

# **Kolmogorov-Arnold Networks: Новая Эра Глубокого Обучения**

## **Николаев М.В.**

Научный руководитель: Р.Ш. Мисбахов, доцент кафедры  
Естественнонаучных дисциплин и информационных технологий  
Альметьевский филиал Казанского национального исследовательского  
технического университета  
им. А.Н. Туполева-КАИ

**Аннотация.** В данной статье представлен обзор современных архитектур глубокого обучения с акцентом на многослойный перцептрон (MLP) и его ограничения. Особое внимание уделено недавно разработанной архитектуре Kolmogorov-Arnold Networks (KAN), которая основана на теореме Колмогорова-Арнольда и использует параметры сплайнов для аппроксимации функций. Рассматриваются основные компоненты KAN, её преимущества и недостатки по сравнению с MLP, а также приводятся примеры реализации для задач регрессии и классификации. В заключение обсуждаются перспективы дальнейшего развития и применения KAN в глубоких нейронных сетях.

**Ключевые слова:** Глубокое обучение, многослойный перцептрон, Kolmogorov-Arnold Networks, функции активации, теорема Колмогорова-Арнольда, аппроксимация функций, сплайны, нейронные сети, интерпретируемость, масштабируемость.

**Nikolaev M.V.**

Supervisor: R.S.Misbakhov, docent of the Department of Natural Sciences and  
Information Technology  
Almetyevsk branch of Kazan National Research Technical University named after  
A.N. Tupolev-KAI

**Abstract.** This article provides an overview of modern deep learning architectures with a focus on the multilayer perceptron (MLP) and its limitations. Special attention is given to the recently developed Kolmogorov-Arnold Networks (KAN) architecture, which is based on the Kolmogorov-Arnold theorem and uses spline parameters for function approximation. The main components of KAN, its advantages and disadvantages compared to MLP, and implementation examples for regression and classification tasks are discussed. The article concludes with a discussion on the future development and application prospects of KAN in deep neural networks.

**Keywords:** Deep learning, multilayer perceptron, Kolmogorov-Arnold Networks, activation functions, Kolmogorov-Arnold theorem, function approximation, splines, neural networks, interpretability, scalability.

## **Введение**

Современные архитектуры глубокого обучения, такие как многослойный перцептрон (MLP), являются основой многих нейронных сетей. Однако, недавно разработанная исследователями из MIT архитектура Kolmogorov-Arnold Networks (KAN) предлагает новый подход, в котором функции активации располагаются на "ребрах" сети, что может существенно изменить подходы к глубокому обучению.

## **Перцептрон и его ограничения**

Многослойный перцептрон (MLP) использует линейные слои с функциями активации. Согласно теореме Цыбенко, любая нейронная сеть с нелинейной функцией активации может аппроксимировать любую непрерывную функцию. Однако, при решении сложных задач такой подход имеет свои ограничения, такие как отсутствие интерпретируемости и сложности с масштабируемостью.

## **Теорема Колмогорова-Арнольда**

Теорема Колмогорова-Арнольда утверждает, что любую непрерывную функцию можно аппроксимировать с помощью одномерных функций. Эта теорема долгое время оставалась не реализованной в практике из-за сложности создания гладких и масштабируемых функций. Однако, она предлагает теоретическую основу для разработки новых архитектур нейронных сетей.

## **Архитектура KAN**

Архитектура KAN использует параметры сплайнов для аппроксимации функций, делая их гладкими и обучаемыми. Сеть состоит из трех основных компонентов:

- **Входной энкодер:** Преобразует входные данные в более высокоразмерное пространство.
- **Сеть смешивания:** Комбинирует закодированные входные данные с использованием обучаемых параметров.
- **Выходной декодер:** Преобразует смешанные признаки обратно в выходное пространство.

В отличие от MLP, где обучаются веса, в KAN обучаются функции, что повышает гибкость и точность модели.

