KAN: Kolmogorov–Arnold Networks

Александров Кирилл

Мотивация

В основе полносвязных нейронных сетей (MLP) лежит универсальная теорема приближения (UAT). Авторы решили пойти другим путем и оттолкнуться от теоремы Колмогорова-Арнольда:

Теорема

(Колмогоров-Арнольд)
Каждая многомерная непрерывная функция может быть представлена в виде суперпозиции непрерывных функций одной переменной.

На основе этой теоремы предложена архитектура KAN.

Основные идеи

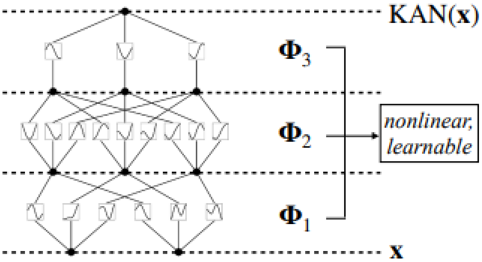

Рис.: Модель нейронной сети

Основные особенности:

- Обучаемые функции активации
- Использование B-сплайнов
- Замена матриц весов на матрицы функций активации
- Индивидуальные импульсы для каждого нейрона
- Нелинейные преобразования между слоями

Реализация: Функции активации

Обучаемые функции активации:

$$\phi(x) = \omega_1 b(x) + \omega_2 \text{spline}(x) \tag{1}$$

$$b(x) = \frac{x}{1 + e^{-x}} = \text{silu}(x) \tag{2}$$

$$\text{spline}(x) = \sum_i c_i B_i(x) \tag{3}$$

где $B_i(x)$ — B-сплайны, c_i — обучаемые параметры.

Реализация: Архитектура сети

Параметры сети:

- L — глубина
- n_i — ширина i-го слоя
- $[n_1, n_2, ..., n_L]$ — массив параметров

Сеть обозначается как KAN$[n_1, n_2, ..., n_L]$. Пример: KAN[1] — простая интерполяция.

Реализация: Математическое описание

Преобразования между слоями:

$$x_{l+1,j} = \sum_{i=1}^{n_l} \phi_{l,j,i}(x_{l,i}) \tag{4}$$

Матричный вид:

$$x_{l+1} = \begin{pmatrix} \phi_{l,1,1}(\cdot) & \cdots & \phi_{l,1,n_l}(\cdot) \\ \vdots & \ddots & \vdots \\ \phi_{l,n_{l+1},1}(\cdot) & \cdots & \phi_{l,n_{l+1},n_l}(\cdot) \end{pmatrix} x_l$$

Итоговое выражение:

$$\mathrm{KAN}(x) = (\Phi_{L-1} \circ \cdots \circ \Phi_0)x$$

Обоснование подхода

Обучаемость: Преобразования дифференцируемы и обучаются методом градиентного спуска. **Теорема аппроксимации:**

Теорема

(Approximation theory, KAT)
Для функции $f(x)$ существуют параметры G, k, такие что:

$$\|f - f_G\|_{C^m} \leq CG^{-k-1+m}$$

где f_G — аппроксимированная функция.

Сравнение KAN с MLP

Преимущества KAN:

- Интерпретируемые данные
- Легкость переобучения

Недостатки:

- Большее число параметров
- Более медленное обучение

Заключение

В данном обзоре были рассмотрены основные аспекты архитектуры KAN, включая:

- Мотивацию разработки
- Преимущества и недостатки по сравнению с MLP
- Реализацию и математическое обоснование

Подробнее о тестах и применении KAN — в оригинальной статье: `https://arxiv.org/abs/2404.19756`.