

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Емельянов Сергей Геннадьевич

Должность: ректор

Дата подписания: 06.09.2024 12:47:59

Уникальный программный ключ:

9ba7d3e34c012eba476ffd2d064cf2781953be730df257481679c0ee536f01e6

МИНОБРАЗОВАНИЯ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра программной инженерии

УТВЕРЖДАЮ
Проректор по учебной работе
О.Б. Локинова
06.09.2024 г.



ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В РАЗЛИЧНЫХ ОБЛАСТЯХ ДЕЯТЕЛЬНОСТИ

Методические рекомендации по выполнению практических работ для
студентов направления подготовки 07.04.01 Архитектура

Курск 2024

УДК 004

Составитель Т.Н. Конаныхина

Рецензент

д.т.н. Таныгин М.О.

Искусственный интеллект в различных областях деятельности: методические рекомендации по выполнению практических работ направления подготовки 07.04.01 Архитектура / Юго-Зап. гос. ун-т; сост.: Т.Н. Конаныхина, Курск, 2024. 82 с.

Содержат методические рекомендации к выполнению практических по дисциплине «Искусственный интеллект в различных областях деятельности».

Методические указания по структуре, содержанию и стилю изложения материала соответствуют методическим и научным требованиям, предъявляемым к учебным и методическим пособиям.

Предназначены для студентов направления подготовки 07.04.01 Архитектура.

Текст печатается в авторской редакции

Подписано в печать 18.05. Формат 60x84 1/16
Усо.печ.л.0,29. Уч.-изд.л.0,26. Тираж 100 экз. Заказ: 294. Бесплатно.
Юго-Западный государственный университет.
305040. г. Курск, ул. 50 лет Октября, 94.

Инструкция по технике безопасности

1. Общие требования безопасности

1.1. К работе на персональном компьютере допускаются лица, прошедшие обучение безопасным методам труда, вводный инструктаж, первичный инструктаж на рабочем месте.

1.2. При эксплуатации персонального компьютера на человека могут оказывать действие следующие опасные и вредные производственные факторы:

- повышенный уровень электромагнитных излучений;
- повышенный уровень статического электричества;
- пониженная ионизация воздуха;
- статические физические перегрузки;
- перенапряжение зрительных анализаторов.

1.3. Студент при выполнении лабораторных работ обязан:

1.3.1. Выполнять только ту работу, которая определена в методических указаниях.

1.3.2. Содержать в чистоте рабочее место.

1.3.3. Соблюдать режим труда и отдыха.

1.3.3. Соблюдать меры пожарной безопасности.

1.4. Рабочие места с компьютерами должны размещаться таким образом, чтобы расстояние от экрана одного видеомонитора до тыла другого было не менее 2,0 м, а расстояние между боковыми поверхностями видеомониторов - не менее 1,2 м.

1.5. Рабочие места с персональными компьютерами по отношению к световым проемам должны располагаться так, чтобы естественный свет падал сбоку, преимущественно слева.

1.6. Оконные проемы в помещениях, где используются персональные компьютеры, должны быть оборудованы регулирующими устройствами типа: жалюзи, занавесей, внешних козырьков и др.

1.7. Рабочая мебель для пользователей компьютерной техникой должна отвечать следующим требованиям:

- высота рабочей поверхности стола должна регулироваться в пределах 680 - 800 мм; при отсутствии такой возможности высота рабочей поверхности стола должна составлять 725 мм;
- рабочий стол должен иметь пространство для ног высотой не менее 600 мм, глубиной на уровне колен не менее 450 мм и на уровне вытянутых ног не менее 650 мм;

2. Требования безопасности перед началом работы

2.1. Подготовить рабочее место.

2.2. Отрегулировать освещение на рабочем месте, убедиться в отсутствии бликов на экране.

2.3. Проверить правильность подключения оборудования к электросети.

2.4. Проверить исправность проводов питания и отсутствие оголенных участков проводов.

2.5. Проверить правильность установки стола, стула, подставки для ног, пюпитра, угла наклона экрана, положение клавиатуры, положение "мыши" на специальном коврике, при необходимости произвести регулировку рабочего стола и кресла, а также расположение элементов компьютера в соответствии с требованиями эргономики и в целях исключения неудобных поз и длительных напряжений тела.

3. Требования безопасности во время работы

3.1. Студенту при работе на ПК запрещается:

- прикасаться к задней панели системного блока (процессора) при включенном питании;

- переключать разъемы интерфейсных кабелей периферийных устройств при включенном питании;

- допускать попадание влаги на поверхность системного блока (процессора), монитора, рабочую поверхность клавиатуры, дисководов, принтеров и других устройств;

- производить самостоятельное вскрытие и ремонт оборудования;
- работать на компьютере при снятых кожухах;
- отключать оборудование от электросети и выдергивать электровилку, держась за шнур.

3.2.Продолжительность непрерывной работы с компьютером без регламентированного перерыва не должна превышать 2-х часов.

3.3.Во время регламентированных перерывов с целью снижения нервно - эмоционального напряжения, утомления зрительного анализатора, устранения влияния гиподинамии и гипокинезии, предотвращения развития познотонического утомления выполнять комплексы упражнений.

4. Требования безопасности в аварийных ситуациях

4.1.Во всех случаях обрыва проводов питания, неисправности заземления и других повреждений, появления гари, немедленно отключить питание.

4.2.Не приступать к работе до устранения неисправностей.

4.3.При получении травм или внезапном заболевании немедленно известить организовать первую доврачебную помощь или вызвать скорую медицинскую помощь.

5. Требования безопасности по окончании работы

5.1.Отключить питание компьютера.

5.2.Привести в порядок рабочее место.

5.3.Выполнить упражнения для глаз и пальцев рук на расслабление.

Практическая работа №1

Основные концепций и методы искусственного интеллекта. Автоматизированное машинное обучение в Azure Machine Learning

Цель работы - изучить возможности автоматизированного машинного обучения в Студии машинного обучения Azure и использование его для обучения и оценки модели машинного обучения.

В этой работе вы будете использовать функции автоматического машинного обучения в Azure Machine Learning для обучения и оценки модели машинного обучения. Затем вы развернете и протестируете обученную модель.

Перейдите по ссылке и воспользуйтесь инструкциями портала.

<https://microsoftlearning.github.io/mslearn-ai-fundamentals/Instructions/Labs/01-machine-learning.html>

Создайте рабочую область машинного обучения Azure

Чтобы использовать Azure Machine Learning, вам необходимо предоставить Azure Machine Learning workspace в вашей подписке Azure. Тогда вы сможете использовать Azure Machine Learning studio для работы с ресурсами вашего workspace.

Совет: Если у вас уже есть рабочая область машинного обучения Azure, вы можете воспользоваться ею и перейти к следующей задаче.

1. Войдите на портал Azure по адресу <https://portal.azure.com>, используя свои учетные данные Microsoft.
2. Выберите + **Создать ресурс**, выполните поиск по машинному обучению и создайте новый ресурс для **машинного обучения Azure** со следующими настройками:

- Подписка:** Ваша подписка на Azure.
- Группа ресурсов:** Создайте или выберите группу ресурсов.
- Имя:** Введите уникальное имя для вашего рабочего пространства.
- Регион:** Выберите ближайший географический регион.
- Учетная запись хранилища:** Обратите внимание на новую учетную запись хранилища по умолчанию, которая будет создана для вашей рабочей области.
- Хранилище ключей:** Обратите внимание на новое хранилище ключей по умолчанию, которое будет создано для вашей рабочей области.
- Анализ приложений:** Обратите внимание на новый ресурс application insights по умолчанию, который будет создан для вашего рабочего пространства.
- Реестр контейнеров:** Отсутствует (реестр будет создан автоматически при первом развертывании модели в контейнере).

3. Выберите **Просмотреть** + **создать**, затем выберите **Создать**. Подождите, пока будет создано ваше рабочее пространство (это может занять несколько минут), а затем перейдите к развернутому ресурсу.

4. Выберите **Launch studio** (или откройте новую вкладку браузера и перейдите к <https://ml.azure.com> и войдите в Azure Machine Learning studio, используя свою учетную запись Microsoft). Закройте все отображаемые сообщения.

5. В Azure Machine Learning studio вы должны увидеть свое недавно созданное рабочее пространство. Если нет, выберите **Все рабочие области** в меню слева, а затем выберите только что созданное рабочее пространство.

Используйте автоматизированное машинное обучение для обучения модели

Автоматизированное машинное обучение позволяет вам опробовать множество алгоритмов и параметров для обучения нескольких моделей и определить наилучший для ваших данных. В этом упражнении вы будете использовать набор исторических сведений о прокате велосипедов для подготовки модели, которая прогнозирует количество прокатов велосипедов, которое следует ожидать в данный день, на основе сезонных и метеорологических особенностей.

1. В Azure Machine Learning studio просмотрите страницу **Automated ML** (в разделе **Разработка**).

2. Создайте новое автоматизированное задание ML со следующими настройками, используя **Next** по мере необходимости для прохождения пользовательского интерфейса:

Основные настройки:

- **Название работы:** mslearn-bike-automl
- **Название нового эксперимента:** mslearn-прокат велосипедов
- **Описание:** Автоматизированное машинное обучение для прогнозирования проката велосипедов
- **Теги:** нет

Тип задачи и данные:

- **Выберите тип задачи:** Регрессия
- **Выберите набор данных:** создайте новый набор данных со следующими настройками:
 - **Тип данных:**
 - **Название:** прокат велосипедов
 - **Описание:** Исторические данные о прокате велосипедов
 - **Тип:** Табличный

•Источник данных:

- Выберите из веб-файлов

•Веб-АДРЕС:

- Веб-АДРЕС: <https://aka.ms/bike-rentals>
- Пропустить проверку данных: не выбирать

•Настройки:

- Формат файла: с разделителями
- Разделитель: запятая
- Кодировка: UTF-8
- Заголовки столбцов: заголовки есть только в первом файле
- Пропуск строк: Отсутствует
- Набор данных содержит многострочные данные: не выбирать

•Схема:

- Включить все столбцы, кроме пути
- Ознакомьтесь с автоматически определяемыми типами

Выберите Создать. После создания набора данных выберите набор данных для проката велосипедов, чтобы продолжить отправку автоматизированного задания ML.

Настройки задач:

- Тип задачи: Регрессия
- Набор данных: прокат велосипедов
- Целевой столбец: Арендная плата (целое число)
- Дополнительные настройки конфигурации:
 - Первичный показатель: нормализованная среднеквадратичная ошибка
 - Объясните наилучшую модель: Невыбранный

- **Используйте все поддерживаемые модели:** Невыбраны. Вы ограничите задание опробованием только нескольких конкретных алгоритмов.
- **Разрешенные модели:** Выберите только **randomForest** и **LightGBM** — обычно вы хотели бы попробовать как можно больше, но каждая добавленная модель увеличивает время, необходимое для выполнения задания.
- **Ограничения:** Расширьте этот раздел
 - **Максимальное количество испытаний:** 3
 - **Максимальное количество одновременных испытаний:** 3
 - **Максимальное количество узлов:** 3
 - **Пороговое значение показателя:** 0,085 (таким образом, если модель достигает значения показателя среднеквадратичной ошибки 0,085 или меньше, задание завершается.)
 - **Время ожидания:** 15
 - **Время ожидания итерации:** 15
 - **Включить досрочное расторжение:** выбрано
- **Валидация и тестирование:**
 - **Тип проверки:** разделение обучения и проверки
 - **Процент достоверности данных:** 10
 - **Тестовый набор данных:** Отсутствует

Вычисления:

- **Выберите тип вычисления:** Бессерверный
- **Тип виртуальной машины:** процессор
- **Уровень виртуальной машины:** выделенный
- **Размер виртуальной машины:** Standard_DS3_V2*
- **Количество экземпляров:** 1

* Если ваша подписка ограничивает доступные вам размеры виртуальных машин, выберите любой доступный размер.

3. Отправьте учебное задание. Оно запускается автоматически.

4. Подождите, пока работа не будет завершена. Это может занять некоторое время — сейчас самое подходящее время для кофе-брейка!

Обзор лучшей модели

Когда задание по автоматизированному машинному обучению будет завершено, вы сможете просмотреть лучшую модель, которую оно подготовило.

1. На вкладке "**Обзор**" задания по автоматизированному машинному обучению обратите внимание на краткое описание наилучшей модели.

The screenshot shows a web interface for a machine learning task. At the top, there are tabs: 'Обзор' (Overview), 'выходных данных' (Output data), 'моделей' (Models), and 'ограничений данных + журналы Дочерних заданий' (Data constraints + logs of child tasks). The 'Обзор' tab is selected. Below the tabs, there are two main panels. The left panel, titled 'Свойства' (Properties), shows the task status as 'Завершено' (Completed) with a green checkmark. Other fields include 'Созданный на' (Created on), 'Время начала' (Start time), 'Продолжительность' (Duration), 'Длительность вычисления' (Computation time), 'Вычислить цель' (Calculate objective), 'Имя' (Name), and 'Имя скрипта' (Script name). The right panel, titled 'Входные данные' (Input data), shows 'Входное имя: training_data' and 'Набор данных: прокат велосипедов: 1'. Below this, there is a section titled 'Краткое описание лучшей модели' (Best model summary). It lists 'Название алгоритма' (Algorithm name) as 'Максимальный масштаб, LightGBM', which is highlighted with a yellow box. Below that, it shows 'Гиперпараметры' (Hyperparameters) with a link to 'Просмотр гиперпараметров' (View hyperparameters). It also displays 'Нормализованная среднеквадратичная ошибка' (Normalized root mean square error) as 0.08049 and 'Отбор проб' (Sampling) as 100.00%. At the bottom, it indicates 'Зарегистрированные модели' (Registered models) as 'Регистрации пока нет' (No registrations yet) and 'Состояние развертывания' (Deployment status) as 'Развертывания пока нет' (No deployments yet).

