

Лекция 11. Нейронные сети и глубокое обучение в Data Mining

Тема: Основы нейросетей, применение в обработке изображений и текстов

1. Введение

Современные методы **интеллектуального анализа данных (Data Mining)** все чаще используют подходы **глубокого обучения (Deep Learning)** — раздела машинного обучения, основанного на **искусственных нейронных сетях (ИНС)**.

Глубокие нейронные сети позволяют извлекать скрытые закономерности и абстрактные представления из данных, которые невозможно обнаружить традиционными статистическими методами.

Благодаря достижениям в вычислительных технологиях, появлению больших данных (Big Data) и графических процессоров (GPU), нейросетевые методы стали ведущими в таких областях, как **распознавание изображений, обработка речи, машинный перевод и генерация текста**.

2. Основные понятия и структура нейронных сетей

2.1. Искусственный нейрон

Прототипом для искусственного нейрона послужил **биологический нейрон** — клетка, передающая электрические сигналы.

Искусственный нейрон принимает входные данные, умножает их на **веса (weights)**, суммирует и пропускает результат через **функцию активации**, определяющую выход.

Формально:

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

где:

- x_i — входные данные,
 - w_i — веса связей,
 - b — смещение (bias),
 - f — функция активации (например, ReLU, сигмоида, tanh).
-

2.2. Архитектура нейронной сети

Нейронная сеть состоит из:

- **входного слоя** — принимает исходные данные;
- **скрытых слоёв (hidden layers)** — преобразуют данные, выявляя зависимости;
- **выходного слоя** — выдает результат (например, вероятность принадлежности классу).

Глубокая нейронная сеть (DNN) — это сеть с несколькими скрытыми слоями, которая способна обучаться сложным нелинейным функциям.

3. Основные типы нейронных сетей

3.1. Полносвязные сети (Fully Connected Networks)

Каждый нейрон связан со всеми нейронами предыдущего слоя.

Используются для табличных данных и простых задач классификации или регрессии.

Однако при большом числе признаков такие сети становятся громоздкими и склонными к переобучению.

3.2. Свёрточные нейронные сети (CNN — Convolutional Neural Networks)

CNN предназначены для **анализа изображений и пространственных данных**.

Они используют **свёртки (convolutions)** — фильтры, которые выделяют важные признаки (края, формы, текстуры).

Ключевые компоненты CNN:

- **Свёрточный слой** — извлекает признаки;
- **Слой подвыборки (Pooling)** — уменьшает размерность;
- **Полносвязный слой** — принимает итоговые признаки для классификации.

Применение:

- распознавание лиц, объектов, символов;
- медицинская диагностика по снимкам;
- автономные автомобили (распознавание дорожных знаков).

Примеры архитектур: *LeNet*, *AlexNet*, *VGG*, *ResNet*.

3.3. Рекуррентные нейронные сети (RNN — Recurrent Neural Networks)

RNN разработаны для работы с **последовательными данными** — текстами, временными рядами, аудио.

Они сохраняют контекст предыдущих состояний через **обратные связи**, что позволяет учитывать порядок элементов.

Модификации RNN:

- **LSTM (Long Short-Term Memory)** — устраняет проблему исчезающего градиента;
- **GRU (Gated Recurrent Unit)** — упрощённая и быстрая версия LSTM.

Применение:

- автоматический перевод;
- анализ тональности текстов;
- прогнозирование временных рядов;
- распознавание речи.

3.4. Трансформеры (Transformers)

Современные модели, такие как **BERT, GPT, T5**, основаны на архитектуре **трансформера**, которая использует механизм **внимания (attention)** для оценки важности элементов входной последовательности.

Трансформеры заменили рекуррентные сети в большинстве задач обработки текста и речи.

Применение:

- генерация и анализ текста (ChatGPT, Bard);
- поиск информации;
- резюмирование и ответы на вопросы;
- мультимодальные модели (текст + изображение).

4. Обучение нейронных сетей

Процесс обучения включает:

1. **Прямое распространение (forward pass)** — вычисление выхода сети;
2. **Функцию потерь (loss function)** — измеряет ошибку;

3. **Обратное распространение ошибки (backpropagation)** — корректирует веса;
4. **Оптимизацию** (например, с помощью Adam, SGD, RMSProp).

Цель обучения — минимизировать ошибку модели, улучшая точность предсказаний.

5. Применение в Data Mining

5.1. В обработке изображений:

- Распознавание объектов и лиц;
- Сегментация изображений (например, в медицине);
- Обнаружение аномалий и дефектов на производстве.

5.2. В обработке текстов:

- Анализ тональности и мнений;
- Классификация документов;
- Автоматический перевод;
- Поиск и суммаризация информации.

5.3. В других областях:

- Рекомендательные системы;
 - Финансовое прогнозирование;
 - Анализ звука, видео и сигналов.
-

6. Преимущества и ограничения

Преимущества:

- способность работать с неструктурированными данными (изображения, тексты);
- высокая точность на больших выборках;
- возможность автоматического извлечения признаков.

Ограничения:

- требовательность к вычислительным ресурсам;
- сложность интерпретации (эффект «чёрного ящика»);
- необходимость большого количества обучающих данных.

7. Заключение

Нейронные сети и глубокое обучение радикально изменили подход к анализу данных.

Они позволяют моделировать сложные нелинейные зависимости и достигать уровня, сопоставимого с человеческим восприятием.

В контексте **Data Mining** нейросети стали универсальным инструментом для анализа изображений, текстов и звука, а также для поиска скрытых закономерностей в больших данных.

Список литературы

1. Хэн, Дж., Камбер, М., Пей, Дж. *Интеллектуальный анализ данных: концепции и методы*. — М.: Вильямс, 2019.
2. Goodfellow, I., Bengio, Y., Courville, A. *Deep Learning*. — MIT Press, 2016.
3. Géron, A. *Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow*. — O'Reilly Media, 2022.
4. LeCun, Y., Bengio, Y., Hinton, G. *Deep Learning*. — Nature, 2015.
5. Chollet, F. *Deep Learning with Python*. — Manning, 2021.