Урок 11. Как ИИ работает с языком и переводом?

5 класс · Урок 11 из 17 · 45 минут

Вся программа 5 класса · Технологии искусственного интеллекта

Почему перевод зависит от соседних слов и обстоятельств. На уроке ученики разберутся, как обучают языковые модели, проверят перевод по контексту и составят запрос для работы с текстом.

Содержание урока

- 01. Цель урока
- 02. Ход урока
- 03. Этап 1. Зачем переводчику обстоятельства · 5 минут
- 04. Этап 2. Чему учится языковая модель · 11 минут
- 05. Этап 3. Когда переводу не хватает контекста · 12 минут
- 06. Этап 4. Как поручить модели работу с текстом · 11 минут
- 07. Этап 5. Что удалось проверить · 6 минут

01. Цель урока

Понять, как компьютер обрабатывает человеческий язык, и научиться сохранять смысл при переводе и сокращении текста.

К концу урока ученики смогут:

- объяснять различие между естественным языком и языком команд компьютеру;
- описывать обучение языковой модели на примерах и выбирать данные для задачи перевода;
- находить в контексте основания для выбора значения слова;
- выбирать перевод по исходнику и контексту, представляя решение в таблице;
- составлять запрос к языковой модели и проверять сохранность фактов в её ответе.

02. Ход урока

Этап	Мин	Учитель	Ученики
1. Зачем переводчику обстоятельства	5	Объясняет тему и предлагает языковую проблему	Обосновывают перевод по сообщению
2. Чему учится языковая модель	11	Объясняет принципы и показывает исторический пример	Создают пары контекстов и объясняют выбор данных
3. Когда переводу не хватает контекста	12	Объясняет выбор перевода по условиям и разбирает основания	Выбирают перевод четырёх строк и объясняют его по контексту
4. Как поручить модели работу с текстом	11	Проводит практику запроса и проверки	Составляют критерии, запрос и краткое сообщение
5. Что удалось проверить	6	Подводит к самостоятельному переносу знаний	Отвечают на новые вопросы и обосновывают решения
Итого	45		

03. Этап 1. Зачем переводчику обстоятельства 5 минут

ГОВОРИМ УЧЕНИКАМ

Сегодня разберёмся, как искусственный интеллект работает с языком и переводом. Автоматический перевод помогает прочитать сообщение на незнакомом языке, но иногда меняет его смысл. Мы выясним, откуда возникают такие ошибки. Затем проверим перевод по контексту и попробуем поручить языковой модели другую работу с текстом.

Реплика без видео

Для школьного видеопроекта прислали английскую реплику «I saw her duck», а сам ролик не прикрепили. Один ученик предложил перевод «Я видел её утку», другой — «Я видел, как она пригнулась». Как выяснить, какой перевод подходит к ролику?

Обсудите сообщение

Решите, хватает ли реплики для выбора. Назовите сведения, которые нужно запросить у автора видео.

Словарь: *I saw* — я видел; *her* — её или сочетание «она» в конструкции «видел, как она…»; *duck* — утка или пригнуться.

Для учителя

Оба перевода возможны. Слово *duck* может называть птицу или действие. Нужны сам ролик, описание происходящего или соседние реплики. Если героиня показывает свою птицу, подходит первый вариант; если она опускает голову, чтобы уклониться от мяча, — второй. Без этих сведений нельзя объявить один перевод ошибочным.

ГОВОРИМ УЧЕНИКАМ

Одна и та же реплика допускает два смысла. Чтобы выбрать, нам нужно знать, что происходит в ролике. Соседние слова и обстоятельства сообщения называют контекстом. Переводчик, который получил только эту реплику, тоже не получил сведений для однозначного выбора.

04. Этап 2. Чему учится языковая модель 11 минут

ГОВОРИМ УЧЕНИКАМ

Русский, английский и другие языки общения людей называют естественными языками. Мы пользуемся ими, чтобы рассказывать, спрашивать и договариваться. Язык программирования нужен для записи команд компьютеру по строгим правилам. В человеческом сообщении многое зависит от того, о чём идёт речь: слово «снять» может относиться и к одежде, и к фотографии.

Обработка естественного языка — работа компьютера с такими сообщениями. Например, программа переводит текст, находит в нём дату или составляет краткий пересказ. Перевод меняет язык сообщения, а краткий пересказ уменьшает его объём.

Языковая модель — программа, которую обучили находить связи в текстах. Компьютер представляет части текста числами и выполняет вычисления. Во время обучения настройки модели меняют так, чтобы она лучше справлялась с учебными задачами. При работе с новым сообщением модель использует найденные связи, а не ищет обязательную готовую копию этого сообщения.

Для перевода полезны пары: предложение на одном языке и его проверенный перевод на другом. Они показывают соответствия между выражениями. Большие наборы текстов на одном языке тоже полезны: по ним изучают сочетания слов. Но они не содержат готового перевода каждого предложения на другой язык. Разные способы обучения могут сочетаться.