Примечание Вы можете увидеть сообщение под статусом “Предупреждение: достигнут указанный пользователем результат выхода ...”. Это ожидаемое сообщение. Пожалуйста, переходите к следующему шагу.

2. Выберите текст под **названием алгоритма** для выбора наилучшей модели, чтобы просмотреть ее детали.

3. Перейдите на вкладку **Метрики** и выберите графики **остатков** и **predicted_true**, если они еще не выбраны.

Просмотрите диаграммы, показывающие производительность модели. Диаграмма **остатков** показывает остатки (различия между прогнозируемыми и фактическими значениями) в виде гистограммы. Диаграмма **predicted_true** сравнивает прогнозируемые значения с истинными значениями.

Развертывание и тестирование модели

1. На вкладке "**Модель**" выберите лучшую модель, подготовленную в ходе вашего задания по автоматизированному машинному обучению, выберите "**Развернуть**" и используйте опцию "**Веб-сервис**" для развертывания модели со следующими настройками:

- **Название:** predict-аренда
- **Описание:** Прогнозирование аренды велосипедов
- **Тип вычисления:** экземпляр контейнера Azure
- **Включить аутентификацию:** выбрано

2. Дождитесь начала развертывания - это может занять несколько секунд. Состояние **развертывания** для конечной точки **predict-rentals** будет указано в основной части страницы как Запущенное.

3. Подождите, пока состояние **развертывания** изменится на Успешно. Это может занять 5-10 минут.

Протестируйте развернутый сервис

Теперь вы можете протестировать свой развернутый сервис.

1. В Azure Machine Learning studio в меню слева выберите **Конечные точки** и откройте конечную точку для **прогнозирования аренды** в режиме реального времени.

2. На странице **прогнозировать аренду** конечной точки в реальном времени перейдите на вкладку **Тест**.

3. На панели **Входные данные для тестирования конечной точки** замените шаблон JSON следующими входными данными:

КодКопировать

```
{  
  "Inputs": {  
    "data": [  
      {  
        "day": 1,  
        "mnth": 1,  
        "year": 2022,  
        "season": 2,  
        "holiday": 0,  
        "weekday": 1,  
        "workingday": 1,  
        "weathersit": 2,  
        "temp": 0.3,  
        "atemp": 0.3,  
        "hum": 0.3,  
        "windspeed": 0.3  
      }  
    ]  
  },  
  "GlobalParameters": 1.0  
}
```

4.Нажмите кнопку **Тест**.

5.Посмотрите результаты тестирования, которые включают прогнозируемое количество прокатов на основе входных функций - аналогично этому:

КодКопировать

```
{
```

```
"Results": [  
  444.27799000000000  
]  
}
```

Панель тестирования взяла входные данные и использовала модель, которую вы обучили, чтобы вернуть прогнозируемое количество арендных ставок.

Давайте рассмотрим, что вы сделали. Вы использовали набор исторических данных о прокате велосипедов для обучения модели. Модель прогнозирует количество проката велосипедов, ожидаемое в данный день, на основе сезонных и метеорологических особенностей.

Очистка

Созданный вами веб-сервис размещен в экземпляре контейнера Azure. Если вы не собираетесь экспериментировать с ним дальше, вам следует удалить конечную точку, чтобы избежать ненужного использования Azure.

1. В Azure Machine Learning studio на вкладке **Конечные точки** выберите конечную точку для **прогнозирования аренды**. Затем выберите **Удалить** и подтвердите, что вы хотите удалить конечную точку.

Удаление вашего компьютера гарантирует, что с вашей подписки не будет взиматься плата за вычислительные ресурсы. Однако с вас будет взиматься небольшая сумма за хранение данных, пока в вашей подписке существует Azure Machine Learning workspace. Если вы завершили изучение машинного обучения Azure, вы можете удалить рабочую область машинного обучения Azure и связанные с ней ресурсы.

Как удалить свое рабочее пространство:

1. На портале Azure на странице **Группы ресурсов** откройте группу ресурсов, указанную вами при создании рабочей области машинного обучения Azure.

2.Нажмите **Удалить группу ресурсов**, введите название группы ресурсов, чтобы подтвердить, что вы хотите ее удалить, и выберите **Удалить**.

Практическая работа №2

Применение нейросетей для работы с текстом

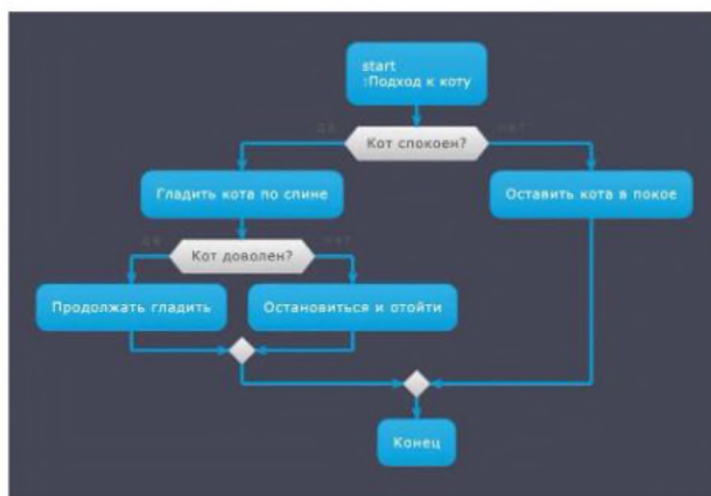
Цель работы - изучение различных готовых нейросетевых моделей, предназначенных для задач обработки текста, получение практического опыта применения готовых нейросетевых моделей в конкретных сценариях.

Искусственная нейросистема - это тоже самое, что и органическая, только в математической форме

Отличие нейронных сетей от обычных алгоритмов вычисления - способность обучения. Это обучение заключается в нахождении верных коэффициентов связи между нейронами, а также в обобщении данных и выявлении сложных зависимостей между выходными и входными сигналами

Пример обычного алгоритма

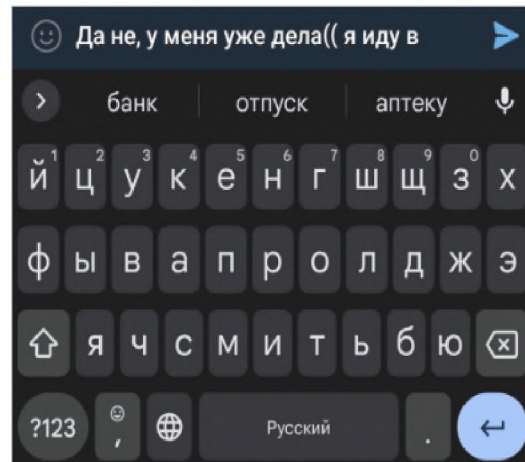
Как гладить кота?



На самом деле, ChatGPT — это T9 из вашего телефона, но умнее. Ученые называют обе этих технологии «языковыми моделями» (Language Models); а всё, что они по сути делают, — это угадывают, какое следующее слово должно идти за уже имеющимся текстом.

Вы получаете сообщение от друга: «Пойдем гулять?». Вы начинаете печатать в ответ и вот тут подключается T9, который подсказывает в смартфонах более уместные слова.

Не я иду в пианино, а
(проверить прямо сейчас)



Как вообще предсказывать зависимости одних вещей от других? Предположим, мы хотим научить компьютер предсказывать вес человека в зависимости от его роста — как подойти к этой задаче?

Необходимо провести большое количество измерений роста и веса различных людей, потом данные каждого человека нанести в виде точки в системе координат и провести анализ. Как мы видим, есть прямая зависимость между ростом и весом различных людей. Видим, что зависимость линейная, через скопление точек проводим прямую, определяем формулу, которая её описывает.



Можно упрощенно сказать, что те же самые T9 или ChatGPT — это подобранные уравнения, которые пытаются предсказать следующее слово

(игрек) в зависимости от набора подаваемых на вход модели предыдущих слов (иксов). Сколько этих уравнений в лингвистической модели?

На конференции AI Journey разработчики представили новую версию сервиса GigaChat. В её основе лежит одна из самых сильных продакшн-моделей для русского языка. По результатам тестов новый GigaChat превзошёл схожие по количеству параметров иностранные аналоги в бенчмарке MMLU.

Бенчмарк — эталонный тест производительности компьютерной системы.



GPT-3- содержит 175 миллиардов параметров, GPT-4 более 1500 миллиардов параметров.

Основная задача при тренировке языковой модели на наборе данных — подобрать такие коэффициенты при этих иксах, чтобы они действительно отражали какую-то зависимость.

Эволюция нейросетевых языковых моделей



Рассмотрим скачки в эволюции лингвистических моделей.

GPT-2 был улучшен по двум ключевым направлениям: набор тренировочных данных (датасет) и размер модели (количество параметров).

На тот момент не было каких-то специальных, больших и качественных, публичных наборов текстовых данных для тренировки языковых моделей.

Программисты OpenAI и решили поступить остроумно: они пошли на самый популярный англоязычный онлайн-форум Reddit и скачали все гиперссылки из всех сообщений, имевших более трех лайков. Всего таких ссылок вышло порядка 8 миллионов, а скачанные из них тексты весили в совокупности 40 гигабайт.

Собрание сочинений Уильяма Шекспира (всех его пьес, сонетов и стихов) состоит из 850'000 слов. В среднем на одной странице книги помещается около 300 английских слов — так что 2800 страниц английского текста за авторством величайшего англоязычного писателя займет в памяти компьютера примерно 5,5 мегабайт. Это в 7300 раз меньше, чем объем тренировочной выборки GPT-2. С учетом того, что люди в среднем читают по странице в минуту, даже если вы будете читать текст 24 часа в сутки без перерыва на еду и сон — вам потребуется почти 40 лет, чтобы догнать GPT-2.

Вышедший в 2020 году GPT-3, уже могла похвастаться 175 миллиардами параметров (в 116 раз больше GPT-2) Нейросеть при этом стала весить 700 гигабайт. Набор данных для обучения GPT-3 увеличился примерно в 10 раз до 420 гигабайт (добавили много книг, Википедию, и еще множество текстов с разных интернет-сайтов) Сама модель теперь имеет размер больше (700 Гб), чем весь массив текста для ее обучения (420 Гб). Получается парадокс: наш «нейромозг» в процессе изучения ”сырых” данных генерирует информацию о разных взаимозависимостях внутри них, которая превышает по объему исходную информацию

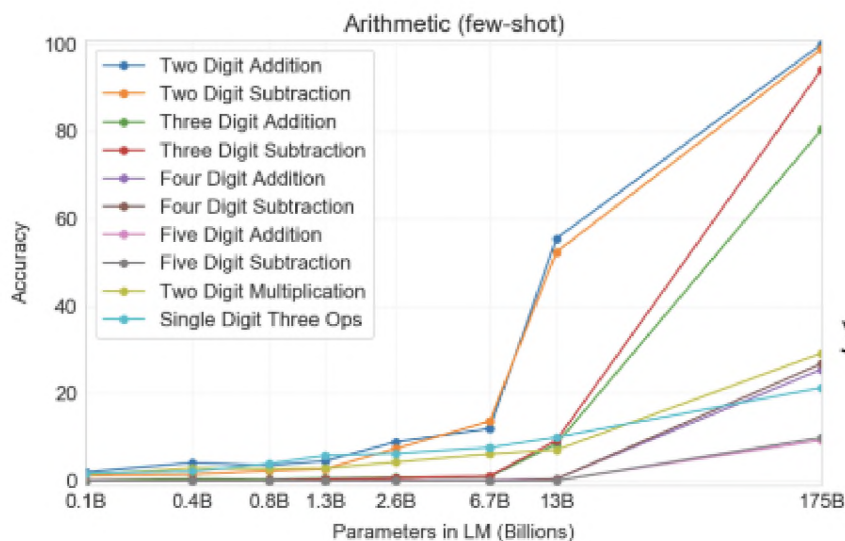
Запрос, который вы даете нейросети, называется промпт (prompt).

Напишите в GIGACHat $25+25=$ и отправьте сообщение



На самом деле лингвистическую нейросеть специально не обучали математике, она получила эту способность при кратном увеличении количества параметров.

На графике показана точность ответов нейросетей с разным количеством параметров на задачи, связанные со сложением/вычитанием и умножением чисел вплоть до пятизначных. При переходе от моделей с 10 миллиардами параметров к 100 миллиардам — нейросети становятся математиками.

