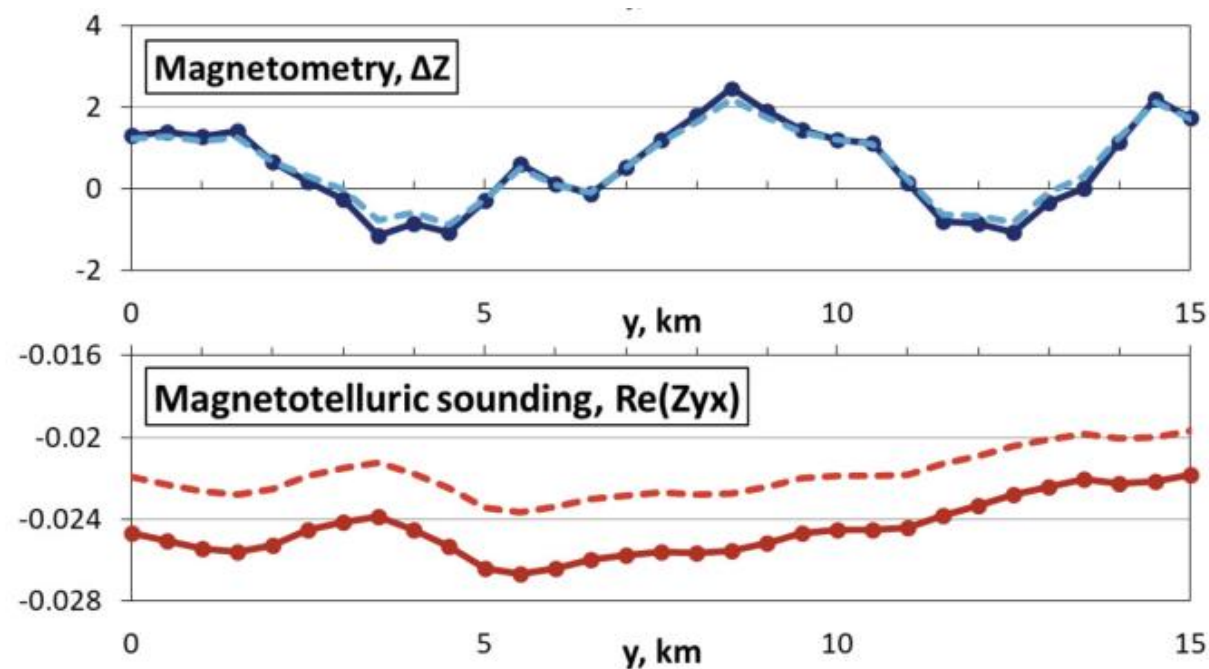
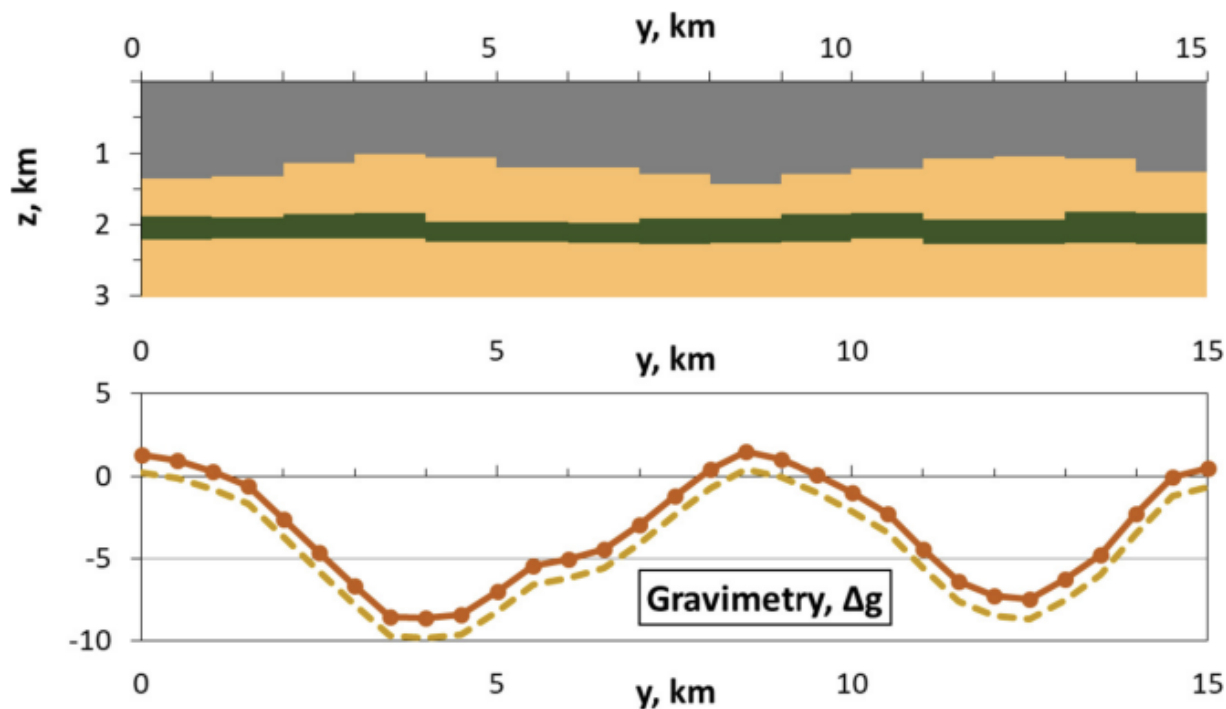
$$\phi(x) = w_b b(x) + w_s \text{spline}(x).$$