## **Преимущества и недостатки KAN**

Преимущества KAN включают в себя:

- Меньшее количество нейронов для достижения высокой точности.
- Лучшая обобщаемость данных.
- Возможность повышения точности без полного переобучения.
- Высокая интерпретируемость и способность аппроксимировать сложные математические функции.

Недостатки KAN:

- Большое количество параметров, необходимых для обучения.
- Медленная скорость обучения по сравнению с традиционными методами.

## Примеры реализации KAN

### Установка библиотеки PyKAN

Установить библиотеку можно через pip:

```
pip install pykan
```

Или клонируя репозиторий:

```
git clone https://github.com/KindXiaoming/pykan.git
cd pykan
pip install -e .
```

### Пример задачи регрессии:

```
from kan import *
import torch
# Определение модели KAN
model = KAN(width=[2,5,1], grid=5, k=3, seed=0)
# Генерация датасета
f = lambda x: torch.exp(torch.sin(torch.pi*x[:,[0]])) + x[:,[1]]**2)
dataset = create_dataset(f, n_var=2)
# Обучение модели
model.train(dataset, opt="LBFGS", steps=20, lamb=0.01, lamb_entropy=10.)
# Визуализация KAN
model.visualize()
```

### Пример задачи классификации:

```
from kan import KAN
import torch
import numpy as np
from sklearn.datasets import make_moons
# Генерация датасета "две луны"
train_input, train_label = make_moons(n_samples=1000, noise=0.1)
test_input, test_label = make_moons(n_samples=1000, noise=0.1)
dataset = {
    'train_input': torch.from_numpy(train_input),
    'test_input': torch.from_numpy(test_input),
    'train_label': torch.from_numpy(train_label),
    'test_label': torch.from_numpy(test_label)
}
```

```
# Определение модели KAN
model = KAN(width=[2,2], grid=3, k=3)
# Обучение модели
results = model.train(dataset, opt="LBFGS", steps=20,
loss_fn=torch.nn.CrossEntropyLoss())
# Точность модели
train_acc = torch.mean((torch.argmax(model(dataset['train_input']), dim=1) ==
dataset['train_label']).float())
test_acc = torch.mean((torch.argmax(model(dataset['test_input']), dim=1) ==
dataset['test_label']).float())
print(f"Train Accuracy: {train_acc}, Test Accuracy: {test_acc}")
```

## **Заключение**

Kolmogorov-Arnold Networks (KAN) представляют собой значительный шаг вперед в развитии нейронных сетей, предлагая новый подход к аппроксимации функций. Эта архитектура имеет потенциал для дальнейшего совершенствования и может существенно повлиять на развитие моделей глубокого обучения в будущем.

## Список литературы

1. Aggarwal, Charu C. Neural Networks and Deep Learning: A Textbook. – Springer, 2018. – 497 с.
2. Deisenroth, Marc Peter, Faisal, A. Aldo, Ong, Cheng Soon. Mathematics for Machine Learning. – Cambridge University Press, 2020. – 398 с.
3. Goodfellow, Ian, Bengio, Yoshua, Courville, Aaron. Deep Learning. – MIT Press, 2016. – 775 с.
4. Kolmogorov-Arnold Networks на arXiv.org. – URL: <https://arxiv.org/abs/2404.19756> (Дата обращения 17.05.2024).
5. KAN: Kolmogorov-Arnold Networks | AI Models FYI. – URL: <https://www.aimodels.fyi/kan-kolmogorov-arnold-networks> (Дата обращения 17.05.2024).
6. KAN: Kolmogorov-Arnold Networks | Papers With Code. – URL: <https://paperswithcode.com/paper/kan-kolmogorov-arnold-networks> (Дата обращения 17.05.2024).
7. The Rise of Kolmogorov-Arnold Networks: A New Frontier in Deep Learning. – Multiplatform AI. – URL: <https://multiplatform.ai/articles/the-rise-of-kolmogorov-arnold-networks> (Дата обращения 17.05.2024).