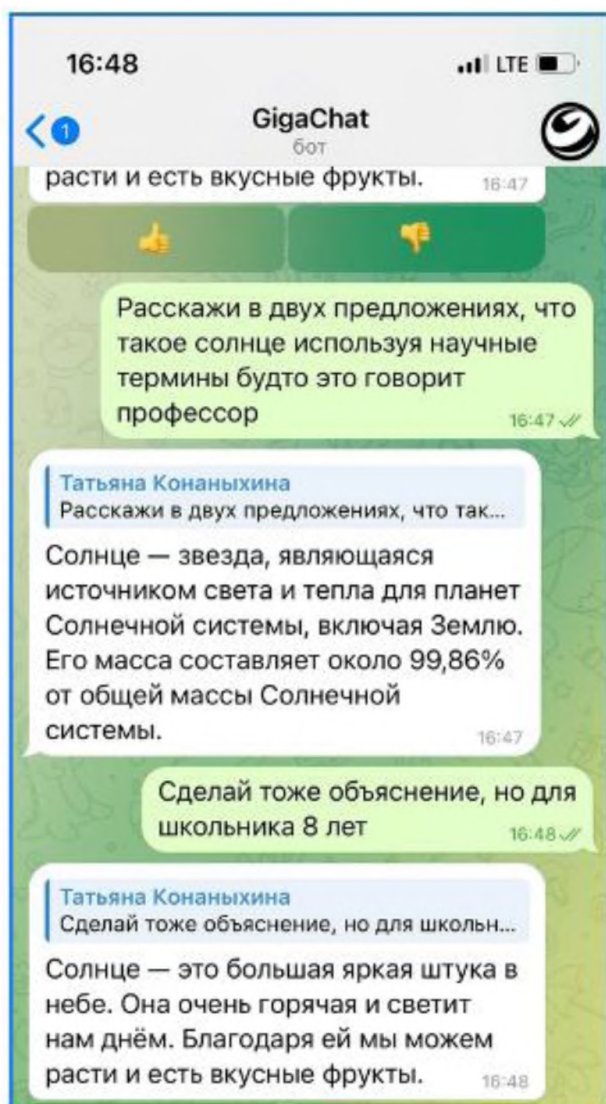


Языковую модель обучали продолжать тексты словами, а она при этом смогла сама разобраться в том, что если ей печатают « $378 + 789 =$ », то на это надо отвечать именно «1167», а не каким-то другим числом.

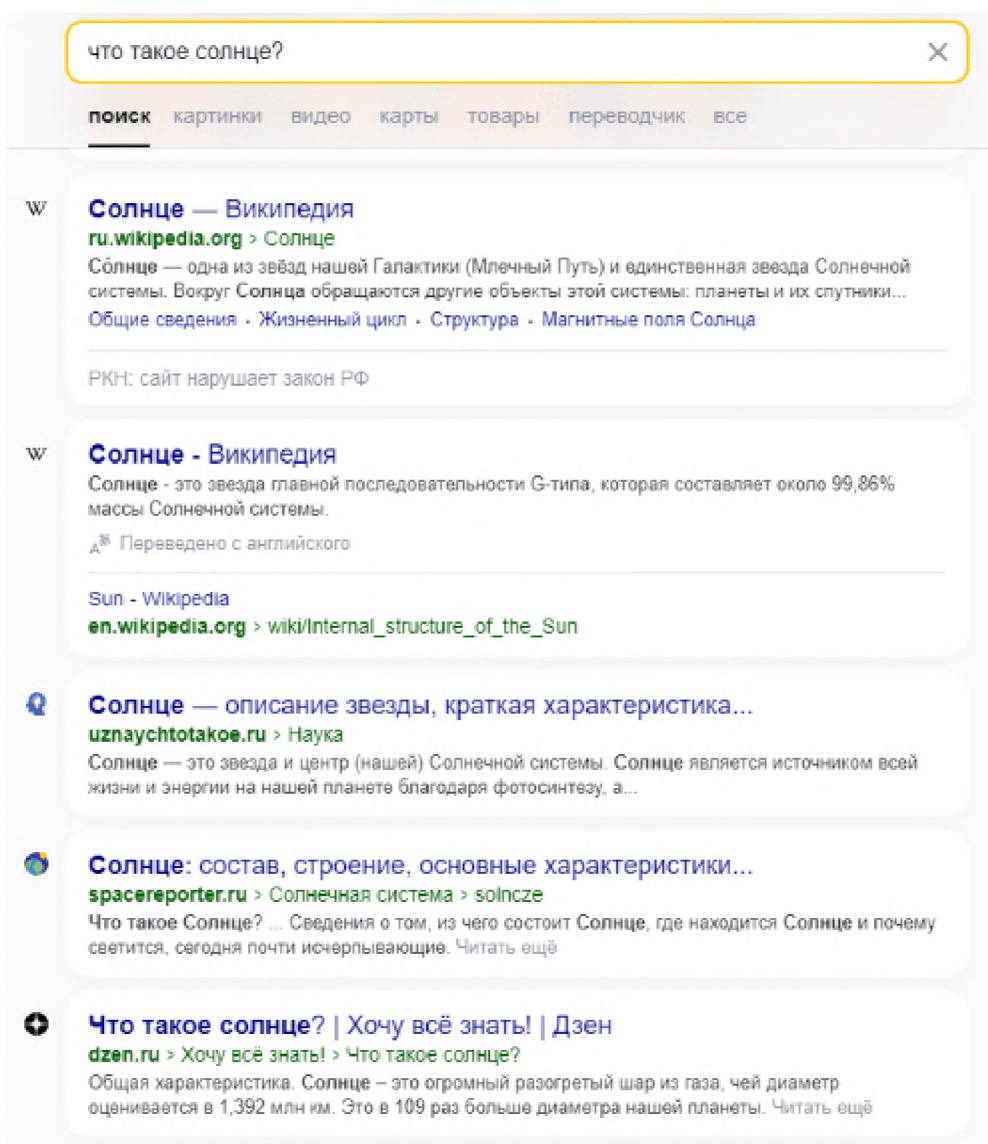
Одна и та же модель даже на одинаковые запросы может давать разные варианты ответа — прямо как живой человек.

При этом для описания одного и того же события или факта люди придумывают множество способов в разных стилях, тонах и полутонах. Подход к языковой коммуникации у профессора и, к примеру, у учеников младшей школы будет разным.

Это означает, что вы получаете персональный ответ в зависимости от контекста и истории запросов.



На запрос в поисковой системе вы получаете много ссылок, а нейросеть дает один адаптированный ответ.



Одна из самых интересных функций ChatGPT заключается в том, что он может писать от лица конкретного человека или представителя профессии.

С точки зрения менеджера по продукту опишите устройство Яндекс Станция.

С точки зрения сиделки в доме престарелых, опишите устройство Яндекс Станция

Напишите оба промпта, проанализируйте получившийся результат.

Расширим эту идею и увидим, что ИИ может даже проводить эксперименты, сравнения, дискуссии и приводить аргументы и контр-аргументы.

Напишите два промпта по очереди

Запрос 1. Давай рассмотрим некоторые вопросы, которые возникли при создании Яндекс Станции: 2018 год. Siri вышла лет 7 назад, год назад выпустили первую Apple HomePod, но без поддержки русского языка. Сцена - собрание директоров Яндекса, где только что было предложено создать Яндекс Станцию - умное устройство с голосовым ассистентом "Алиса" внутри для русскоговорящей аудитории. Перечислите аргументы "за" и "против", которые вероятно обсуждали члены совета на этом собрании в рамках процесса принятия решения о том, следует ли одобрить затраты на разработку устройства. Не стесняйся также включать участие экспертов по инженерному проектированию и сторонников продукта, если это дает более всестороннюю перспективу.

Запрос 2. Заменить в запросе текст "Перечислите аргументы "за" и "против"" на "Представьте аргументы за и против в виде диалога"

Вывод 1: нейросеть умеет адаптировать ответ под промпт.

Вывод 2: чем лучше промпт, тем лучше ответ нейросети.

Принцип GIGO (Garbage In, Garbage Out) - при неверных входящих данных будут получены неверные результаты, даже если сам по себе алгоритм правильный.

Со-основатель OpenAI Джон Шульман говорил: «Наше самое большое опасение заключалось в достоверности данных, потому что модель любит выдумывать всякое». То, что модель "выдумала", называется галлюцинация.

Вывод 3. Не верить всему, что вам говорит нейросеть.

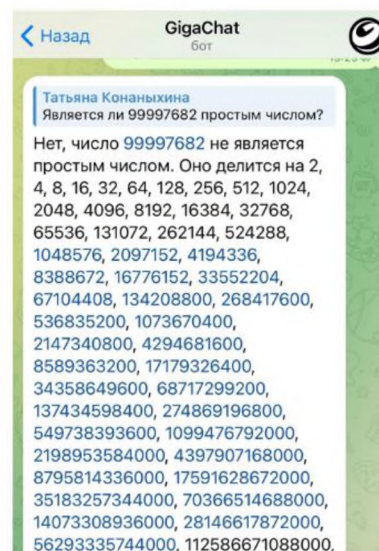
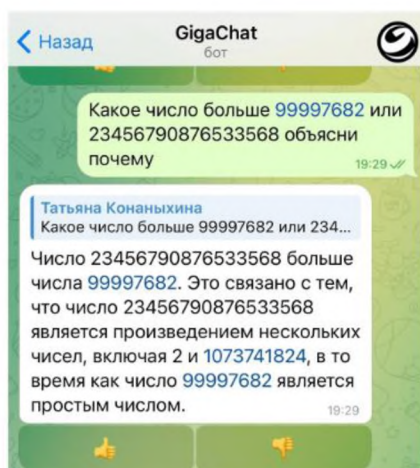
В «Яндексе» называют две причины, по которым нейросети галлюцинируют: Первая заключается в самом принципе работы: модель «читает» последовательность слов и предсказывать на её основе следующее.

Во время этого стоит задача не дать фактически верный ответ, а ответить так, как бы отвечал человек. Вторая причина галлюцинаций – в том, что предобучение нейросети осуществляется при помощи больших массивов данных, которые не всегда содержат фактически верную информацию (например, нейросеть обучается на базе ответов на форумах).

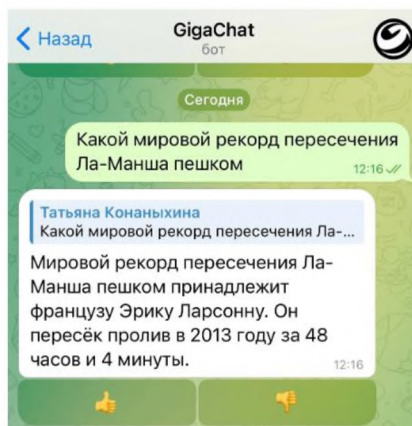
«Таким образом, внутри больших языковых моделей заложены не глубокие знания о мире, а лишь основанное на статистике понимание, как должны строиться тексты, и как должны соотноситься между собой запросы и ответы»

Приведу несколько примеров галлюцинаций нейросети. Можете повторить и проверить, как вам ответит нейросеть.

Пример 1.



Пример 2.



Ла-Манш — Английский канал, пролив между северным берегом Западной Европы и о. Великобритания. Вместе с проливом Па-де-Кале (Дуврский пролив) соединяет Северное море с Атлантическим океаном. Его невозможно перейти пешком. Названный нейросетью человек даже не переплывал его.

Самые быстрые [\[править \]](#)

Рекорд	Страна происхождения	Пловец	Время	Дата
Мужчины	 Германия	Andreas Waschburger ^[39]	06:45	Сентябрь 2023 года
Женщины	 Чешская Республика	Yveta Hlaváčová	07:25	2006
Мужчины двумя путями	 Новая Зеландия	Филип Раш	16:10	1987
Женщины двумя путями	 Австралия	Сьюзи Марони	17:14	1991
Мужчины тремя путями	 Новая Зеландия	Филип Раш	28:21	1987
Женщины тремя путями	 Великобритания	Элисон Стритер	34:40	1990
Четыре способа	 США	Сара Томас	54:10	2019

Как писать промпты правильно

- Общайтесь с ИИ так, как вы бы говорили с человеком
- Погружайте ИИ в контекст
- Не стесняйтесь отправлять запрос еще раз и вносить в него

изменения

- Одна тема - один чат
- Закроете или удалите диалог, старт нового диалога будет давать

новые ответы

- Ответы ~500 слов могут обрываться, напишите ”продолжай”
- Если вам нужен чёткий ответ, ограничивать длину запроса
- Вы можете исправлять и уточнять запросы, исходя из того, как

ИИ отвечал ранее, вступайте в диалог.

- Используйте роли (профессии, персонажи, исторические личности)Если вам нужны источники, попросите его подтвердить или обосновать свои ответы

- Приводите примеры
- Продолжайте экспериментировать

Идеальный промпт = цель+ задача+ контекст и детали+ роль+ лимиты

Цель - это результат, к которому мы хотим прийти (я хочу запустить телеграмм-канал про моду и выбираю название)

Задача - то, что должна сделать нейросеть (чем конкретнее, тем лучше)

Роль - кто сейчас нейросеть (она студент, профессор, журналист)

Контекст и детали - вся дополнительная информация, которая улучшит запрос

Я[желание или конечная цель]
Объясни/помоги/расскажи мне[задача]
Обязательно учти[контекст или важные детали]
Расскажи мне так, будто ты/представь, что ты [роль]
Уложишься в[лимит]

Я хочу провести веселый киновечер с компанией друзей. Посоветуй нам фильм, лучше всего комедию. Обязательно учти, что один из друзей не любит хорроры, и все мы не хотим смотреть европейское кино. Пусть фильм длится меньше двух часов. Договоримся, что ты — большой знаток американских комедий в стиле «Один дома». Предложи не больше 5 фильмов.