$$b(x) = \text{silu}(x) = x / (1 + e^{-x})$$

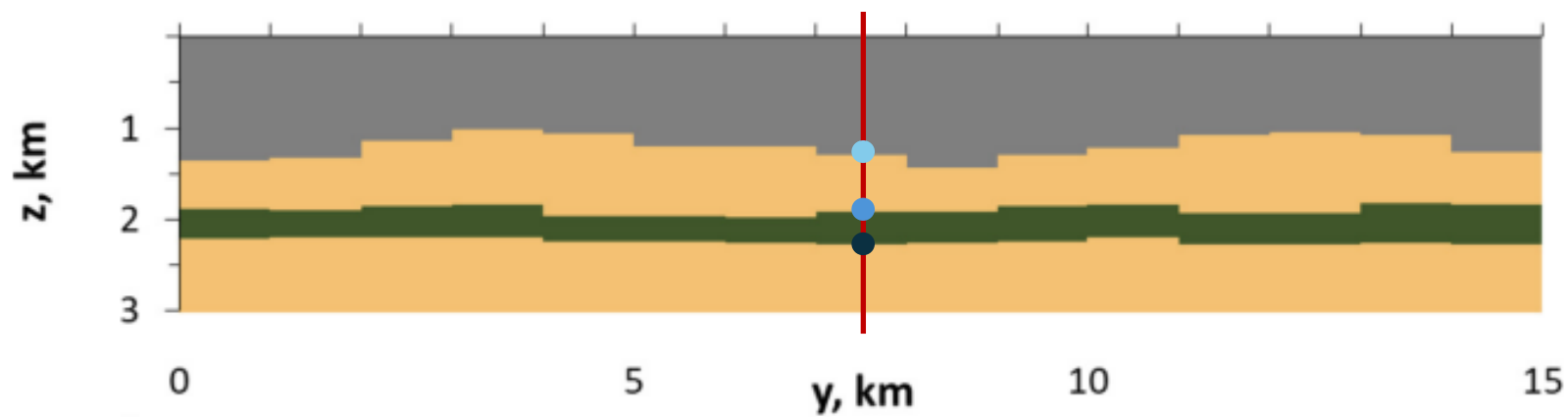
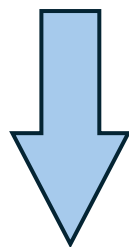
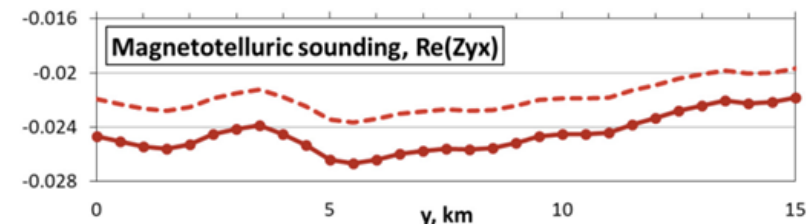
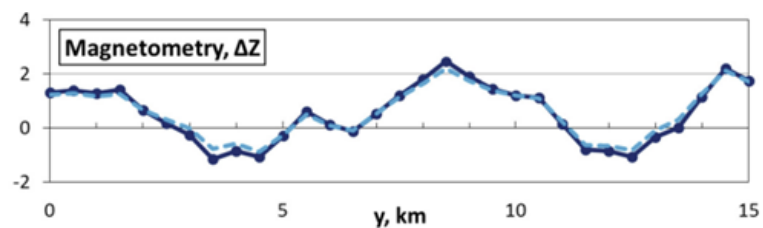
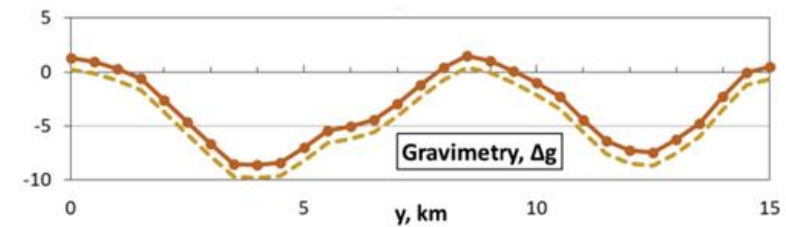
- Сетка, на которой строятся базисные сплайны, может адаптивно меняться в процессе обучения КАН
- Порядок В-сплайнов  $\sim 3$

