Урок 26. Обучение с учителем и классификация

7 класс · Урок 26 из 34 · 45 минут

Обучение с учителем и классификация: урок для 7 класса о примерах с метками, предсказании категории, разметке и отличии от регрессии.

Один учебный вопрос требует объяснить понятие, другой — проверить рассуждение. Как научить программу различать такие обращения? Разберём обучение с учителем и составим правила разметки, по которым смогут работать другие люди.

▼ Содержание урока

1. Цель урока
2. Ход урока
3. Этап 1. Вход в тему: какая помощь нужна
4. Этап 2. Признаки, метки и предсказания
5. Этап 3. Метка должна иметь смысл
6. Этап 4. Редакция разметки
7. Этап 5. Итоги урока

▼ 01. Цель урока

Понять обучение по примерам с известными ответами и разработать понятную разметку для задачи классификации.

К концу урока ученики смогут:

- Объяснять роль признаков и меток при обучении с учителем.
- Различать классификацию и регрессию по виду требуемого ответа.
- Составлять правила разметки и выявлять неоднозначные случаи.
- Предлагать отдельную проверку качества классификатора.

▼ 02. Ход урока

Этап	Мин	Учитель	Ученики
1. Вход в тему: какая помощь нужна	5	Просит записать реальные учебные вопросы, вводит задачу направления помощи.	Формулируют вопрос и обсуждают полезный тип ответа.
2. Признаки, метки и предсказания	10	Объясняет обучение по меткам и различие видов выхода.	Различают признаки, метки, категории и числа.
3. Метка должна иметь смысл	9	Объясняет, как решения сортировщика становятся метками для обучения модели.	Объясняют зависимость правильного ответа от стандарта.
4. Редакция разметки	15	Организует независимую разметку и разбор споров.	Создают руководство и испытывают его на отложенных вопросах.
5. Итоги урока	6	Обсуждает общие выводы и ответы на итоговые вопросы.	Формулируют знания и объясняют ответы.

03. Этап 1. Вход в тему: какая помощь нужна 5 минут

Говорим ученикам

Запишите один настоящий вопрос по предмету, который вам сейчас непонятен. Не пытайтесь придумывать сложную формулировку: важно понять, какой ответ вам помог бы. Поменяйтесь вопросами и скажите, какую помощь вы бы предложили. Возможны разные варианты.

Если учебный сервис должен направлять вопрос к нужному виду помощи, ему надо выбрать категорию. Это задача классификации. Сегодня разберём обучение с учителем: как известные ответы помогают модели научиться выбирать категории для новых примеров.

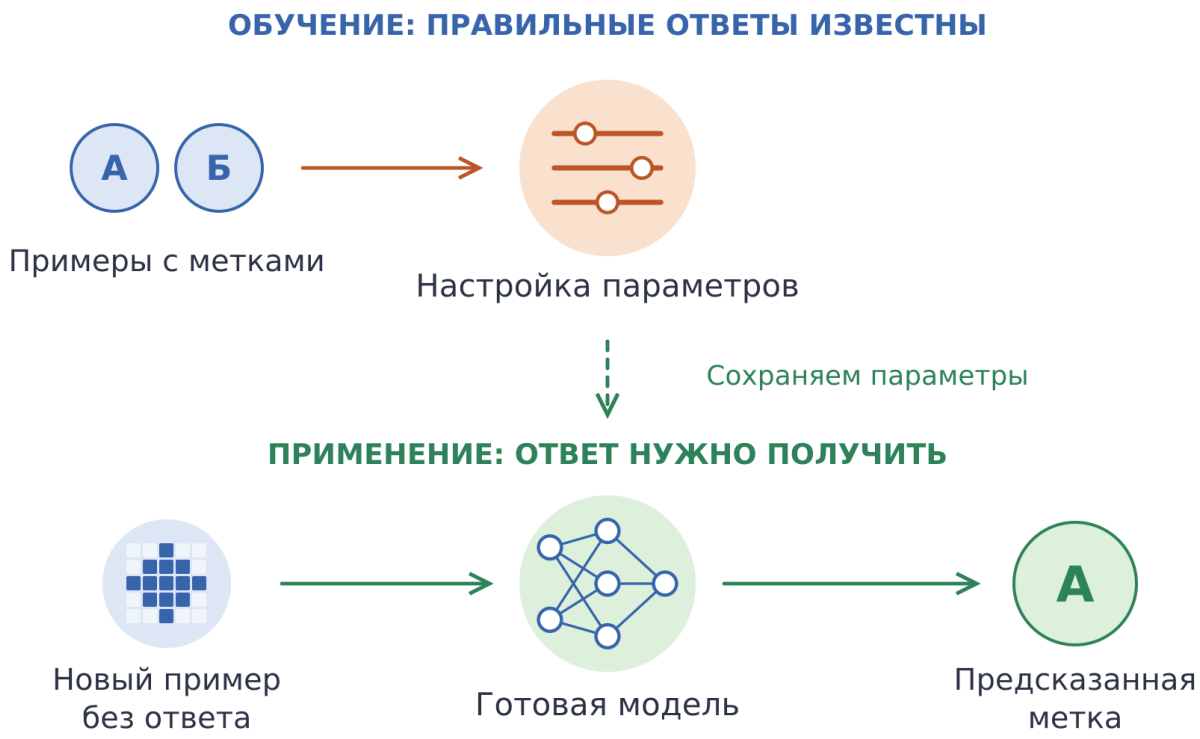
04. Этап 2. Признаки, метки и предсказания 10 минут