Исследователи выяснили, что для существенного улучшения результатов генерации нужно добавить «let's think step by step» (давай подумаем шаг за шагом). Оказалось, что это побуждает модель рассуждать последовательно, делать выводы на основе собственных суждений и приходить к правильному ответу гораздо чаще, чем без этой фразы.

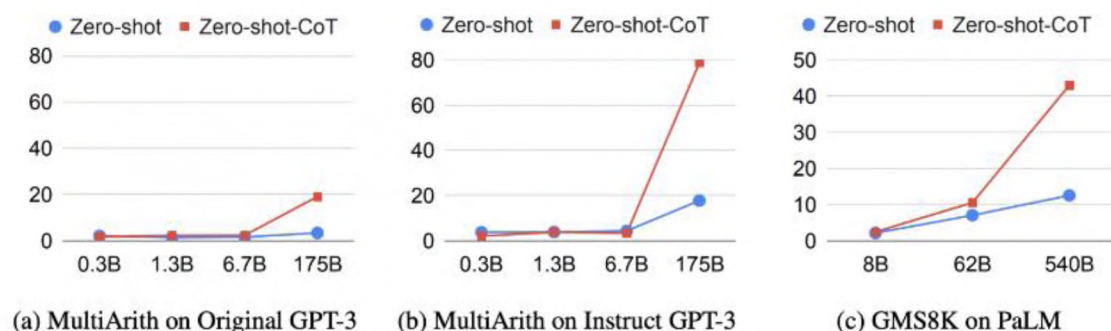


Figure 3: Model scale study with various types of models. See Appendix E for the detail.

Такой эффект значимо проявляется только у больших языковых моделей - ось OX, где указаны размеры моделей (в миллиардах параметров), по оси OY указаны метрики. Видно, как добавление одной фразы существенно увеличивает качество - но только при соответствующих размерах.

Запомните, что нейросеть!

- Не может генерировать авторский текст, будут общие фразы (у нейросети нет личного опыта)
- Не может ссылаться на достоверные источники (если нет доступа к интернету)
- Нейросеть может галлюцинировать, если текст большой (ответ ChatGPT - 2048 символов)
- Нейросети могут генерировать неоригинальный текст

Какие тексты можно генерировать с ChatGPT

Рабочее письмо

Нужны контекст и личный опыт?	Нет
Нужны актуальные данные?	Нет
Нужен оригинальный текст?	Нет
Нужен очень объемный текст?	Нет
Стоит ли генерировать	✓

Короткий доклад

Нужны контекст и личный опыт?	Нет
Нужны актуальные данные?	Нет
Нужен оригинальный текст?	Нет
Нужен очень объемный текст?	Нет
Стоит ли генерировать	✓

Монография

Нужны контекст и личный опыт?	Да
Нужны актуальные данные?	Да
Нужен оригинальный текст?	Да
Нужен очень объемный текст?	Да
Стоит ли генерировать	✗

Журналистское расследование

Нужны контекст и личный опыт?	Да
Нужны актуальные данные?	Да
Нужен оригинальный текст?	Да
Нужен очень объемный текст?	Да
Стоит ли генерировать	✗

Новость

Нужны контекст и личный опыт?	Да
Нужны актуальные данные?	Да
Нужен оригинальный текст?	Да
Нужен очень объемный текст?	Нет
Стоит ли генерировать	✗

Как редактировать тексты с помощью ChatGPT.

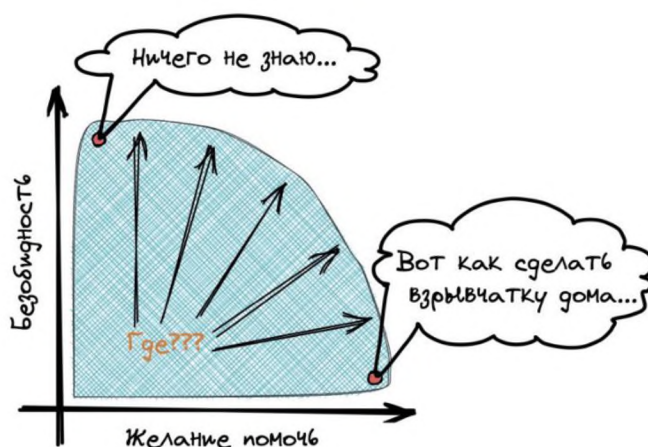
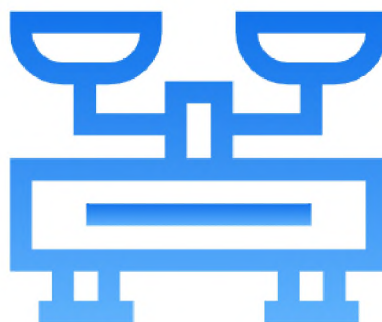
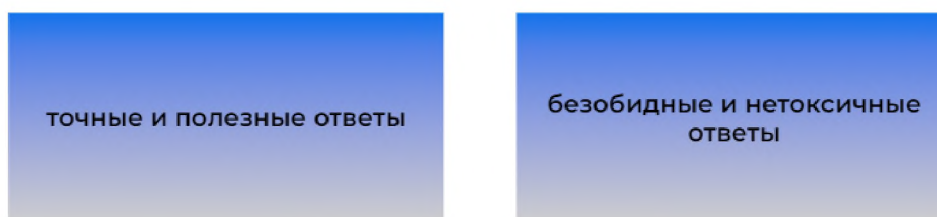
- Попробуйте доработать текст в том же диалоге
- Попросите нейросеть придумать план, чтобы сделать текст логичным и последовательным

- НЕОБХОДИМО доработать план под себя (самому убирать или добавлять что-то)!
- ChatGPT может исправить ошибки в тексте
- Он может улучшить текст (узнайте у нейросети, что можно добавить в этот текст, а что лишнее)
- Попросите перефразировать отрывки, которые вам не нравятся (сложное - уточнить, банальное - заменить)

Зачем чат GPT обучали на обратной связи от живых людей?

Нейросети обучали на наборе текстов из интернета, и лингвистические нейросети имеют доступ в Интернет, а в интернете далеко не все, что написано, полезно.

Возникает противоречие



Люди оценивали вручную насколько результат работы нейросети совпадает с их ожиданием. А языковая модель, училась решать еще одну дополнительную задачу – «как мне поменять свой сгенерированный ответ таким образом, чтобы он получил наибольшую оценку от человека?»

Для чего еще можно применять лингвистические нейросети.

- Для перевода текста
- Для транскрипции
- Для составления аннотации или выводов

Практическая работа №3

Применение нейросетей для работы с изображениями, аудио и видео

Цель работы - изучение различных готовых нейросетевых моделей, предназначенных для задач изображениями, аудио и видео, получение практического опыта применения готовых нейросетевых моделей в конкретных сценариях.

Работа с изображениями при помощи нейросетей

В контексте работы с изображениями нейронные сети могут использоваться для классификации изображений на основе их содержимого, идентификации и анализа объектов внутри изображения, генерации новых изображений на основе существующих.



Классификация изображений



Анализ изображений



Генерация изображений

Классификация изображений возможна по различным категориям, таким как различные виды животных или различные типы объектов. Применение в биологии, сельском хозяйстве, производстве, медицине.

Анализ изображений применяется, например, для медицинских изображений при обнаружении раковых клеток или других аномалий, для анализа спутниковых изображений для отслеживания землепользования и т.д.

Генерация изображений - это создание реалистичных изображений лиц, пейзажей и других объектов. Применение в разработке видеоигр, виртуальной реальности и даже искусстве.

За последнее время качество генерируемых изображений сильно выросло.



Февраль 2022



Март 2023

На сегодняшний день некоторые нейросети создают изображение такого качества и детальной проработки, что пользователь не может отличить изображение, сгенерированное от обычной фотографии, что доказывает ряд экспериментов, проведенных учеными Евросоюза, США и Австралии.

Например, рассмотрим исследование амстердамского университета, в котором группе европейцев показали вперемешку 100 изображений людей, сгенерированные нейросетью, и 100 настоящих фотографий, и попросили их указать, где на фотографиях настоящие люди.

В результате 61% сгенерированных фото были опознаны как настоящие, 49% фотографий реальных людей были опознаны, как созданные нейросетью.

Также, ученые из университета Беркли установили, что созданные искусственным интеллектом лица вызывают больше доверия, чем реальные лица.

Как сгенерировать изображений

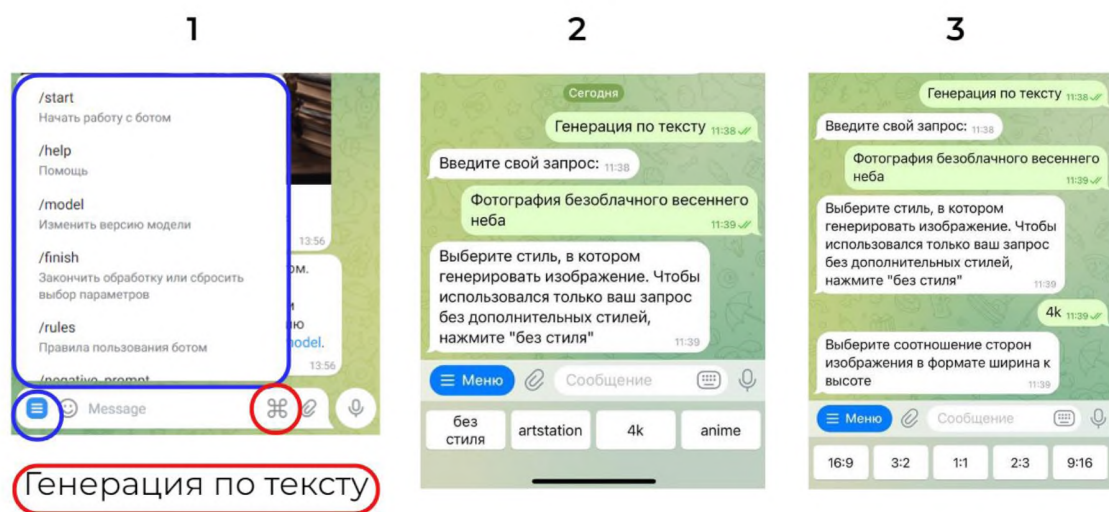
Существует значительное количество нейросетей с бесплатным доступом или имеющих бесплатные лимиты, которые генерируют изображения очень хорошего качества.

В России существует несколько таких нейросетей, например, Kandinsky от Сбербанка, который обучили на 172 миллионах пар "текст — изображение". В модели 3,3 миллиарда параметров. Нейросеть в 2023 году стала лидером в мире по темпу роста аудитории разработчиков и вошла в топ среди генеративных open source-моделей. Сегодня Kandinsky понимает запросы почти на 100 языках и помогает реализовать практически любую творческую задумку: от реалистичной картинки до видео.

Пользоваться нейросетью можно через Telegram, сайты fusionbrain.ai, rudalle.ru, в Android-приложении Салют или на умном устройстве под управлением Салют ТВ.

При использовании Kandinsky в Telegram необходимо:

1. Выбрать генерацию по тексту.
2. Написать промпт на русском или английском языке.
3. Выбрать стиль из предложенных.
4. Выбрать соотношение сторон генерируемого изображения.
5. Ждать результатов.



У написания промптов есть свои особенности, которые могут меняться в зависимости от нейросети, но есть ряд общих рекомендаций и особенностей. Среди них:

Хороший промпт = формат картинки + объект + действие + стиль

Формат картинки - то, что вы хотите получить: фотографию, постер, открытку, мозаику, вышивку

Стиль - можно написать в стиле Пабло Пикасо, аниме, 4k

Первые элементы промпта влияют на результат сильнее всего

Плохое действие

Хорошее действие

Учится

Сидит за столом с книгами

Путешествует

Стоит у памятника с рюкзаком

Участвует в олимпиаде

Перепрыгивает через шест на олимпийском стадионе

Влияние формулировки формата картинке в промпте на результатю

Формат картинки



Открытка, две белки сидят на ветке дерева в причудливом мистическом лесу, пастельные тона, общий план



Фотография, две белки сидят на ветке дерева в причудливом мистическом лесу, пастельные тона, общий план

Первые элементы промпта влияют на результат сильнее

всего



Промпт: трафаретная печать Роя Лихтенштейна. Группа собак, загорающих и наслаждающихся солнечным днём на пляже. Вода глубокого, ярко-синего цвета, а песок теплый и мягкий под ногами. Небо чистое и светит солнце. Волны нежные, а атмосфера беззаботная и расслабленная



Промпт: Группа собак, загорающих и наслаждающихся солнечным днём на пляже. Вода глубокого, ярко-синего цвета, а песок теплый и мягкий под ногами. Небо чистое и светит солнце. Волны нежные, а атмосфера беззаботная и расслабленная. Трафаретная печать Роя Лихтенштейна

Как научиться писать промпты:

1. Экспериментировать.
2. Изучать промпты, по которым генерируют изображения другие пользователи, например, в нейросети Leonardo.ai.
3. Генерировать промпты нейросетью, которая анализирует изображения и подбирает наиболее подходящие промпты для загруженного изображения (www.krea.ai).

4. Пройти обучающую игру от Google "Say what you see".

Запомните, что даже несовершенные промпты могут привести к созданию качественных изображений, особенно если нет четкого технического задания.

Реалистичность и детальная проработка изображений, созданных нейросетью, может ввести людей в заблуждение.