# Обратная задача геофизики

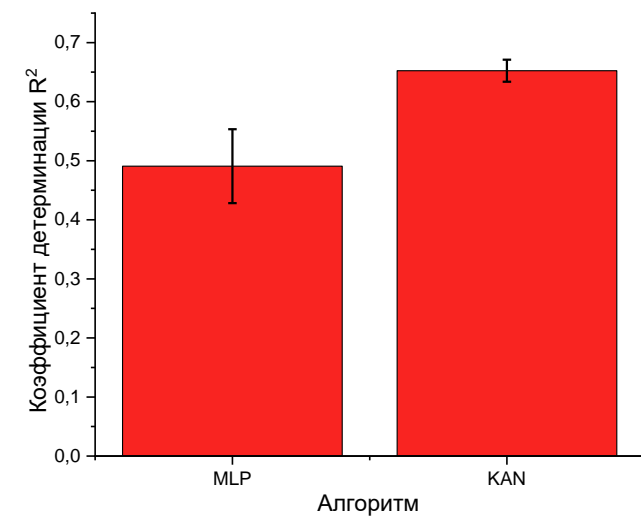
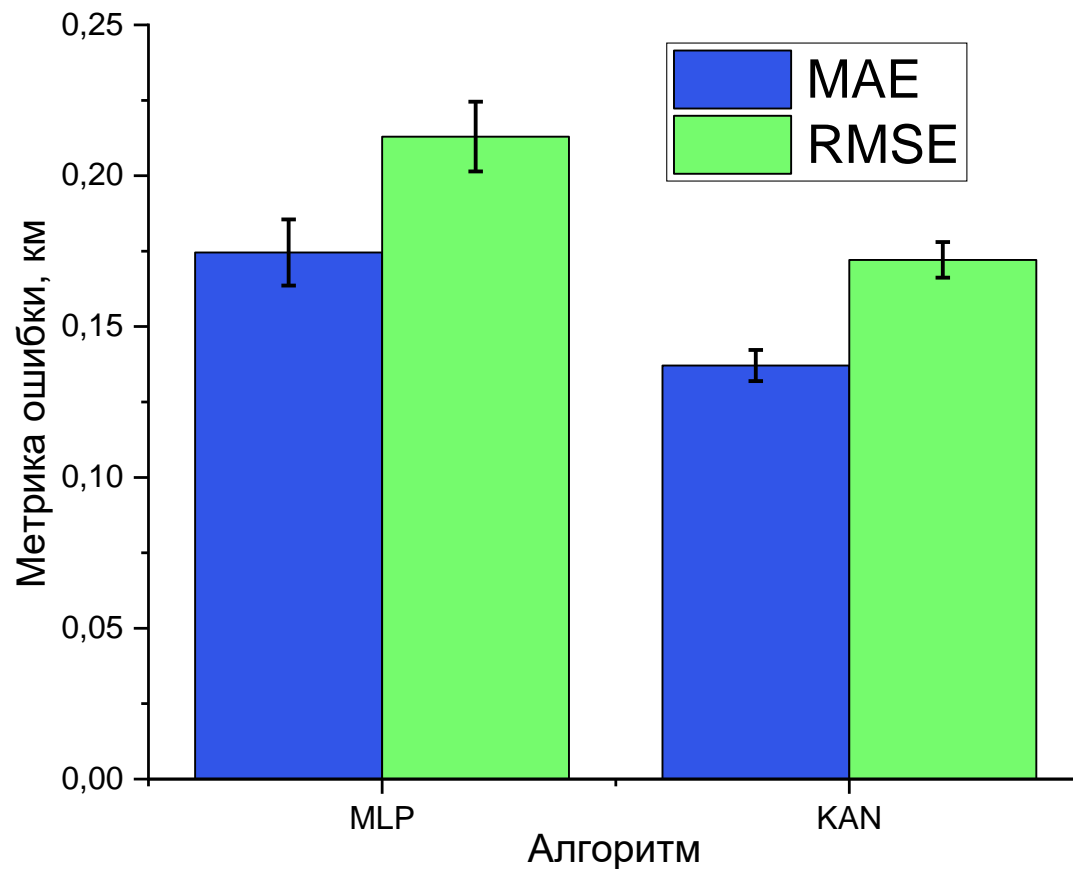
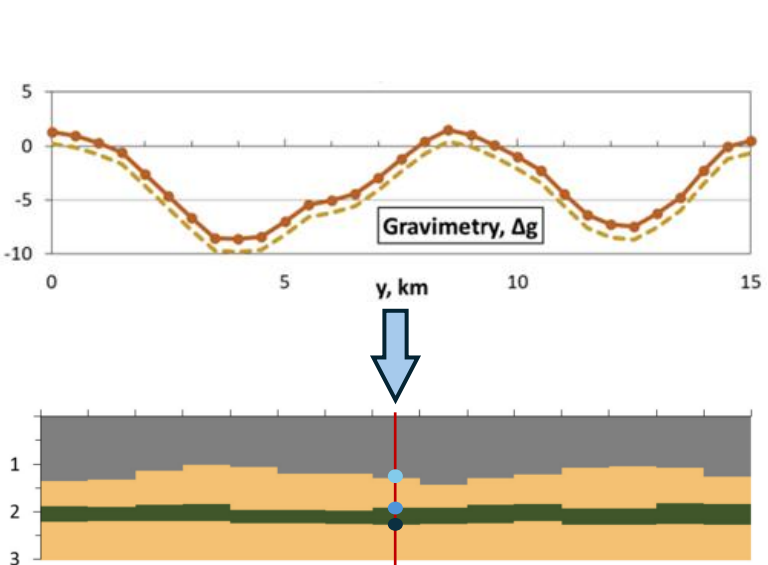
**Задача:** По данным неинвазивных исследований (Гравиметрия, Магнитометрия, Магнитотеллурика) восстановить профиль залегающих пород



