Урок 21. Как обучается нейронная сеть?

7 класс · Урок 21 из 34 · 45 минут

Обучение нейронной сети в 7 классе: как ошибка помогает изменять веса, зачем нужны новые данные и как отличить обучение от запоминания.

Сеть верно отвечает на знакомые картинки, но ошибается на новых. Как понять, что она научилась решать задачу, а не только запомнила образцы? Разберём, что меняется после ошибки и как проверить результат обучения.

▼ Содержание урока

1. Цель урока
2. Ход урока
3. Этап 1. Вход в тему: знакомое решение
4. Этап 2. Ответ, ошибка и изменение параметров
5. Этап 3. Почему одной удачной поправки мало
6. Этап 4. Исследование: чему поверить после обучения
7. Этап 5. Итоги урока

▼ 01. Цель урока

Понять, как сравнение ответа с образцом позволяет изменять параметры нейронной сети и проверять результат обучения.

К концу урока ученики смогут:

- Объяснять, чем обучение отличается от применения готовой сети.
- Связывать ошибку с изменением весов и смещений.
- Объяснять, зачем повторять обучение на разных примерах.
- Предлагать проверку, которая отличает запоминание от переноса на новые данные.

▼ 02. Ход урока

Этап	Мин	Учитель	Ученики
1. Вход в тему: знакомое решение	5	Выслушивает два воспоминания, различает воспроизведение и перенос.	Приводят собственный пример, формулируют вопрос об обучении.
2. Ответ, ошибка и изменение параметров	12	Объясняет прямой проход, потери и настройку параметров.	Прослеживают цикл по схеме, различают ответ и изменение весов.
3. Почему одной удачной поправки мало	10	Разбирает разнообразие почерка и риск переобучения.	Объясняют, зачем нужны отдельные новые примеры.
4. Исследование: чему поверить после обучения	12	Организует разработку и обмен испытаниями переноса.	Создают карточку исследования, обсуждают замечание читателей.
5. Итоги урока	6	Обсуждает общие выводы и ответы на итоговые вопросы.	Формулируют знания и объясняют ответы.

03. Этап 1. Вход в тему: знакомое решение 5 минут

Говорим ученикам

Вспомните задачу, решение которой вы однажды выучили, а потом не узнали похожую задачу с другими числами. Что именно изменилось? Поговорите с соседом и приведите один такой случай без оценки, кто лучше учится.

Нейронная сеть тоже может хорошо отвечать на знакомые примеры и ошибаться на новых. Сегодня разберём, как устроено обучение нейронной сети и почему уменьшение ошибки ещё не означает, что сеть освоила задачу. На прошлом занятии мы меняли входы отдельного нейрона. Теперь вопрос другой: как подобрать его веса, чтобы получать нужные ответы?

