Урок 14. Как нейросеть создаёт изображения по текстовому запросу?

9 класс · Урок 14 из 34 · 45 минут

Вся программа 9 класса

Генерация изображений нейросетью зависит от того, как описана сцена. Начнём с рисунка, который показывает это различие.

Содержание урока

- 01. Цель урока
- 02. Ход урока
- 03. Этап 1. Введение: Где регистрация · 8 мин
- 04. Этап 2. От подписи к изображению · 13 мин
- 05. Этап 3. Свой промпт и проверка · 19 мин
- 06. Этап 4. Итоги урока · 5 мин

01. Цель урока

Объяснить, как модель учится связывать текст с изображениями, и сформулировать запрос, по которому результат можно проверить.

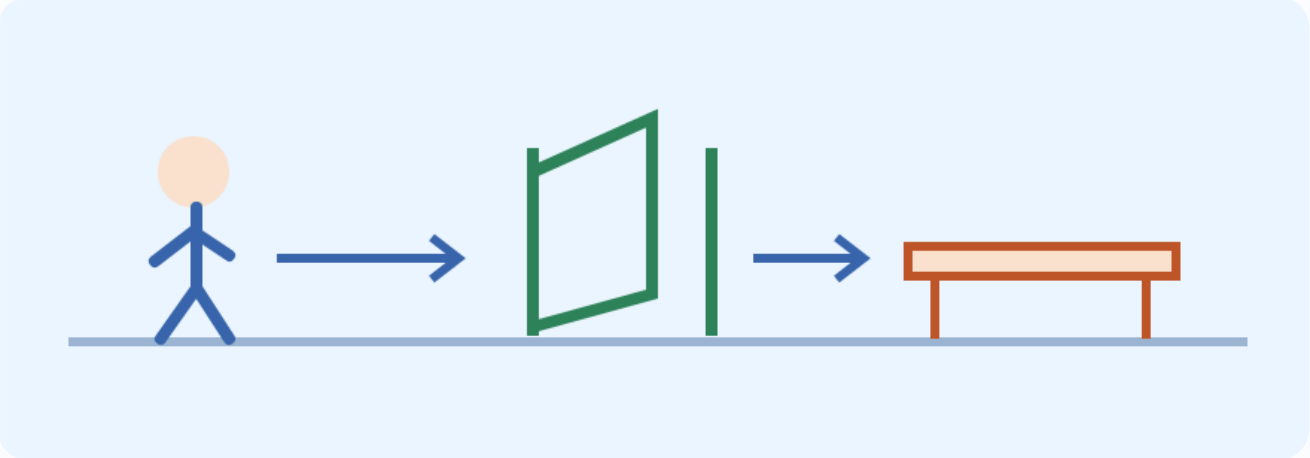
02. Ход урока

Этап	Мин	Учитель	Ученики
Этап 1. Введение: Где регистрация	8	Показывает схему с неполным описанием.	Сопоставляют описание со схемой и называют различие.
Этап 2. От подписи к изображению	13	Разводит обучение и создание новой сцены.	Сравнивают подписи и объясняют, чему помогает каждая.
Этап 3. Свой промпт и проверка	19	Предлагает записать критерии до проверки.	Готовят запрос и проверяют, можно ли увидеть условия.
Этап 4. Итоги урока	5	Подводит итог и сверяет ответы.	Письменно отвечают и сверяют решение.
Итого	45		

03. Этап 1. Введение: Где регистрация 8 мин

ГОВОРИМ УЧЕНИКАМ

Для дня открытых дверей нарисовали маршрут по запросу «Покажи калитку и стол регистрации». Редактор хотел, чтобы человек регистрировался до входа. Посмотрите на схему и попробуйте найти условие, которого не было в запросе. Сегодня разберём, как генерация изображений нейросетью связывает слова со сценой.



Учебная схема входа. Линия показывает путь посетителя.

Для учителя

Дайте время проследить путь по стрелкам. На схеме человек сначала проходит калитку. В запросе названы два объекта, но не сказано, в каком порядке посетитель их встречает.

