

КАЗАХСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ ИМЕНИ АЛЬ-ФАРАБИ

А.Т. Агишев

МЕТОДЫ И ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

**Сборник лекций для студентов магистратуры,
обучающихся по образовательной программе
«7M06201 - Радиотехника, электроника и телекоммуникации»**

Алматы, 2025

Лекция 8. Основы Глубокого обучения

Цель лекции

Дать базовое понимание архитектуры и принципов работы глубоких нейронных сетей. Рассмотреть ключевые концепции, лежащие в основе deep learning, включая функции активации, распространение ошибки, оптимизацию, а также показать применение глубоких сетей в инженерных и телекоммуникационных задачах.

Основные вопросы:

1. Что такое глубокое обучение и чем оно отличается от классического машинного обучения.
2. Архитектура искусственных нейронных сетей (ANN).
3. Математическая основа нейрона и функции активации.
4. Обратное распространение ошибки и оптимизация.
5. Типы нейронных сетей: полносвязные, сверточные, рекуррентные.
6. Регуляризация, нормализация и методы улучшения обучения.
7. Применение глубокого обучения в радиотехнике, электронике и телекоммуникациях.

Краткие тезисы:

1. **Определение глубокого обучения.** По Goodfellow, Bengio & Courville (2016): *Глубокое обучение (Deep Learning) - это набор методов машинного обучения, основанных на построении и обучении нейронных сетей с множеством скрытых слоёв, способных автоматически извлекать иерархические признаки из данных.* В отличие от классического МО, где признаки задаются вручную, DL **самостоятельно извлекает признаки** из «сырых» данных - изображений, сигналов, текста.

2. **Архитектура нейронной сети.**

- **Нейрон** - базовый элемент, вычисляющий:

$$y = f\left(\sum_{i=1}^n w_i x_i + b\right)$$

где f - функция активации.

- **Слой:**

- входной (input layer),
- скрытые (hidden layers),
- выходной (output layer). **Глубина** сети - количество скрытых слоёв, **ширина** - число нейронов в слое. Сеть обучается на основе минимизации ошибки между предсказанным и реальным выходом.

3. **Функции активации.** Определяют нелинейность и способность сети моделировать сложные зависимости.
Основные типы:

- **Sigmoid:** $f(x) = \frac{1}{1+e^{-x}}$
- **Tanh:** $f(x) = \tanh(x)$
- **ReLU:** $f(x) = \max(0, x)$ - современный стандарт (устойчива к затухающим градиентам).
- **Leaky ReLU, ELU, GELU** - модификации для повышения стабильности.

В современных моделях (ResNet, Transformer) активации подбираются под архитектуру для лучшей сходимости.

4. **Обратное распространение ошибки (Backpropagation).** Ключевой механизм обучения нейронных сетей (Rumelhart et al., 1986).

- Прямое распространение: вычисление выходов сети.
- Обратное распространение: вычисление градиентов функции потерь по весам.
- Обновление весов по правилу:

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} - \eta \frac{\partial L}{\partial w_{ij}}$$

где η - скорость обучения. Используются оптимизаторы: **SGD, Adam, RMSProp**. Преимущество метода - эффективность при большом числе параметров.

5. Основные архитектуры глубоких сетей.

а) Полносвязные сети (MLP, Multilayer Perceptron): Универсальные аппроксиматоры, применяются для задач регрессии и классификации.

б) Сверточные нейронные сети (CNN): Извлекают пространственные признаки, применяются для изображений и двумерных сигналов.

- Компоненты: сверточный слой (convolution), pooling, fully connected layer.
- Примеры: *LeNet, AlexNet, VGG, ResNet*. В радиотехнике CNN используются для **распознавания форм сигналов**, спектрального анализа и классификации модуляций.

с) Рекуррентные сети (RNN, LSTM, GRU): Обработывают временные зависимости. Применяются для анализа временных рядов, предсказания каналов связи, анализа речи.

- *LSTM* и *GRU* решают проблему исчезающих градиентов.

д) Современные архитектуры:

- **Transformer (Attention-based)** - архитектура, использующая механизм внимания для выделения значимых частей данных.

- **Autoencoders, GANs, Diffusion models** - для генерации, восстановления и анализа данных.

6. Регуляризация и методы улучшения обучения.

- **Dropout:** случайное выключение нейронов при обучении для предотвращения переобучения.
- **Batch Normalization:** нормализация активаций между слоями для ускорения сходимости.
- **Data Augmentation:** искусственное расширение обучающего набора (шум, повороты, масштабирование).
- **Early Stopping:** остановка обучения при ухудшении метрики на валидации.
- **Transfer Learning:** дообучение уже обученной модели на новых данных - популярно в инженерных задачах.

7. Применение глубокого обучения в радиотехнике и телекоммуникациях.

- **Анализ сигналов:** определение типа модуляции, классификация помех.
- **Обнаружение неисправностей:** по спектрам, временным рядам и вибросигналам.
- **Прогнозирование трафика и управления ресурсами:** использование LSTM и RL.
- **Адаптивные антенны и beamforming:** CNN для оптимизации направленности.
- **Сжатие и восстановление данных:** Autoencoder для оптимизации передачи.
- **Когнитивные радиосистемы:** применение DQN и Policy Gradient для динамического выбора каналов.

Современные телекоммуникационные стандарты 5G/6G интегрируют DL в системы управления и мониторинга.

Вопросы для контроля, изучаемого материал:

- 1) В чём заключается основное отличие глубокого обучения от классического машинного обучения?
- 2) Как устроен искусственный нейрон и какие функции активации используются?
- 3) Объясните принцип обратного распространения ошибки.
- 4) Какие архитектуры относятся к глубоким сетям и для чего они применяются?
- 5) Как работают сверточные и рекуррентные нейронные сети?
- 6) Какие методы используются для предотвращения переобучения в глубоких сетях?
- 7) Приведите примеры применения глубокого обучения в радиотехнике и телекоммуникациях.

- 8) Что такое transfer learning и почему он эффективен в инженерных задачах?

Рекомендуемый список литературных источников:

1. Russell, S., Norvig, P. Artificial Intelligence: A Modern Approach. 4th Edition. Pearson, 2021.
2. Goodfellow, I., Bengio, Y., Courville, A. Deep Learning. MIT Press, 2016.
3. Alpaydin, E. Introduction to Machine Learning. 4th Edition. MIT Press, 2020.