04. Этап 2. Ответ, ошибка и изменение параметров 12 минут

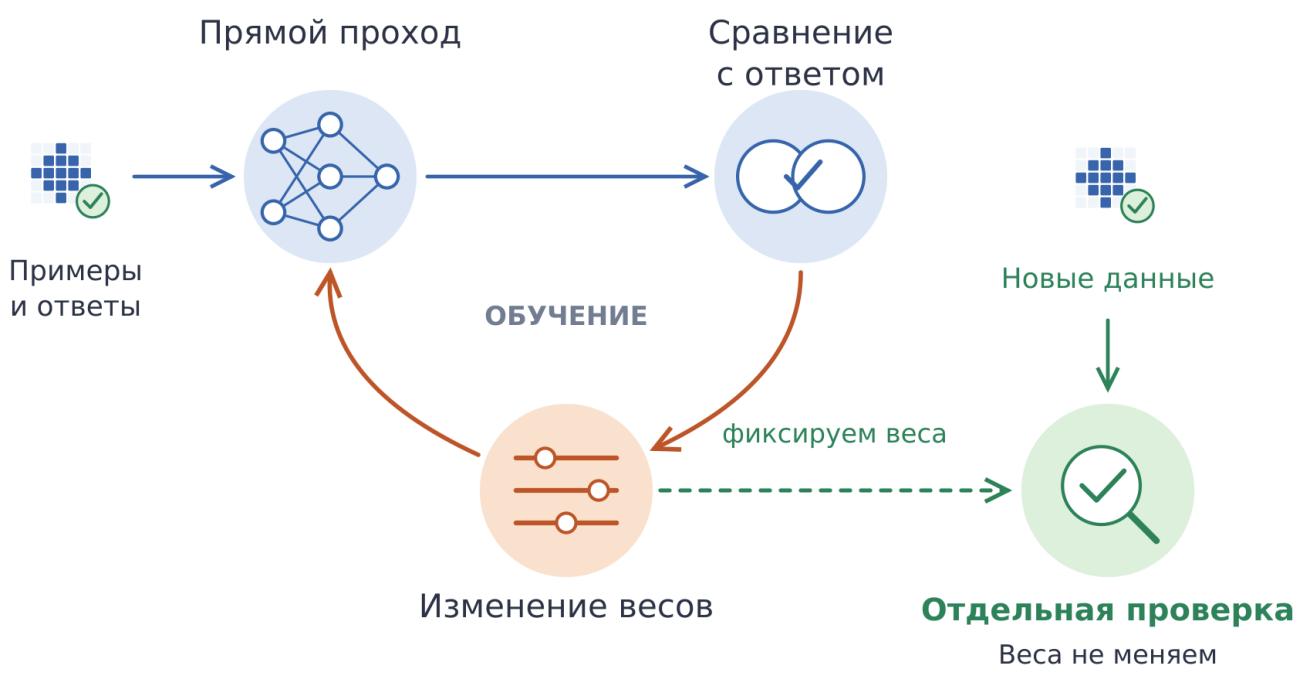
Говорим ученикам

У сети есть параметры — числа, которые определяют преобразование входа в выход. К ним относятся веса связей и смещения нейронов. Когда сетью пользуются после обучения, эти числа обычно остаются прежними. Новый вход может дать другой ответ. При обучении программа изменяет сами параметры.

Рассмотрим обучение с известными правильными ответами. Пример проходит через сеть, затем результат сравнивается с нужным ответом. Для сравнения задают функцию потерь — правило, которое превращает расхождение в число. Если сравнивать результаты на одних и тех же данных, меньшие потери означают более близкие к цели ответы. Для числового прогноза можно, например, брать квадрат разности.

Пусть сеть должна выдать 1, а выдала 0,4. Разность по модулю — 0,6, квадрат разности — 0,36. Эти числа описывают расхождение, но ещё не говорят, какой именно вес изменить. В многослойной сети алгоритм обратного распространения ошибки вычисляет, как малое изменение каждого параметра повлияет на потери. Алгоритм настройки использует эти сведения, чтобы сделать следующий шаг.

Слова «обратное распространение» не означают, что правильный ответ идёт назад через сеть как новая картинка. Алгоритм вычисляет это влияние от выходных слоёв к входным. Сам ответ получается прямым проходом от входа к выходу.



Стрелки по кругу показывают обучение с изменением весов. Пунктир ведёт к отдельной проверке на новых данных: в этом испытании веса не меняют и ответы не используют для настройки.

Для учителя

Короткая проверка: сеть получила новую фотографию, но веса не изменились. Это применение. Ответ 1 вместо 0,4 уменьшает указанные потери до нуля только на данном примере; вывод о других примерах пока не сделан.

05. Этап 3. Почему одной удачной поправки мало

10 минут

Говорим ученикам

Одна поправка может помочь одному примеру и ухудшить другой. Поэтому обучение повторяют на множестве примеров. Один полный проход по обучающему набору называют эпохой. Шаг настройки делают осторожным: слишком большой шаг способен перескочить удачное значение,

слишком маленький требует больше повторений.

Если сеть подстраивается под частные особенности знакомых примеров и хуже работает на новых, это переобучение. Чтобы заметить его, часть примеров заранее оставляют отдельно. На проверочных данных параметры не подгоняют. Если по ним снова и снова выбирают настройки, для итогового испытания нужен ещё один нетронутый набор.

В 1998 году Ян Лекун и соавторы описали системы чтения рукописных документов, включая банковские чеки — документы с записанной суммой. Разные люди пишут цифры по-разному, и запомнить одну красивую цифру недостаточно. После обучения сеть должна узнавать и незнакомые варианты написания.