Говорим ученикам

При обучении с учителем у примеров известны целевые ответы. Метка — ответ, который мы хотим научиться предсказывать. В классификации это класс: например, один из видов обращения. Признаки — то, по чему модель должна получить ответ. Если дать модели правильную метку на вход, её повторение не покажет, что модель научилась решать задачу.

Учитель в этом названии — источник известных ответов, а не обязательно человек, стоящий рядом с компьютером. Метки могут быть получены людьми, измерением или из надёжной записи результата. Важно, чтобы они соответствовали задаче. Во время обучения программа сравнивает свой ответ с меткой и меняет параметры.

Когда моделью пользуются, известны признаки нового примера, а нужный ответ ещё неизвестен. Модель предсказывает класс. Позже ответ можно проверить. Если классификация ошиблась, нужно выяснить не только количество ошибок, но и какие классы она путает.



При обучении верные метки участвуют в настройке параметров. Затем готовая модель получает новый пример без ответа и предсказывает метку. Буквы обозначают условные классы.

Обучение с учителем не ограничено категориями. Если нужен числовой ответ — длительность, масса, значение температуры, — это регрессия. Различие определяется требуемым выходом. Одинаковые исходные данные могут использоваться для разных задач: предсказать минуты ожидания или класс «короткое/долгое ожидание».

Вернитесь к своей таблице опроса из урока 24. Какой столбец мог бы стать числовым ответом для регрессии, а какой — категорией для классификации? Какие другие признаки пришлось бы собрать для такого прогноза?

Для учителя

Например, минуты чтения — числовой ответ, предпочитаемый формат — категория. Опирайтесь на фактические столбцы группы. Номер записи не объясняет ответ; для прогноза нужны содержательно связанные признаки. Маленький опрос не обещает возможности обучить надёжную модель. Достаточно одного устного ответа на каждый тип задачи.

05. Этап 3. Метка должна иметь смысл 9 минут

Говорим ученикам

Категории не возникают из красивого названия. Нужно решить, какие различия полезны человеку, и описать их так, чтобы разные разметчики — люди, которые ставят метки, — понимали их одинаково. Если один и тот же текст получает разные метки из-за неясного правила, модель учится на противоречивых примерах.

31 августа 2016 года Google Cloud рассказала о Макото Коикэ, который помогал родителям на огуречной ферме в Японии. Его мать сортировала огурцы по принятой на ферме системе качества. Макото фотографировал уже отсортированные плоды, чтобы учить модель по этим образцам.

Целевой ответ был связан с принятой системой сортировки, а не с единственным «истинным классом огурца». Перед использованием на другой ферме нужно проверить её правила сортировки и условия съёмки. Нельзя просто перенести название категории и считать задачу прежней.