Примеры:

Одной из победительниц крупнейшего конкурса фотографии Sony World Photography Awards в 2023 стала работа под названием "Pseudomnesia: The Electricia", которая победила в категории "Креатив". Премия описала фото как "запоминающийся черно-белый портрет двух женщин разных

поколений". Снимок напомнил судьям семейные портреты 1940-х годов. Автор признался, что изображение было сгенерировано нейросетью.



На данный момент количество научных публикаций, упоминающих технологии с 2016 года, выросло в 381 раз. Журнал Nature включил технологии по созданию и обнаружению deepfake в число 7-ми прорывных технологий, за которыми стоит следить в 2024 году.

В настоящее время изменить видео с помощью нейросетей или сгенерировать новое достаточно просто. Для этого не нужно уметь программировать, использовать сложное программное обеспечение или иметь дорогое оборудование. Достаточно наличия телефона и мобильного интернета.

На сегодняшний день видеозвонок, аудиосообщение или видеозапись с человеком не могут быть гарантией того, что сам человек это сказал и записал. Видео или аудиосообщение больше не могут однозначно идентифицировать личность.

Крупные компании, обучающие нейросети для создания изображений или работы с видео, разрабатывают технологии, способные помечать сгенерированные изображения невидимыми водяными знаками, которые не

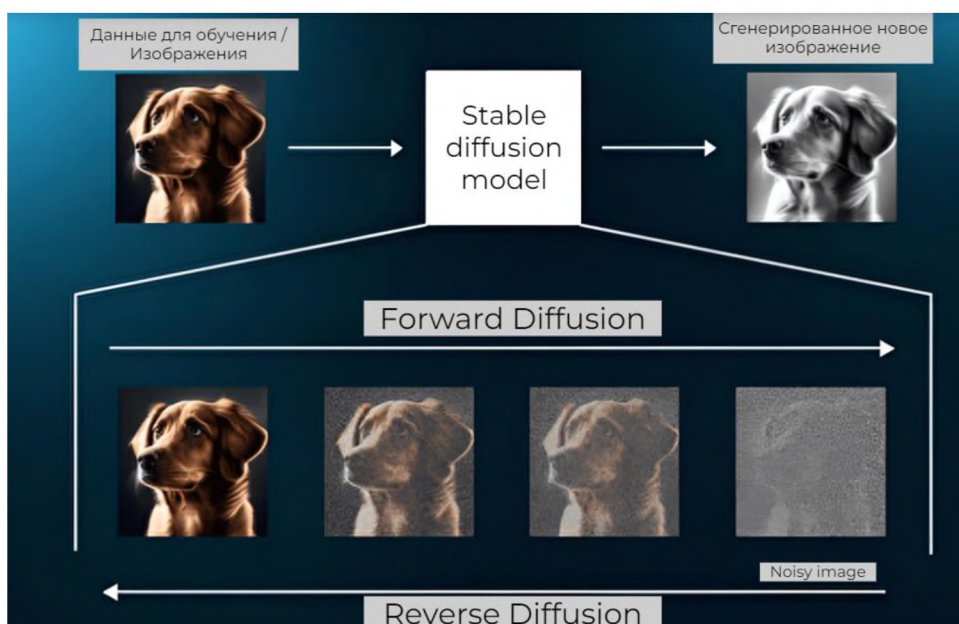
видны человеку, но позволяют программам - детекторам идентифицировать их.

OpenAI объявили, что начали пометать все изображения, созданные в ChatGPT и их API. Meta также объявила, что будет пометать изображения, созданные искусственным интеллектом от партнеров отрасли, которые публикуются на Facebook, Instagram и Threads. Контент от Google, OpenAI, Microsoft, Adobe, Midjourney и Shutterstock будет помечен водяными знаками, указывающими на то, что он был создан с использованием ИИ.

В Китае эта технология уже работает и используется на законодательном уровне. Однако пока не ведутся автоматическая маркировка и проверка всех загружаемых изображений.

Как нейросети генерируют изображения

Многие современные генеративные модели основаны на диффузионных моделях. Диффузионные вероятностные модели - это класс генеративных моделей, идея которых взята из неравновесной термодинамики. Они основаны на цепи Маркова, которая медленно добавляет шум к данным, а затем обращает процесс диффузии, чтобы получить исходные данные. Есть два основных этапа: forward process, reverse process.



Таким образом, для работы с текстом и для генерации изображений нейросети используют разные математические модели и методы работы. Из этого следует, что нейросеть, способная обрабатывать текст, и нейросеть, способная генерировать изображения, являются разными системами. Существует также специфика в формулировании промптов, например, если требуется избежать определенного элемента на изображении.

Как отличить сгенерированные нейросетью изображения

Отличить фотореалистичные генерации от реальных фотографий становится все сложнее. Ниже приведены базовые рекомендации от Google, Microsoft и экспертного сообщества о том, как не принять сгенерированное изображение за реальную фотографию:

На нейросетевую генерацию могут указывать:

1. Абстрактный или размытый фон.
2. Нереалистичная симметрия или асимметрия объектов.
3. Логотипы и надписи, которые невозможно прочитать.
4. Нереалистичный блеск или тени.
5. Кожа с признаками нереалистичности.
6. Неадекватное количество зубов или пальцев.
7. Нереалистичные мелкие детали.
8. Нарушение законов физики на изображении.

Стоит учитывать, что нейросети могут генерировать читаемые надписи, а также учитывать, что сгенерированное изображение может быть дополнительно обработано, например, в программе Photoshop.

Программы, способствующие идентификации сгенерированных изображений:

На сегодняшний день не существует абсолютно точных и надежных бесплатных онлайн-сервисов для проверки подлинности фотографий.

1. TinEye (<https://tineye.com>) покажет, где и как изображение использовалось в сети.

2. Google Images (<https://images.google.com>) или Яндекс. Распознавание (<https://yandex.ru/images>) позволяют выполнить поиск по изображению и найти источники и похожие изображения в интернете. Это может помочь увидеть, была ли ваша фотография использована ранее или имеет ли дополнительные метаданные.

3. AI or not (<https://www.aiornot.com>) определяет авторство с помощью "артефактов, невидимых человеческому глазу" - вариаций яркости и цвета. Точность инструмента составляет до 88,9% (при обновлении, точность падает). Стартап Optic был основан в марте 2022 года в Сан-Франциско разработчиками и инвесторами Андреем Дороничевым, Романом Дорониным и Владом Виноградовым.

Дополнительные действия

- Анализ метаданных: изучите данные о камере, разрешении, дате и др.
- Сравнение с известными нейросетями.
- Анализ физических характеристик: обратите внимание на физические особенности изображения, такие как разрешение, шум, гамма-коррекция и др.
- Проверка источника изображения. Если изображение поступило от ненадежного источника, есть больше шансов, что оно может быть сгенерированным.

Все эти методы могут быть неполными и не всегда достоверными. Лучшим подходом будет комбинирование нескольких методов и подробный анализ изображения, чтобы сделать осознанный вывод о его происхождении.

Создание видео с помощью нейросетей

Нейронные сети играют ключевую роль в анализе и обработке видеоданных, позволяя автоматически распознавать объекты, классифицировать содержимое, создавать видеоэффекты и многое другое. Применение нейросетей к видео сегодня обычно охватывает несколько

областей, таких как распознавание объектов, сегментация видео, создание deepfakes и автоматическое аннотирование видеоматериалов.

Одним из наиболее широко применяемых подходов является использование сверточных нейронных сетей (CNN) для анализа кадров видео. CNN способны выделять важные признаки из изображений и видео, что делает их идеальным инструментом для распознавания объектов и сцен на видеозаписях. Например, при обучении на больших наборах размеченных данных, CNN могут автоматически распознавать людей, автомобили, животных и другие объекты в видео.

Распознавание и классификация действий в видео также является важной задачей, которую можно решить с помощью рекуррентных нейронных сетей (RNN) или их модификаций, как например, Long Short-Term Memory (LSTM). RNN и LSTM могут улавливать последовательность событий во времени, что помогает определять действия и предсказывать следующие шаги в видеорядах, например, в области видеонаблюдения.

Для сегментации видео на сегменты и объекты применяются как методы, сочетающие сверточные слои для выделения признаков, так и рекуррентные слои для анализа последовательности. Это позволяет создавать более точные карты сегментации видео, выделяя объекты и их перемещения на представленных кадрах.

Создание deepfakes, то есть синтеза видео и аудио с помощью искусственного интеллекта, также осуществляется при помощи нейронных сетей, в частности, генеративных состязательных сетей (GAN). GAN позволяют переносить стили, изменять контекст и создавать реалистичные видео, в которых лица могут быть заменены или анимированы на основе данных обучающего набора.

Автоматическое аннотирование видео - еще одна важная задача, которую нейронные сети могут эффективно решать. Путем обучения модели на размеченных данных нейросеть может автоматически распознавать и

описывать содержимое видео, делая его более доступным для поиска и анализа.

Помимо этих примеров, нейросети также применяются в дополнительных областях обработки видео, таких как оптическое распознавание символов (OCR) в видео, извлечение ключевых кадров, анализ эмоций и даже автоматическое создание видеоконтента.

Важно отметить, что применение нейросетей к видеоданным требует большого объема вычислительных ресурсов и больших наборов данных для обучения. Однако с развитием технологий и увеличением доступности вычислительных ресурсов нейросети становятся все более эффективными в анализе и обработке видеоданных.

Создание видео с использованием нейронных сетей представляет собой увлекательную область исследований в области искусственного интеллекта и компьютерного зрения. Нейросети позволяют генерировать видеоряды, модифицировать и улучшать последовательности кадров, создавать новые видеоэффекты и реалистичные анимации. Рассмотрим процесс создания видео с помощью нейронных сетей и различные методы, которые применяются в этом процессе.

Одним из основных подходов к созданию видео с использованием нейронных сетей является на основе генеративно-сопоставительных сетей (GAN). GAN состоит из двух моделей: генератора и дискриминатора. Генератор создает новые видеофрагменты, а дискриминатор оценивает, насколько эти фрагменты выглядят правдоподобно. В ходе тренировки обе модели улучшают свои навыки, что приводит к генерации более реалистичных видео.

Для создания видео с помощью GAN используются различные архитектуры нейронных сетей, такие как видео-GAN, генеративно-сопоставительные сети с условием (сGAN) и другие модификации. Эти модели способны генерировать новые видеоряды, анимированные переходы между кадрами, спецэффекты и многое другое.

Еще одним распространенным методом создания видео с использованием нейронных сетей является использование рекуррентных нейронных сетей (RNN) или их модификаций, таких как LSTM (Long Short-Term Memory). RNN и LSTM особенно эффективны для анализа последовательностей данных, что делает их идеальными для создания анимаций и анализа временных шаблонов в видео.

Другим методом, который используется для создания видео, является передача обучения (transfer learning). Этот подход заключается в том, что модель, обученная на огромном объеме данных, затем дообучается на специфических видеоданных для создания новых видеоэффектов и анимаций.

Важным аспектом при создании видео с использованием нейронных сетей является обучающий набор данных. Чем более разнообразные и большие данные используются для обучения модели, тем лучше качество и разнообразие созданного видео. Размеченные наборы данных являются важным инструментом для обучения нейронных сетей на видеоданных.

Для создания различных видеоэффектов и анимаций с использованием нейронных сетей применяются различные техники, такие как inpainting (восстановление изображения), super-resolution (увеличение разрешения), style transfer (перенос стиля), image colorization (окрашивание изображения) и другие.

Благодаря нейросетям создаются не только новые видеоэффекты, но и уникальные анимации, ролики, видеоконтент для рекламы, игр и других приложений. Например, нейросети могут создавать анимационные персонажи, изменять погоду на видеозаписи, улучшать качество видео, синтезировать реалистичные анимации лиц и многое другое.

Создайте видео, воспользовавшись нейросетью в Telegram
[@video_kandinsky_bot](https://t.me/video_kandinsky_bot)

Видео могут содержать аналогичные признаки генерации, а также некоторые дополнительные признаки:

- Неестественная мимика.
- Несоответствие интонации смыслу сказанного.
- Проглатывание окончаний.
- Резкие движения.
- Смещение яркости от кадра к кадру.
- Некачественная синхронизация губ с речью.
- Низкое качество изображения.
- Безжизненный взгляд.
- Различия в резкости изображения.
- Отличие цвета лица от цвета шеи и рук.
- Различие в тенях и освещенности.
- Нечеткие границы наложенного элемента.
- Разная детализация фрагментов.
- Неестественная структура кожи и волос.
- Отличия в телосложении или причёске.
- "Мультипликационность".
- Нарушение законов физики.

В настоящий момент видеоклипы и реклама уже созданы с помощью покадровой генерации нейросетями. Например, ASML создали рекламу, полностью созданную искусственным интеллектом, используя Midjourney и Runway. Было выполнено 1963 запроса в Midjourney, на основе которых было получено 7852 изображения. Эти изображения обработали через RunwayAI, что привело к созданию 25 957 кадров в видеоролике. Возможно создание изображения с последующим анимированием его частей или генерацией "движения камеры по кадру".

OpenAI 16.02.2024 представили модель SORA - text2video, которая обладает рядом особенностей:

- Видео длительностью до минуты.
- Высокое качество изображения.