Реальный случай: рукописная сумма вместо единого шрифта

В статье Лекуна, Ботту, Бенжио и Хаффнера 1998 года описана коммерческая система чтения банковских чеков, читавшая несколько миллионов чеков в день. Это уже была работа с разнообразным почерком, а не только с печатным шрифтом.

Общий вывод: полезное обучение позволяет применять найденные зависимости к новым примерам той же задачи. Поэтому успех оценивают отдельно от процесса настройки.

Для учителя

Не утверждайте, что нейросеть читала любой документ без ошибок. Статья описывает систему из нескольких модулей. Для разбора механизма обучения достаточно различия настройки и применения; производительность системы не равна точности отдельного нейрона.

06. Этап 4. Исследование: чему поверить после обучения 12 минут

Говорим ученикам

Перенос — это применение найденного принципа к новым примерам. До проверки нужно решить, по какому результату мы его заметим. Копия знакомого изображения проверяет повторение знакомого ответа. Новый документ другого автора проверяет более широкий круг вариантов. Но совершенно другая задача не становится честным испытанием только потому, что она сложнее.

Сейчас вы спроектируете небольшое исследование своей учебной трудности. На собственном опыте проверим, чем запоминание образца отличается от умения применить правило к новой задаче. В этом сравнении нас интересует только способ проверки.

Задание. Составьте испытание переноса

Выберите изучаемое сейчас умение: различать языковые конструкции, решать задачи определённого типа или читать музыкальный ритм. Составьте карточку, по которой другая пара сможет провести проверку.

1. Запишите, чему именно учат образцы, и придумайте два разных учебных примера.

2. Подготовьте новый пример для проверки и объясните, почему он проверяет тот же принцип, но не копирует образец.
3. Предложите один способ усложнить проверку, сохранив учебную цель.
4. Обменяйтесь карточками. Читатели находят место, где проверка может пройти за счёт запоминания. Исправьте план.

Критерии: понятна цель; обучение и проверка разделены; новое испытание требует того же умения; замечание читателя учтено.

Для учителя

Допустимы разные предметы. Например, для отличия главного и зависимого слова нужна новая фраза, а не перестановка имён в выученной. Этот пример не предлагайте заранее как готовую карточку. Не оценивайте реальную точность нейросети по человеческому опыту. Спросите, что в сравнении помогает понять проверку, а что остаётся различием человека и программы.

07. Этап 5. Итоги урока 6 минут

Говорим ученикам

При обучении программа настраивает параметры и тем самым меняет способ вычисления будущих ответов. Уменьшить потери на знакомых данных — промежуточная цель. Практический смысл обучения появляется, когда сеть справляется с новыми примерами своей задачи.

Главное за урок

- При обучении нейронной сети программа изменяет веса и смещения, ориентируясь на выбранную меру ошибки.
- Обратное распространение ошибки помогает определить влияние параметров на потери.
- Повторение обучения может привести к переобучению: успеху на знакомых данных и слабому переносу на новые.
- Отдельные данные для итогового испытания позволяют оценить сеть без подгонки под ответы проверки.

Итоговые вопросы

Ответьте коротко. Затем обсудим объяснения двух-трёх участников.

1. Чем отличается изменение ответа из-за нового входа от обучения нейронной сети?
2. Почему нулевая ошибка на одном примере не доказывает, что нейронная сеть обучилась всей задаче?

3. Как проверить, что система использует найденную закономерность на новых данных, а не только воспроизводит знакомые ответы?

<details>
<summary>Для учителя</summary>

1. Новый вход проходит через прежние параметры; обучение меняет сами параметры.
2. Один ответ не описывает результат на других вариантах; поправка может им повредить.
3. Заранее отложить новые примеры той же задачи, не использовать их ответы при настройке и отдельно оценить ошибки.

Принимаются ответы по смыслу. Важна связь с общим знанием, а не воспроизведение деталей реального случая.
</details>

<details>
<summary>Для учителя: источники</summary>

1. LeCun, Bottou, Bengio, Haffner. Gradient-Based Learning Applied to Document Recognition. Дата материала: ноябрь 1998. Материал старше 2024 года. Подтверждающая фраза: «reads several million cheques per day».
</details>