Исходное выражение	Второй язык	Что связывает пара
Идёт дождь.	It is raining.	Сообщение об осадках
Поезд идёт к станции.	The train is moving towards the station.	Сообщение о движении

ГОВОРИМ УЧЕНИКАМ

В обеих русских строках есть слово «идёт». В английских предложениях используются разные выражения, потому что и действия разные. Такие пары помогают обучать перевод с учётом сообщения. Если каждый раз заменять «идёт» одним английским словом, смысл будет теряться. Даже обученная модель может неверно применить найденную связь к новому тексту.

Задание 1. Дайте модели разные контексты

В паре составьте два коротких предложения со словом «снять», чтобы в них были разные действия. У каждого предложения подпишите действие другими словами. Затем решите: для обучения переводу чего не хватает вашему набору?

Для учителя

Возможная пара: «Пассажир снял рюкзак» — убрал с плеч; «Репортёр снял матч» — записал на видео. Другие сочетания допустимы, если значение действительно различается. Нужны проверенные переводы на второй язык. Два придуманных предложения помогают понять устройство данных, но не достаточны для обучения полноценного переводчика.

ГОВОРИМ УЧЕНИКАМ

Вы составили сообщения, в которых одно слово обозначает разные действия. Для перевода нужны соответствующие выражения на втором языке. Поэтому обучающие данные выбирают по задаче и проверяют их качество. Ошибочные примеры тоже могут научить модель неверным связям.

Реальный случай: предложение вместо отдельных кусочков

ГОВОРИМ УЧЕНИКАМ

27 сентября 2016 года исследователи Google сообщили о запуске нейросетевого перевода с китайского на английский в Google Translate. Новый подход рассматривал предложение целиком и помогал учитывать связи между его частями. Разработчики назвали и оставшиеся ошибки: программа могла пропускать слова и неверно переводить имена. История показывает, зачем переводчику связи внутри предложения и почему улучшение системы не отменяет проверку.

05. Этап 3. Когда переводу не хватает контекста 12 минут

ГОВОРИМ УЧЕНИКАМ

Короткие подписи иногда отправляют на перевод без фотографии или предыдущего сообщения. Тогда переводчик видит слова, но не знает, о каком предмете говорят. Несколько переводов могут быть допустимы. Проверим, какие сведения позволяют выбрать нужный смысл.

Задание 2. Проверьте памятку мастерской

Ученики готовят двуязычную памятку для школьной мастерской. Английские строки прислали отдельно от фотографий. Для каждой строки предложены два перевода. Прочитайте описание фотографии и выберите подходящий вариант.

№	Английская строка	Варианты перевода	Что происходит на фотографии
1	Check the scale.	«Проверьте весы». «Проверьте масштаб».	На экране чертёж детали. Перед печатью сравнивают размеры на чертеже с размерами будущей детали.
2	Move the file.	«Переместите файл». «Переложите напильник».	Рядом с деревянной заготовкой лежит металлический инструмент с насечками. Он мешает закрепить заготовку.
3	Close the case.	«Закройте футляр». «Закройте корпус».	После замены детали компьютер собран и проверен. Осталось поставить на место боковую крышку.
4	The glasses are in the box.	«Очки в коробке». «Стаканы в коробке».	В мастерской готовят выставку. Для гостей принесли кувшин лимонада и ищут, во что его налить.

Словарь: *check* — проверьте; *scale* — весы или масштаб; *move* — переместите, переложите; *file* — файл или напильник; *close* — закройте; *case* — футляр или корпус; *glasses* — очки или стаканы; *in the box* — в коробке.

1. В тетради составьте таблицу: номер строки → выбранный перевод → основание из описания фотографии.
2. Выберите одну строку и придумайте другую ситуацию, в которой подошёл бы второй перевод. Английскую строку не меняйте.
3. Обсудите: можно ли без фотографии уверенно отвергнуть второй вариант? Какую информацию нужно передать переводчику?

Результат: четыре решения с основаниями и одна новая ситуация для другого перевода.

Для учителя

1. «Проверьте масштаб»: сравнивают размеры чертежа и детали. 2. «Переложите напильник»: речь о физическом инструменте, который мешает работе. 3. «Закройте корпус»: боковая крышка относится к компьютеру. 4. «Стаканы в коробке»: их ищут для лимонада.

Для второго перевода подходят, например: проверка весов перед взвешиванием материала; перенос файла с чертежом в другую папку; уборка инструмента в футляр; поиск защитных очков перед шлифовкой. Засчитывайте другую жизненную ситуацию, если она поддерживает выбранное значение и не противоречит английской строке.

В таблице даны варианты смысла, а не ответы конкретной программы. Без контекста нельзя считать один из вариантов безусловно неверным. Причина неопределённости — у слова несколько значений, а переводчику не передали фотографию или её описание.

Засчитывайте перевод с конкретным основанием. Одного ответа «так логичнее» недостаточно. Разрешите пользоваться словарём: здесь проверяют работу с контекстом.

ГОВОРИМ УЧЕНИКАМ

В каждой строке было слово с несколькими значениями. Размеры чертежа, инструмент у заготовки, крышка компьютера и лимонад помогли выбрать перевод. Когда мы изменили обстоятельства, подошёл другой вариант той же строки. Поэтому перевод проверяют по словам и обстоятельствам сообщения.