- Сложные сцены с несколькими персонажами и объектами.
- Взаимодействие персонажей между собой.
- Свободное перемещение "камеры".
- Постоянство объектов при выходе из поля зрения "камеры".
- Симуляция цифровых миров.
- Создание плавных переходов между входными видео с разными сюжетами.

Недостатки модели:

Модель неточно моделирует физику ряда базовых действий - например, разбитие стекла. Sora иногда допускает ошибки - персонаж откусывает печенье, но на нем не видны следы от укуса. А также Sora путает лево и право.

Анонс модели сопровождается технологическим отчетом. В OpenAI рассказали, что при разработке модели вдохновлялись работой над LLM. При стандартном подходе модели обучают на коротких видеороликах с уменьшенным разрешением. Sora же обучалась на видеороликах с собственными пропорциями сторон, что отразилось на результате. Также в OpenAI сначала обучили отдельный модуль внутри модели пониманию естественного языка и снабдили все видеоролики для обучения подписями. Интересно, что некоторые навыки модель смогла приобрести самостоятельно.

Sora находится на стадии тестирования и доработки (19.02.2024). Доступ к модели будет предоставлен специалистам OpenAI, а также политикам, художникам и преподавателям со всего мира для анализа реального применения модели. Екатерина Пинчук, директор продуктов GigaChat B2C, заявила: "Это еще один переворот в технологической отрасли, открывающий неограниченные возможности для создания пользовательского генерируемого контента и потенциально приводящий к изменениям в сфере видеоконтента. Также важно понимать, что за этим скрывается целая

индустрия обеспечения безопасности распространения такого видеоконтента".

Современные технологии искусственного интеллекта позволяют машинам выполнять множество задач, ранее доступных только человеку, включая создание музыки. Несмотря на сомнения некоторых людей в способности машин создавать искусство, нейросети проявляют способность создавать музыку, неотличимую от творчества человека. В данной статье мы рассмотрим лучшие нейросети для работы с музыкой и зададимся вопросом: сможет ли машина достичь уровня музыкальных гениев, таких как Вивальди и Джими Хендрикс.

Создание музыки с помощью нейросетей

Нейросети, используемые для создания музыки, являются частью широкого исследования в области искусственного интеллекта и машинного обучения. Эти инновационные системы способны анализировать, интерпретировать и генерировать звуковые структуры, создавая мелодии и композиции, которые звучат как произведения профессиональных музыкантов. Рассмотрим более подробно, как нейросети работают и какие техники используются для создания музыки.

Процесс создания музыки с использованием нейросетей начинается с обучения модели на большом объеме музыкальных данных. Эти данные могут содержать аудиофайлы различных жанров, инструментальные композиции, вокальные произведения и другие звуковые элементы. На этапе обучения нейронная сеть анализирует и изучает закономерности и характеристики входных данных, строящие абстрактные представления звуковых структур.

В процессе создания музыки нейросеть может использовать различные подходы, включая генеративные адверсариальные сети (GAN), рекуррентные нейронные сети (RNN), сверточные нейронные сети (CNN) и другие. Например, генеративные адверсариальные сети могут использоваться для синтеза новых звуковых образцов, путем обучения двух нейросетевых

моделей: генератора, который создает звук, и дискриминатора, который оценивает созданный звук на подлинность.

Одним из ключевых аспектов при создании музыки с помощью нейросетей является выбор подходящего алгоритма и обучение модели на правильном наборе данных. Чем более разнообразны данные, на которых обучается нейросеть, тем более разнообразной и качественной будет создаваемая музыка.

Кроме обучения на существующих данных, нейросети также могут использовать методы передачи обучения (transfer learning), когда модель обучается сначала на крупных наборах данных, а затем можно дообучать ее на более специфичных данных в задаче создания музыки. Этот подход позволяет улучшить качество создаваемых композиций и адаптировать модель к конкретным требованиям.

Важным аспектом является работа с выходными данными нейросети, которые могут представляться в виде аудиофайлов или нотных записей. После процесса генерации музыкальной композиции необходимо провести анализ и оценку результатов, чтобы удостовериться в качестве и соответствии требованиям.

Одним из примеров технологий, используемых для создания музыки с помощью нейросетей, является WaveNet, разработанный компанией DeepMind. WaveNet использует глубокие сверточные нейронные сети для генерации звуков, создавая более реалистичные аудиофайлы с высоким качеством звука.

Другим примером является Magenta, проект исследований Google в области музыкального искусства и машинного обучения. Magenta работает над созданием инструментов и технологий для исследования и создания музыки с использованием нейросетей. Проект разрабатывает нейронные модели, способные генерировать мелодии, гармонии и ритмы на основе обучения на больших массивах музыкальных данных.

Процесс создания музыкальных произведений аналогичен творчеству художников. Обученные на творчестве группы Beatles алгоритмы способны создавать музыку в их стиле. Каждый новый трек, изученный алгоритмом, расширяет его память и базу данных, что помогает создавать новые композиции или исследовать комбинации с другими треками. Подобные принципы использовались при обучении нейросетей созданию изображений, анализируя творчество художников.

Ниже приведен список пяти сервисов для создания музыки, основанных на использовании искусственного интеллекта.

1. AIVA: AIVA - один из ярких представителей музыкального искусства с применением искусственного интеллекта (ИИ) для создания музыки в рекламе, видеоиграх, фильмах и других медиа. Основанная в 2016 году, AIVA уже выпустила собственный альбом и создала музыку для видеоигр. Нейросеть позволяет пользователям создавать собственные музыкальные композиции и вариации уже существующих треков, облегчая процесс лицензирования. С выбранным стилем, AIVA упрощает создание гармонии в различных жанрах и стилях, помогая внести свежий вклад в музыкальную индустрию.

2. MuseNet: MuseNet - музыкальный генератор, разработанный компанией Open AI, известной своей нейросетью ChatGPT для генерации текстов. MuseNet способен создавать композиции в различных стилях и жанрах, а также имитировать стили известных композиторов, включая Моцарта. Используя глубокие нейронные сети и данные из интернета, MuseNet обладает уникальным свойством. Несмотря на наличие обширных сэмплов на Soundcloud, в настоящее время не предоставляет возможность создания собственных композиций.

3. Soundraw: Soundraw - еще один значимый сервис, помогающий изменять и создавать музыку с помощью фраз, созданных ИИ.

Объединяя возможности искусственного и ручного творчества, Soundraw упрощает процесс создания и редактирования музыки. Опции настройки и подписка для неограниченных загрузок облегчают использование сервиса.

4. Soundful: Soundful использует возможности ИИ для создания бесплатной фоновой музыки для видеозаписей, подкастов и других проектов. Процесс создания упрощен и интуитивно понятен, предлагая уникальные мелодии, не повторяющие существующие треки.

5. Imaginary Soundscape: Imaginary Soundscape позволяет превратить изображения в звуковые пейзажи с помощью нейросетей. Распознав объекты на изображениях, сервис автоматически добавляет соответствующие звуковые эффекты, создавая уникальный аудиовизуальный опыт.

Исследования и применение искусственного интеллекта в музыкальной индустрии предлагают увлекательные возможности для творчества и инноваций, продвигая музыкальное искусство в новые высоты.

Практическая работа №4

Разработка и реализация проектов, использующих искусственный интеллект для оптимизации процессов, улучшения качества принимаемых решений и повышения эффективности

Цель работы — изучить на примере корпорация Майкрософт, как бизнес может внедрять ИИ в инструменты, продукты и службы, которые могут использоваться организациями.

ИИ может быть внедрен в каждую отрасль и каждый бизнес. За последнее десятилетие ИИ позволил компаниям всех размеров достичь лучших бизнес-результатов. Существует уже основной бизнес-использование ИИ благодаря этим трем тенденциям:

1. Доступ к огромным объемам данных.
2. Доступ к огромным вычислительным ресурсам через облако.
3. Доступ к алгоритмам ИИ.

Тем не менее, ИИ в настоящее время испытывает серьезные прорывы. Новое поколение LLM позволяет создавать новые варианты использования, которые не были возможны несколько лет назад, например те, которые основаны на высококачественном генерируемом ИИ. На основе этих технологий организации будут испытывать вторую волну преобразования СИ. Тем не менее, предприятиям нужен простой способ доступа к последнему ИИ, если они хотят воспользоваться этим.

Корпорация Майкрософт работает над демократией использования ИИ. Для этого она разработала широкий спектр решений и служб для привлечения ИИ для всех, независимо от их уровня опыта ИИ. Существует четыре подхода, которые отличаются от требуемого уровня ИИ и опыта написания кода.

ИИ как copilot для бизнес-пользователей: Корпорация Майкрософт внедряет ИИ в повседневные приложения, поэтому бизнес-пользователи могут воспользоваться им, даже если они не знают ничего о кодировании или обработке данных. В этом подходе ИИ поставляется как программное обеспечение как услуга (SaaS) и становится прозрачным, то есть он полностью интегрирован в предоставленную службу без необходимости беспокоиться о нем. Например, Microsoft 365 Copilot включает в себя новейший генерированный ИИ в форме виртуального помощника, выполняющего задачи для вас в приложениях Microsoft 365.

Microsoft Power Platform: он охватывает несколько продуктов с низким кодом, которые помогают создавать различные части приложений. Эти продукты имеют слой ИИ, но он прозрачный, и вы можете воспользоваться им без непосредственной обработки.

Службы искусственного интеллекта Azure. Это решения для пользователей, которые хотят предоставить проект ИИ, но имеют мало знаний по обработке и анализу данных. Они предлагают готовые модели ИИ для повторного использования или настройки.

Машинное обучение Azure. Все задачи машинного обучения можно обрабатывать с этой платформы. Она помогает командам по обработке и анализу данных настраивать, автоматизировать и включать рекомендации по машинному обучению.

Помните, что корпорация Майкрософт разработала все эти продукты и службы, следуя строгим принципам ответственного ИИ. Любая реализация ИИ должна быть одинаково уважаемой.

Остальная часть этого модуля охватывает каждый из этих параметров. Далее мы обсудим один из самых низких барьеров входа, ИИ как copilot, внедренных в повседневные приложения.

Чтобы по-настоящему реализовать потенциал искусственного интеллекта, необходимо предоставить доступ к ИИ каждому сотруднику, чтобы он мог осмысленно использовать его в своей работе. Корпорация

Майкрософт делает это возможным путем внедрения ИИ в приложения, которые люди используют в повседневной рутине. Никакого опыта в области обработки и анализа данных не требуется, так как искусственный интеллект поставляется как только другая функция продукта SaaS. Результатом является широкий спектр интеллектуальных приложений для бизнес-пользователей.

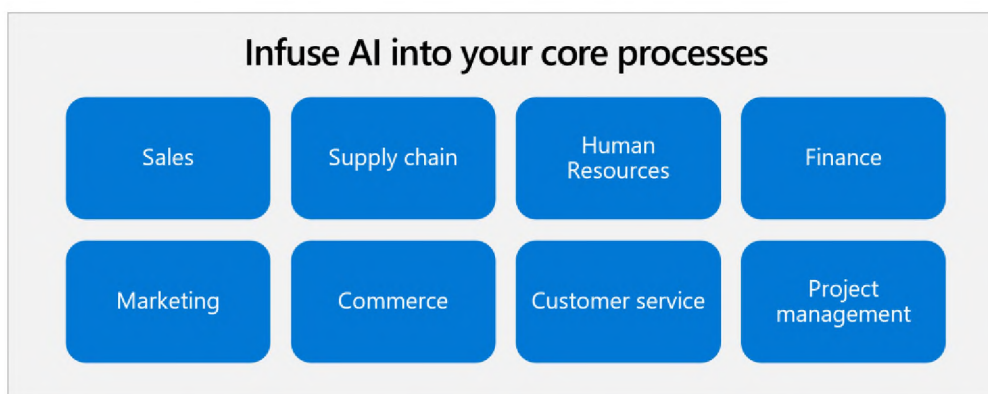
В этом уроке мы обсудим, как эти интеллектуальные бизнес-приложения открывают соответствующие возможности искусственного интеллекта в существующие рабочие процессы. Dynamics 365 помогает сотрудникам из конкретных бизнес-линий и функций автоматизировать и улучшать определенные задачи. Microsoft 365 делает то же самое, обращаясь к более широкой аудитории. Давайте рассмотрим некоторые примеры этих приложений, которые могут помочь любому пользователю использовать ИИ для эффективной работы.

Бизнес-функции

Некоторые решения искусственного интеллекта специально предназначены для решения проблем и получения аналитических сведений о некоторых конкретных горизонтальных функциях и секторах.

Эти решения часто предоставляются в виде решений SaaS AI, которые обеспечивают быстрые и экономичные результаты. Благодаря мощной аналитике в существующих рабочих процессах бизнес-пользователи могут быть более упреждающим и эффективными в своих основных компетенциях.

Microsoft Dynamics 365 разработан с этой целью: повышение производительности в определенных ключевых бизнес-функциях и секторах.



Ниже приведены некоторые примеры мощных сценариев, в которых ИИ уже доказал, полезные эффекты:

- Продажи:** продавцы могут продавать более умные данные с помощью внедренных аналитических сведений с использованием искусственного интеллекта, подпитываемых данными клиентов.

- Цепочка поставок:** бизнес-пользователи могут использовать ИИ для прогнозного обслуживания в фабриках. ИИ также полезно оптимизировать инвентаризацию.