Реальный случай: опыт сортировщика стал метками

Макото Коикэ собрал около 7 тысяч фотографий огурцов, отсортированных его матерью, и использовал их для обучения. Так практический опыт сортировки был представлен как примеры с известными категориями.

Вывод: обучение с учителем зависит от того, какой ответ считается правильным. Поэтому качество и смысл разметки — часть задачи, а не мелкая техническая подготовка.

Для учителя

Фотографии не содержали всех свойств, доступных человеку. Не утверждайте, что модель полностью заменила сортировщика. Число фотографий относится к описанному проекту 2016 года.

06. Этап 4. Редакция разметки 15 минут

Говорим ученикам

Сейчас мы определим категории помощи для вопросов, которые действительно возникли у класса. Наша работа — создать понятное руководство разметки и испытать его. Это работа разработчиков набора, а не ручное изображение того, как нейросеть считает веса.

Задание. Разметьте учебные вопросы

В группе соберите 6–8 коротких вопросов по текущим школьным темам, без имён. Обсудите, какие 2–3 вида помощи по ним полезно различать. Сохраните хотя бы два вопроса отдельно до составления правил.

1. Опишите каждую категорию через цель ответа. Не используйте только слова-маркеры вроде «почему».
2. Два разметчика независимо размечают рабочие вопросы. Сравните решения и выберите один спорный вопрос для обсуждения.
3. Уточните правила так, чтобы объяснить спорный случай. Если спорных не было, придумайте вопрос, который трудно отнести только к одной категории и проверьте различие категорий.
4. Дайте другой паре правила и отложенные вопросы. Сопоставьте её разметку с вашим пониманием цели вопроса.
5. Решите, что изменить: формулировку вопроса, определение категории или возможность нескольких меток. Объясните выбор.

Результат — руководство с категориями и запись одного обсуждённого решения. Критерии: категории нужны для полезного ответа; спор разобран содержательно; отдельная проверка проведена; неоднозначность не спрятана.

Для учителя

Допустимы «объяснить понятие», «помочь выбрать способ», «проверить рассуждение» или другие обоснованные категории. Вопрос может требовать нескольких видов помощи; важно заранее решить, разрешены ли несколько меток. Если ученики выбрали разные метки для одного вопроса, выясните, что именно каждый понял по-своему. Даже совпавшие метки стоит сверить с тем, какой ответ нужен автору вопроса. На независимую разметку достаточно нескольких вопросов, не превращайте этап в длинную сортировку. Объясните отдельно: полноценный классификатор здесь не обучался.

07. Этап 5. Итоги урока 6 минут

Говорим ученикам

При обучении известный ответ помогает настраивать модель. Когда её используют на новом примере, ответ предстоит получить. Классификатор полезен, когда категории связаны с задачей, метки надёжны, а результат проверен на новых случаях. Числовой прогноз тоже можно учить по известным ответам, но это уже регрессия.

Главное за урок

- Обучение с учителем использует примеры с известными целевыми ответами.
- Классификация предсказывает категорию, а регрессия — числовое значение.
- Смысл меток и согласованность разметки влияют на то, чему учится модель.
- Новые примеры для проверки помогают обнаружить ошибки, которые не видны на знакомых данных.

Итоговые вопросы

Ответьте коротко. Затем обсудим объяснения двух-трёх участников.

1. В чём различие входного признака и правильной метки в обучении с учителем?
2. Как определить, является задача классификацией или регрессией?
3. Почему два согласных друг с другом разметчика ещё не гарантируют полезность категорий?
4. Что нужно сделать, если новый вопрос одинаково хорошо подходит двум категориям помощи?

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

1. Признак доступен для получения ответа; метка служит известной целью при обучении.
2. По выходу: категория или число.
3. Категории могут одинаково неверно отражать цель пользователя; согласованность — лишь одно условие.
4. Уточнить цель и правила; обоснованно выбрать главную метку, разрешить несколько меток или изменить категории.

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

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

1. Google Cloud. How a Japanese cucumber farmer is using deep learning and TensorFlow. Дата материала: 31.08.2016. Материал старше 2024 года. Подтверждающая фраза: «Makoto spent about three months taking 7,000 pictures of cucumbers sorted by his mother».
</details>