04. Этап 2. От подписи к изображению 13 мин

ГОВОРИМ УЧЕНИКАМ

Модель учат на множестве примеров. Один из способов — дать ей изображение рядом с текстовым описанием. Подпись «остановка» сообщает, что в кадре за предмет. Подпись «синяя остановка справа от дороги» добавляет цвет и положение. Когда в обучении много таких примеров, модель может научиться связывать слова и их сочетания с деталями и отношениями в кадре.

Когда вы отправляете свой промпт, модель уже обучена. Она переводит текст в числовое представление, с которым умеет работать. Из этого представления модель получает подсказки, какую сцену строить. Она может сочетать знакомые признаки по-новому, но не обязана выполнить каждый пункт. Нечёткая фраза, необычное сочетание и сложные отношения между предметами делают задачу труднее.

Вот почему схема из начала урока могла выполнить короткий запрос и всё равно не подойти. На схеме есть нужные предметы, но запрос не уточняет, где стол должен стоять относительно калитки. Фраза «посетитель сначала подходит к столу перед калиткой, потом проходит внутрь» задаёт порядок. После уточнения проверьте, что на картинке можно пройти этим путём.

Момент	Что получает система	Что происходит
Обучение	Много изображений и подписей к ним	Настраивает внутренние параметры, связывая слова с признаками.
Генерация	Новый текстовый запрос	Строит новый вариант изображения с учётом текста.
Проверка	Изображение и исходная задача	Человек сверяет, выполнены ли важные условия.

Реальный случай. Яндекс обновил YandexART

В апреле 2024 года Яндекс сообщил об обновлении модели, которая работает в том числе в Шедевруме. При подготовке данных к изображениям добавляли более подробные описания. Подписи нужны на этапе обучения, чтобы модель могла лучше связать слова запроса с тем, что должно оказаться в кадре.

05. Этап 3. Свой промпт и проверка 19 мин

ГОВОРИМ УЧЕНИКАМ

Сейчас вам понадобится промпт, по которому другой человек сможет выбрать подходящий результат. Если он по вашему тексту не понимает, что с чем должно соприкасаться или в каком порядке идут действия, обозначьте это словами. Следующая попытка нужна там, где есть конкретный промах.

Иллюстрация для обмена семенами

Для школьного пункта обмена семенами нужен рисунок, который покажет одно понятное действие посетителя. Слова на карточку добавят позже. Сами решите, какие предметы и какое действие показать.

1. Составьте запрос. Опишите, какие два предмета связаны действием и какую деталь важно сохранить.
2. Запишите два признака, по которым узнаете удачный вариант. Передайте запрос соседней паре. Она делает быстрый эскиз или описывает, как поняла кадр.
3. По одной обнаруженной двусмысленности уточните запрос. Учитель показывает на общем экране одну генерацию по запросу группы. Сравните результат с этими же

признаками.

Для учителя

Допустимы, например, рука кладёт пакет в открытый ящик или посетитель берёт пакет с подписанной полки, если смысл понятен визуально. Проверяем конкретное действие и связь двух предметов, например пакет уже в руке или ещё лежит на полке. Для пробы используйте разрешённый режим Шедеврума на учительском устройстве. При отсутствии доступа оцените эскизы и запишите, какое условие затем нужно испытать в модели.

06. Этап 4. Итоги урока 5 мин

> ГОВОРИМ УЧЕНИКАМ
>
> Чтобы объяснить промах модели, сначала сравните её изображение с запросом. Потом решите, что проще: уточнить сцену или повторить генерацию с тем же условием.

Главное за урок

- Обучение связывает текстовые описания с визуальными признаками.
- Промпт направляет генерацию, а важные детали проверяет человек.
- Композицию легче оценивать по конкретному признаку, который можно увидеть.

Итоговые вопросы

1. Почему при обучении полезны подписи «лодка» и «лодка у правого берега»?
2. Как уточнить запрос, если лодка оказалась рядом с другим берегом?

Для учителя

Краткая подпись помогает связать слово с предметом, подробная добавляет отношение. Во втором случае нужно назвать положение лодки относительно нужного берега, например поставить её на воде вплотную к правому берегу.

Источники

1. Яндекс. Яндекс обновил YandexART · 24.04.2024. «сгенерированные нейросетью подробные описания изображений» Источник