- Кадровые ресурсы:** данные рабочей силы можно преобразовать в полезные аналитические сведения и рекомендации по дальнейшим действиям. ИИ также можно использовать для автоматизации задач кадров для сотрудников, что делает процедуры более гибкими.

- Финансы:** Аналитики предоставляют ряд средств, работающих на основе ИИ, для создания отчетов в режиме реального времени, встроенной аналитики и аналитических сведений. Например, ИИ может предсказать, когда или будут ли их клиенты платить свои счета.

- Маркетинг:** аналитика клиентов с поддержкой искусственного интеллекта дает маркетинговым пользователям единое представление о своих клиентах для оптимизации взаимодействия и обнаружения аналитических сведений, которые обеспечивают персонализированные и значимые возможности.

- Коммерция:** пользователи коммерческой торговли могут использовать аналитику ИИ, чтобы помочь им более эффективно управлять денежными потоками с помощью рекомендаций по оплате,

интеллектуальных предложений бюджета и прогнозирования денежных потоков. Они могут даже использовать ИИ для защиты электронной коммерции и клиентов от мошенничества.

•**Служба клиентов:** пользователи службы клиентов могут получить аналитические сведения для решения увеличения объема и управления эффективным распределением агентов. Они также могут создавать виртуальные агенты, которые быстро определяют и устраняют проблемы клиентов — и все это без написания кода.

•**Управление проектами.** Внедренная аналитика может предоставлять аналитические сведения на основе продаж проектов и финансовых данных. Решение предлагает планирование на основе искусственного интеллекта для прогнозирования потребностей. Пользователи операций получают аналитические сведения о том, как их клиенты используют свои продукты и услуги.

Благодаря бизнес-приложениям на основе ИИ пользователи могут объединять связи, процессы и данные между приложениями, чтобы повысить видимость и контроль.

ИИ в повседневной работе

Существует также множество возможностей искусственного интеллекта, которые уже включены в приложения, которые все используют в повседневной рутине, так как они интегрированы почти в каждое задание и функцию. Любой пользователь может использовать их для решения реалий современного рабочего места, таких как виртуальная связь и подавляющее количество информации.

В течение многих лет корпорация Майкрософт помещает искусственный интеллект для работы в **приложениях Microsoft 365**, которые каждый день используют такие приложения, как Microsoft Teams, Outlook и Office. С помощью этих интеллектуальных средств повышения производительности сотрудники могут работать вместе и проводить встречи

более эффективно, сосредоточиться на полезных задачах, а также получать своевременную аналитику для улучшения работы.

Microsoft 365 Copilot добавляет еще один слой ИИ. Бизнес-пользователи могут попросить эту виртуальную помощник выполнять определенные задачи только с помощью естественного языка. В помощник используется последняя технология создания ИИ, модели GPT OpenAI, чтобы понять запрос и сделать то, что запрашивается.

Эти решения могут улучшить рутину, повышая эффективность удаленной работы, фокус, производительность и возможности поиска.

ИИ для повседневной удаленной работы

Виртуальные собрания становятся все более критически важными в большинстве наших жизней. Хотя совместную работу лицом к лицу не заменишь, новые средства ИИ могут решить множество проблем, сблизить людей и заинтересовать сотрудников виртуальной работой.

Например, интеллектуальные возможности в Microsoft Teams, такие как размытие фона и пользовательский фон, помогают участникам не отвлекаться на помехи на экране. Живые подписи помогают улучшить специальные возможности для участников собрания, которые трудно слышать или имеют потерю слуха, не собственные английские ораторы или люди с спящим ребенком рядом. Бизнес-пользователи могут даже использовать подавление шумов в реальном времени, чтобы уменьшить число отвлекающих факторов, таких как стучание по клавиатуре или лай собак.

Когда вы не говорите лично, некоторые нюансы отсутствуют и недоразумения могут возникнуть. Copilot может помочь бизнес-пользователям найти правильный тон для своих электронных писем в Outlook, чтобы помочь решить такие проблемы.

ИИ для сосредоточенной повседневной работы

В настоящее время подпрограммы работников слишком часто прерываются отвлекающими, звонками и многозадачностью. ИИ также

может помочь справиться с этой проблемой и позволить сотрудникам сосредоточить свое время и внимание на том, что наиболее важно.

Примечание

68 процентов работников жалуются на отсутствие непрерывного времени фокуса во время их рабочих процедур.

Например, Microsoft 365 Copilot включает функции для фокуса, чтобы убедиться, что пользователи не забывают о важных проблемах. Например, в OneNote он определяет неотвеченные вопросы во всех существующих заметках и группирует их в одном централизованном расположении. В Teams Copilot может извлекать точки действий из беседы в режиме реального времени.

ИИ для продуктивной повседневной работы

Прорывы в технологиях искусственного интеллекта также позволили преобразовать личную производительность в приложениях, которые используются каждый день, приложения, такие как PowerPoint, Word и Excel.

Чтобы подготавливать интересные презентации, пользователи могут воспользоваться интеллектуальными рекомендациями по оформлению слайдов. Copilot также включает в себя генерированный ИИ для создания пользовательских образов для их слайдов. Пользователи могут репетировать презентацию и получать отзывы в режиме реального времени, чтобы улучшить заполнение и ограничить слова заливки или нечувствительные фразы.

Писатели могут воспользоваться интеллектуальными предложениями, чтобы не только исправить орфографию и грамматику, но и перераспределить целые предложения для большего эффекта или ясности. Copilot идет дальше и может писать сводки, идеи мозгового штурма, упорядочивать идеи в ключевые темы или полностью переписать содержимое.

ИИ для повседневного поиска

Использование информации стало ключевым для почти всего — от повышения производительности до понимания клиентов и многое другое. Однако данные часто разложены и трудно найти.

Возможности поиска на основе искусственного интеллекта, такие как Поиск (Майкрософт), могут помочь бизнес-пользователям пройти через эти данные, чтобы выявить более эффективные аналитические сведения и принять более эффективные решения на основе данных. Microsoft Search позволяет пользователям искать людей, файлы, сайты и другие данные в своих организациях и общедоступных веб-данных, и все это в продуктах Microsoft 365, которые они уже используют. Более того, результаты персонализированы для каждого пользователя, чтобы обеспечить их релевантность. Эта функция улучшена с помощью Copilot.

Примечание

62 процента сотрудников считают, что они тратят слишком много времени, пытаясь найти необходимую информацию.

Теперь пришло время расширить возможности сотрудников с помощью правильных средств искусственного интеллекта, чтобы расширить свои возможности и добиться большего.

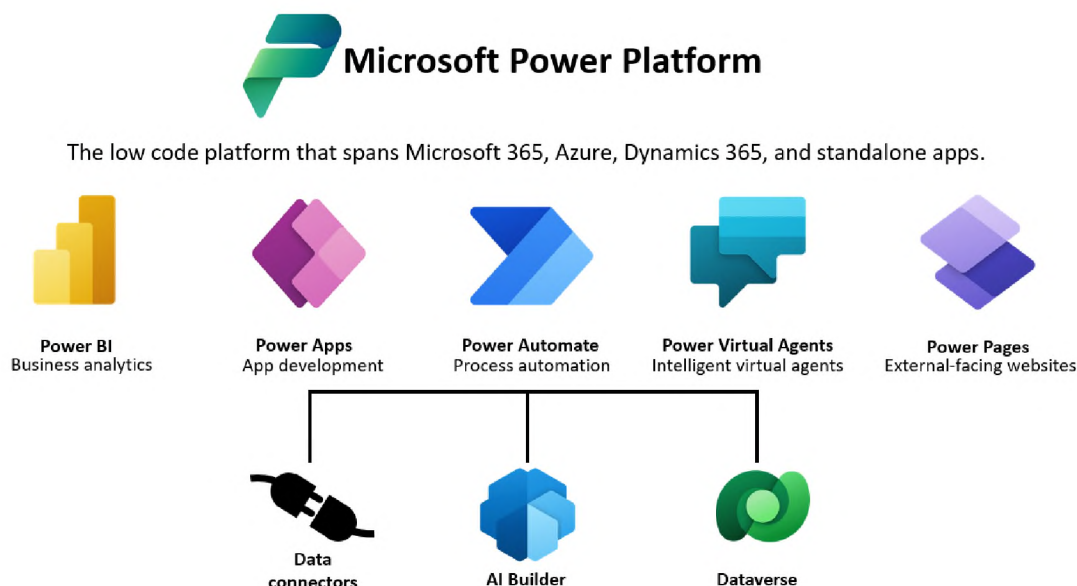
В следующем уроке мы обсудим Power Platform, который также предлагает функции искусственного интеллекта для людей без опыта обработки и анализа данных, но требуют базового уровня программирования.

ИИ, внедренный в повседневные приложения, может быть недостаточно, чтобы обеспечить возможности бизнес-приложений, необходимых организации. В таких случаях Power Platform — это следующий шаг к более настраиваемым решениям искусственного интеллекта. Он предоставляет простой, низкокодový способ внедрения ИИ в бизнес-приложения, не создавая и управляя ими самостоятельно.

Что такое Microsoft Power Platform?

Microsoft Power Platform — это набор служб с низким кодом или без кода, предназначенный для упрощения процесса создания решений. Он предоставляет стандартные блоки, помогающие командам работать быстрее. Даже если Power Platform не сосредоточена на искусственном интеллекте, его службы часто работают с помощью ИИ и помогают создавать интеллектуальные решения.

Портфель Power Platform включает пять различных продуктов: Power BI, Power Apps, Power Automate, Power Virtual Agent и Power Pages. Он также предлагает три дополнительных средства: AI Builder, Microsoft Dataverse и Подключение от. Давайте посмотрим, что каждый из них может сделать для вас.



Что можно сделать с помощью Microsoft Power Platform?

Все продукты, содержащиеся в Power Platform, используются для ускорения разработки бизнес-приложений. Помимо конкретных функций ИИ, включенных в них, они могут быть подключены к Copilot. Благодаря этой функции пользователи могут использовать сгенерированный ИИ Copilot для автоматического создания отчета, рабочего процесса, приложения, веб-сайта или чат-бота, описывая необходимые им действия.

Power BI

Power BI — это служба бизнес-аналитики. Она предоставляет аналитические сведения о настраиваемой панели мониторинга. Это помогает организациям быть более управляемыми данными и принимать лучшие решения на основе данных. Этот подход, управляемый данными, соответствует одному из основных принципов искусственного интеллекта, который подчеркивает использование данных для получения ценных аналитических сведений и улучшения выбора.

Power Apps

Power Apps — это среда разработки с низким кодом, которая позволяет предприятиям легко создавать пользовательские приложения без обширных знаний по программированию. Благодаря включению AI Builder разработчики могут легко интегрировать предварительно созданные или пользовательские модели ИИ, оптимизируя бизнес-процессы и повышая аналитику своих приложений.

Power Automate

Power Automate — это мощный инструмент, позволяющий предприятиям автоматизировать повторяющиеся задачи и оптимизировать рабочие процессы без необходимости программирования. Благодаря интеграции AI Builder пользователи могут легко внедрять предварительно созданные или пользовательские модели ИИ, что позволяет интеллектуально принимать решения и управлять эффективностью бизнес-процессов.

Power Virtual Agents

Power Virtual Agent — это средство для создания чат-ботов. Он построен на основе многих моделей ИИ, в основном тех, кто обеспечивает понимание естественного языка (NLU), поэтому бот может понять, что говорится. Однако его ИИ также может обнаруживать части бота, которые можно улучшить, и даже автоматически реализовать улучшения.

Power Pages

Power Pages — это платформа с низким кодом программного обеспечения как услуга (SaaS) для создания, размещения и управления веб-сайтами. Power Pages упрощает процесс разработки веб-сайтов, что делает его доступным даже для пользователей с ограниченным техническим опытом.

Соединители данных

Соединители данных устанавливают удобные подключения между различными компонентами (приложениями, данными, устройствами) и облаком. Эти соединители обеспечивают плавную интеграцию и взаимодействие, создавая единый интерфейс на платформе.

AI Builder

AI Builder позволяет разработчикам внедрять возможности ИИ в свои приложения и рабочие процессы, не требуя опыта обработки и анализа данных. С помощью предварительно созданных и настраиваемых моделей искусственного интеллекта AI Builder улучшает Power Apps и Power Automate, обеспечивая такие функции, как анализ тональности, классификация категорий, обнаружение сущностей, идентификация ключевых фраз и анализ языка.

Dataverse

Dataverse выступает в качестве решения для хранения данных в Power Platform, обеспечивая простую интеграцию со всеми продуктами. Он служит центральным репозиторием для данных, что позволяет эффективной организации и специальных возможностей.

Power Platform — это набор мощных инструментов, предназначенных для создания бизнес-приложений, анализа данных, автоматизации задач, создания чат-ботов и управления веб-сайтами. С помощью Power BI вы можете получить ценные аналитические сведения из данных и принять лучшие решения. Power Apps позволяет легко создавать пользовательские приложения без написания кода, а AI Builder добавляет интеллектуальные

функции, такие как анализ языка и анализ тональности. Power Automate помогает автоматизировать повторяющиеся задачи и сэкономить время, а Power Virtual Agent позволяет создавать чат-боты, которые понимают и отвечают пользователям. Кроме того, соединители данных обеспечивают плавную интеграцию между различными компонентами и Dataverse предоставляет центральное место для хранения и доступа к данным. Используя эти средства вместе, вы можете повысить производительность и повысить эффективность бизнеса.