06. Этап 4. Как поручить модели работу с текстом 11 минут

ГОВОРИМ УЧЕНИКАМ

Работа с языком нужна не только для перевода. Извлечение сведений помогает отдельно получить из сообщения срок или место. Сокращение помогает передать тот же материал меньшим числом слов. Чтобы результат подошёл, человек должен назвать задачу и проверить, какие сведения сохранились. Сейчас мы сами зададим условия, а затем сравним с ними ответ модели.

Задание 3. Объявление для школьного подкаста

«До пятницы, 18:00, пришлите в школьный радиокружок один вопрос о звёздах. Вопросы принимают от пятиклассников. Имя можно не указывать. Выпуск запишут в понедельник, а ответы опубликуют во вторник».

1. Выпишите в таблицу все сведения, которые нельзя потерять. Разделите срок приёма вопросов, день записи и день публикации.
2. Напишите запрос: получите из исходного текста объявление из двух предложений с тем же смыслом.
3. Учитель вводит исходник и запрос одной пары в доступную школе языковую модель. Класс проверяет фактический ответ по своей таблице.

4. Исправьте расхождения или объясните, почему ответ можно оставить. Подготовьте окончательные два предложения.

Результат: таблица сведений, ваш запрос и проверенное сокращение. Заимствовать из ответа можно только то, что подтверждается исходником.

Для учителя

Обязательные сведения: адресат — школьный радиокружок; участники — пятиклассники; один вопрос о звёздах; срок — до пятницы 18:00; имя необязательно; запись в понедельник; публикация ответов во вторник.

Возможный запрос: «Сократи объявление до двух предложений. Сохрани адресата, участников, тему и число вопросов, срок, необязательность имени, день записи и день публикации. Новых сведений не добавляй».

Возможное сокращение: «Пятиклассники могут до пятницы, 18:00, прислать в школьный радиокружок один вопрос о звёздах, не указывая имя. Выпуск запишут в понедельник, ответы опубликуют во вторник». Это образец для сверки, а не заранее известный ответ сервиса.

Отправляет текст учитель со своего устройства. Используйте только приведённое обезличенное сообщение. Если модель недоступна, пары пишут сокращения сами и обмениваются ими для проверки. Такой вариант отрабатывает работу с текстом; опыт с действующей моделью остаётся невыполненным.

Не предсказывайте, где модель ошибётся: фиксируйте то, что получилось. Полный ответ сохраняет все сведения. В запросе должны быть ясны действие, длина ответа и условия сохранения смысла.

ГОВОРИМ УЧЕНИКАМ

Мы заранее составили таблицу, с которой можно сравнить ответ. Благодаря этому дни записи и публикации не смешиваются со сроком приёма вопросов. Точный запрос задаёт проверяемые условия, но не гарантирует их выполнение. При переводе, извлечении сведений и сокращении нужен один и тот же контроль: совпадает ли результат с исходным сообщением.

07. Этап 5. Что удалось проверить 6 минут

ГОВОРИМ УЧЕНИКАМ

Мы выяснили, что компьютер учится языковым задачам по примерам. При переводе значение слова зависит от сообщения и обстоятельств. Если часть сведений не передана, иногда подходят несколько вариантов. Проверка требует исходника и понятных условий результата. Теперь каждый применит эти выводы самостоятельно.

Итоговые вопросы

1. Почему фраза на человеческом языке может требовать контекста, а команда на языке программирования должна соблюдать строгие правила? Назовите одну задачу обработки естественного языка.
2. Для переводчика собрали фотографии предметов и тексты на одном языке. Какие примеры прямо покажут соответствия между языками?
3. Подпись «The crane is near the water» допускает варианты «Журавль у воды» и «Подъёмный кран у воды». Автор добавил: «It is lifting a boat» — «Он поднимает лодку». Какой перевод теперь подходит? Почему по первой строке нельзя было уверенно выбрать один вариант?
4. Модель верно сократила одно сообщение. Можно ли без проверки публиковать её следующий ответ? Что сравнить с исходником?

Для учителя

1. Одни слова могут передавать разные значения в разных обстоятельствах; компьютерные команды записывают по заданным правилам. Пример языковой задачи: перевод, краткий пересказ или извлечение сведений. 2. Предложения с проверенными переводами на другой язык. 3. «Подъёмный кран у воды»: он поднимает лодку. В первой строке не было сведений, различающих птицу и механизм; слово *crane* допускает оба значения. Уточнение дало основание для выбора. 4. Нет. Нужно сравнить участников и действия, числа, сроки, условия и отрицания, существенные для нового текста.

> **ГОВОРИМ УЧЕНИКАМ**
>
> Обработка естественного языка помогает компьютеру выполнять задачи с сообщениями людей. Модель находит связи в обучающих данных и использует их для новых текстов. Мы проверяем не только отдельные слова, но и смысл действий, сроков и условий. Если контекста недостаточно, уточняем сообщение.

← Урок 10 · Все уроки 5 класса · Урок 12 →