# Обратная задача геофизики

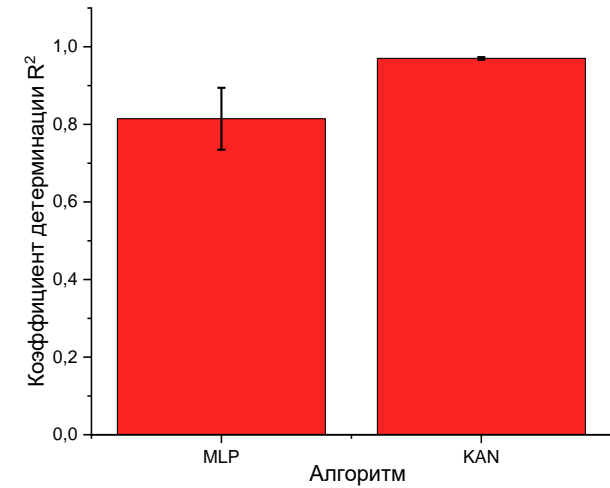
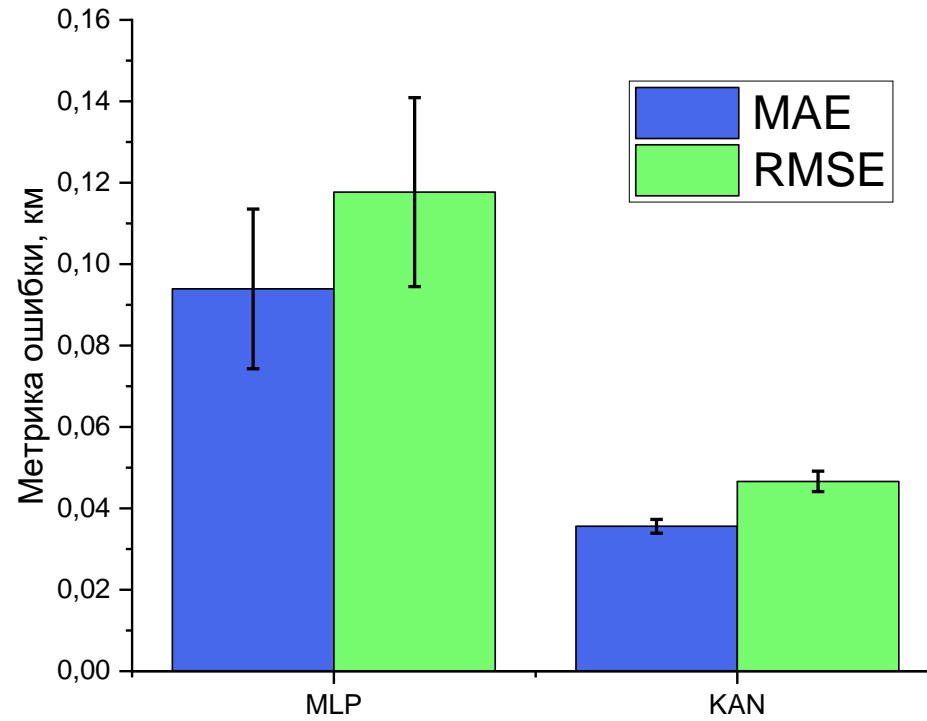
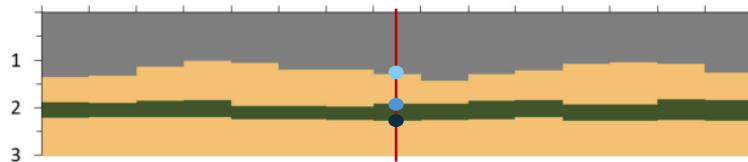
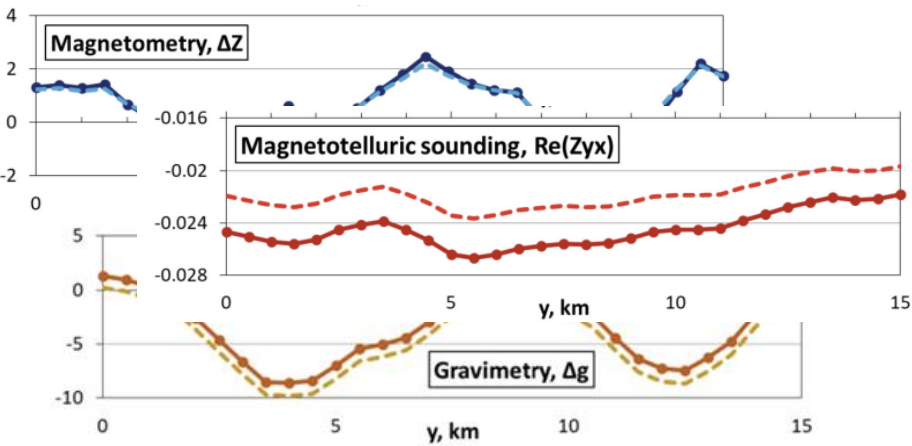


# Сравнение персептрона и КАН



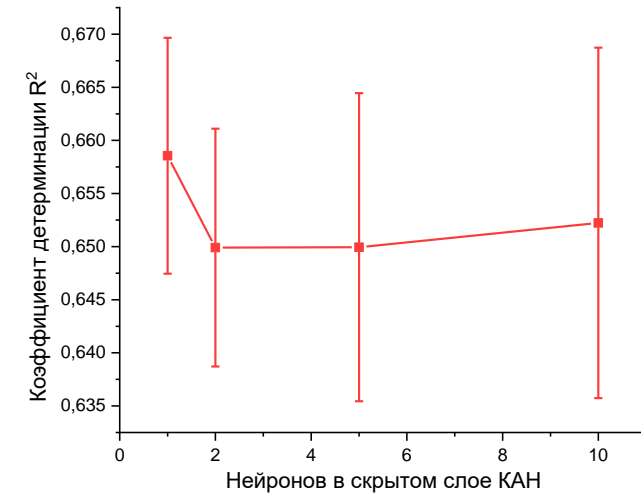
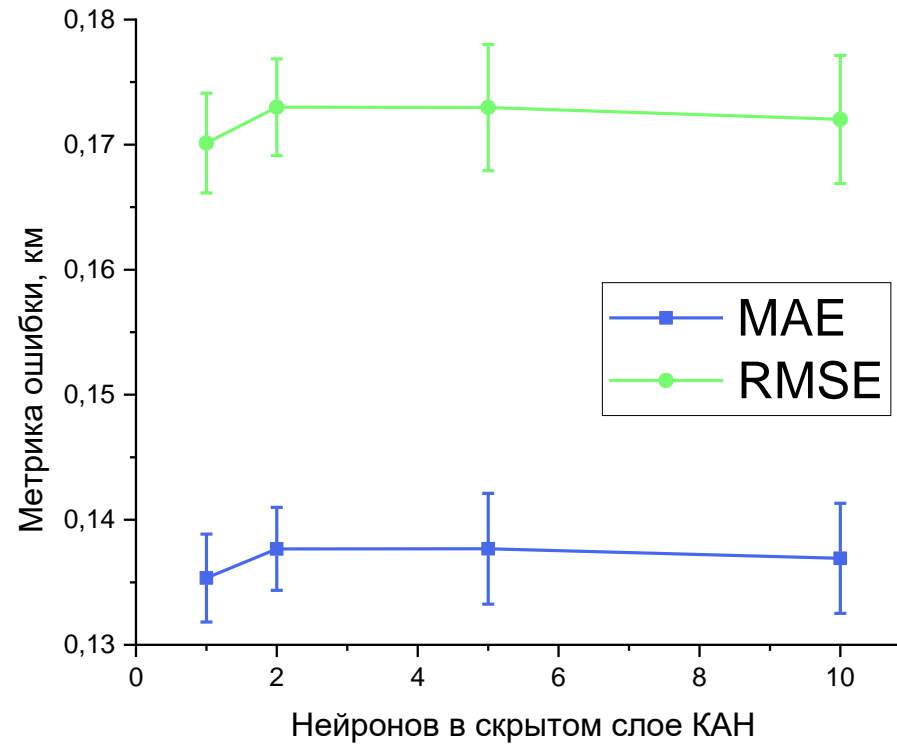
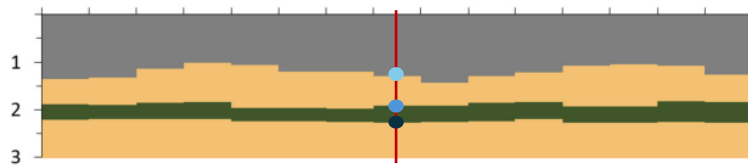
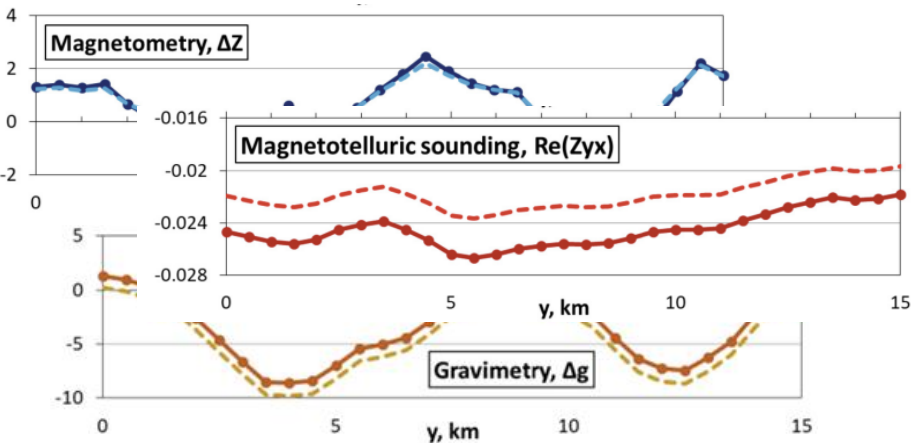
# Сравнение перцептрона и КАН

## Максимальный набор входных данных



Алгоритм

# КАН с разным количеством нейронов в скрытом слое



# Выводы

1. Сети Колмогорова-Арнольда (КАН) на обратной задаче георазведки превзошли персептрон как при ограниченном, так и при полном наборе входных данных.
2. Оптимальное количество нейронов в скрытом слое у КАН – 1 нейрон.



# Спасибо за внимание!

## Детали реализации КАН:

1. Используемая библиотека: ruKan
2. Порядок сплайнов:  $K=3$
3. Размер сетки сплайнов:  $\text{grid}=3$
4. Функция потерь: MSE
5. Оптимизатор: LBFGS
6. Остановка обучения: по callback на валидационном наборе ( $\text{tol}=0.001$ ,  $\text{iter\_no\_change}=25$ )