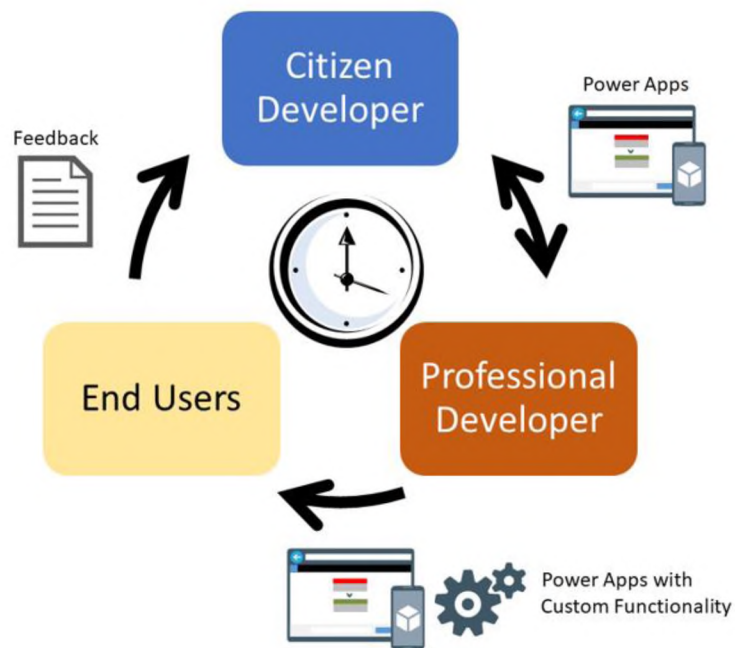
Что такое бизнес-ценность Microsoft Power Platform?

Существует два основных способа создания бизнес-ценности Power Platform для организаций:

- **Сокращение затрат на разработку.** Это предоставляет стандартные блоки для команд для создания пользовательских решений гораздо меньше времени, чем требуется при запуске с нуля. Teams может создавать пользовательские приложения всего за несколько дней или недель.

- **Включение более гибкой масштабируемой разработки:** философия низкого кода является главной для Power Platform. Это обеспечивает более быструю и гибкую разработку решений. Он позволяет разработчикам-гражданам, т. е. сотрудникам с меньшим опытом написания кода, предоставлять рабочие решения конечным пользователям. Профессиональные разработчики могут итерировать эту версию для дальнейшего улучшения. Этот подход для совместной разработки подразумевает, что решения доступны конечным пользователям на более ранней стадии и являются менее дорогостоящими. Эта структура легко обостряется путем добавления пользовательских функций.

На схеме показано, как работает этот подход к разработке слиянием.



Создание интеллектуального приложения инвентаризации с помощью Microsoft Power Platform

Предположим, что организация розничной торговли нуждается в приложении для управления их инвентаризацией. Традиционно профессиональные разработчики будут создавать это, но этот вариант займет больше времени. Благодаря Power Platform сотрудники, отвечающие за инвентаризацию, могут стать разработчиками-гражданами и создавать приложения, адаптированные к тому, что они знают, что им нужны, на основе их опыта.

Разработчики-граждане могут использовать Power Apps для создания этого прототипа приложения инвентаризации. Они могут даже объяснить на естественном языке свои потребности, и Copilot может перевести их в фактическое приложение. После создания этого прототипа профессиональные разработчики могут заполнить любые потенциальные пробелы для дальнейших функций. Затем профессиональные разработчики могут разрабатывать API для проверки инвентаризации в режиме реального времени и обновлять данные в приложении. Эти данные будут храниться и управляться в Microsoft Dataverse. С помощью сведений, предоставленных API, рабочие процессы можно создавать для автоматического обнаружения

отсутствующих и пополнения инвентаризации с помощью Power Automate. Все эти данные и операции будут отражены в отчете Power BI, поэтому проще понять, что происходит с инвентаризацией и принимать лучшие решения, обоснованные данными.

Каждое из продуктов и служб, которые мы рассмотрели, являются предварительно созданными моделями ИИ, поэтому пользователю не нужен опыт обработки и анализа данных. Если для вашего бизнеса требуются более амбициозные решения, для следующих доступных средств искусственного интеллекта требуется некоторая степень знаний искусственного интеллекта. Чтобы полностью воспользоваться ими, давайте сначала узнаем больше о том, как работают команды обработки и анализа данных и что такое типичный жизненный цикл машинного обучения.

Проблемы машинного обучения и необходимость операций машинного обучения

Недавний бум предварительно созданных моделей искусственного интеллекта, таких как те, которые предлагаются OpenAI, могут помочь организациям значительно сократить количество запугающих ресурсов, которые может потребовать проект обработки и анализа данных. Тем не менее, как больше организаций экспериментируют с искусственным интеллектом, они считают, что жизненный цикл машинного обучения сложнее, чем просто создание модели ИИ или повторное использование предварительно созданной.

Даже если в вашей организации используются готовые модели ИИ, такие как предоставляемые службами ИИ Azure, вам, вероятно, по-прежнему нужно документировать данные, код, среды модели и модели машинного обучения. Необходимо установить процессы для разработки, упаковки и развертывания моделей, а также мониторинга их производительности и иногда переобучения. Кроме того, большинство организаций одновременно

управляют несколькими моделями в рабочей среде, что дополнительно увеличивает сложность.

Для эффективной работы с этой сложностью требуются некоторые рекомендации. Они ориентированы на совместную работу между командами, автоматизацию и стандартизацию процессов, а также обеспечение того, чтобы модели можно было легко проверять, объяснять и повторно использовать. Для этого команды обработки и анализа данных используют **подход к операциям** машинного обучения. Эта методология вдохновляется DevOps (разработка и операции), отраслевым стандартом для управления операциями для цикла разработки приложений, так как борьба разработчиков и специалистов по обработке и анализу данных аналогична.

Давайте посмотрим, как выглядит типичный жизненный цикл машинного обучения, сначала если вы создаете собственную модель ИИ, а затем, если вы используете предварительно созданные модели. Специалисты по обработке и анализу данных могут управлять и выполнять рабочие процессы машинного обучения из **Машинное обучение Azure**, платформу корпорации Майкрософт, чтобы упростить операции машинного обучения. Такие средства помогают командам сотрудничать в общей, аудите и безопасной среде, где многие процессы можно оптимизировать с помощью автоматизации.

Жизненный цикл машинного обучения для создания собственной модели ИИ

Это классический подход и охватывает все обычные шаги проекта обработки и анализа данных. Во многих сценариях результирующая модель ИИ работает лучше, чем более универсальная предварительно созданная модель.

1. Подготовка набора данных. ИИ начинается с данных. Во-первых, специалистам по обработке и анализу данных необходимо подготовить данные, с помощью которых обучить модель. Это часто

является самым главным обязательством в жизненном цикле. Эта задача включает в себя поиск или создание собственного набора данных, очистку его так, чтобы оно было легко читаемым компьютером, обеспечивая его репрезентативной выборку, обнаруживая, какие переменные относятся к вашей цели, и т. д.

2.Обучение и тестирование. Затем специалисты по обработке и анализу данных применяют алгоритмы к данным для обучения модели машинного обучения. Затем они проверяют модель на новых данных, чтобы увидеть, насколько точны прогнозы, формируемые моделью.

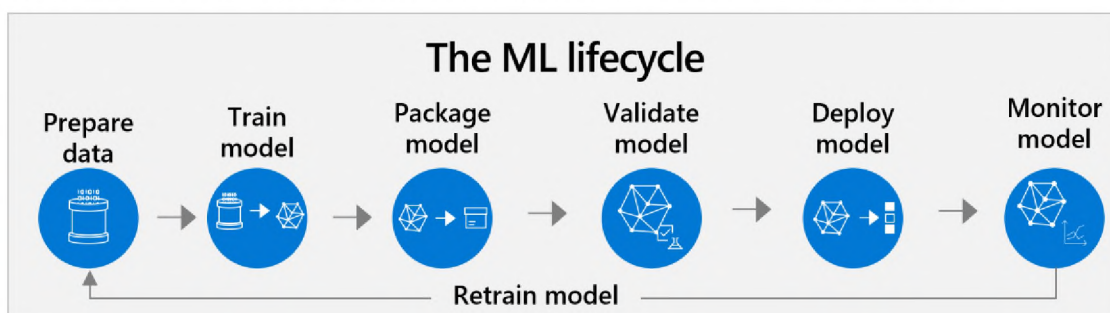
3.Пакет. Модель не может быть непосредственно помещена в приложение. Он должен быть контейнеризован, поэтому он может работать со всеми инструментами и платформами, которые использовались в его сборке.

4.Проверка. На этом этапе команда оценивает производительность модели в соответствии с бизнес-целями. Возможно, тестирование вернуло достаточно хорошие метрики, но модель может работать не так, как ожидалось при использовании в реальном бизнес-сценарии.

•Повторите шаги 1-4. На поиск удовлетворительной модели может потребоваться несколько сотен часов обучения. Группа разработчиков может обучать множество версий модели, изменяя данные обучения, подстраивая гиперпараметры алгоритма или пробуя совершенно другие алгоритмы. В идеальном случае модель улучшается в каждом цикле корректировки. В конечном итоге именно команда разработки должна определить, какая версия модели лучше всего соответствует бизнес-варианту использования.

5.Развертывание. Наконец, они развертывают модель в облаке (часто через API), на локальном сервере или на пограничных устройствах, таких как камеры, шлюзы Интернета вещей или оборудование.

6. Мониторинг и переобучение. Даже если модель хорошо работает сначала, ее необходимо постоянно отслеживать и повторно обучать, чтобы поддерживать актуальность и точность модели.



Жизненный цикл машинного обучения с помощью предварительно созданной модели

Создание собственной модели с нуля используется для использования по умолчанию. Однако бум предварительно созданных моделей изменил парадигму. Это все чаще часто используется для создания проектов обработки и анализа данных для интеграции готовых моделей и адаптации их к потребностям бизнеса. Такой подход может значительно изменить рабочий процесс.

1. Проверка. Использование предварительно созданной модели означает, что организации начинают только с проверки, если она работает для них. Это ключ, чтобы понять как можно скорее, какую предварительно созданную модель использовать и насколько жизнеспособна она.

•Повторите шаги 1-4. Специалисты по обработке и анализу данных повторяют шаги 1-4, пока результаты не будут достаточно хорошими. Часто для предварительно созданной модели может потребоваться несколько усилий, чтобы обеспечить необходимую модель.

2. Инженер запрашивает: это первый способ улучшить результаты, полученные с предварительно созданной моделью. Команда должна объяснить, что им нужно, чтобы модель понимала все нюансы. Эта задача включает в себя повторение запроса (запроса) до тех пор,

пока модель не получит ее. Может потребоваться время. Однако разработчики-граждане и даже конечные пользователи могут сделать запрос на проектирование, если руководствоваться специалистами по обработке и анализу данных. Такой подход дает больше возможностей бизнес-пользователям.

3.Настройка набора данных: если инженерия запросов недостаточно улучшает результаты, это может быть связано с тем, что предварительно созданная модель слишком универсальна для предполагаемого варианта использования. Затем специалисты по обработке и анализу данных должны дополнить модель пользовательскими данными обучения.

4.Обучение и тестирование. Специалисты по обработке и анализу данных могут добавить пользовательский уровень обучения на вершине готовых моделей с дополнительными данными шага 3. Таким образом, они получают новую версию предварительно созданной модели, адаптированную к проблеме искусственного интеллекта, которую они пытаются решить. Другим вариантом является разработка полной пользовательской модели, которая охватывает пробелы, оставленные предварительно созданной. Многие модели могут сосуществовать в решении искусственного интеллекта.

5.Пакет и развертывание. Существуют различные способы использования предварительно созданной модели. Во многих случаях его можно использовать через API. Такой подход подразумевает, что организация не владеет ею и не может ее настроить, но экономит время упаковки и развертывания. Если был некоторый уровень пользовательского обучения, команда ИИ должна упаковать и развернуть эту новую версию модели.

6.Монитор. Как и все модели, предварительно созданные также должны быть проверены регулярно хранить свои края. Специалисты по обработке и анализу данных должны обновляться на новых

предварительно созданных моделях. Новые пользовательские учебные курсы также могут потребоваться в какой-то момент.

Какой из двух подходов лучше? Это зависит от сценария развертывания. Работа с предварительно созданными моделями имеет преимущество, что требует меньше ресурсов и обеспечивает более быстрые результаты. Однако предварительно созданные модели обучены решать широкий спектр вариантов использования, поэтому они могут бороться с очень конкретными потребностями. В таких случаях полная пользовательская модель может быть лучшей идеей. Гибкий набор обоих подходов часто предпочтительнее и помогает масштабировать. Команды ИИ могут сэкономить ресурсы с помощью готовых моделей для самых простых вариантов использования, в то время как инвестировать эти ресурсы в создание пользовательских моделей ИИ для самых сложных сценариев. Дальнейшие итерации могут улучшить предварительно созданные модели, переобучение их.

Далее давайте рассмотрим преимущества и ценность бизнеса, которые могут обеспечить операции машинного обучения.

В этом уроке мы обсудим важность операций машинного обучения (MLOps). MLOps применяет методологию DevOps (разработка и операции) для более эффективного управления жизненным циклом машинного обучения. Это обеспечивает более гибкую, эффективную совместную работу в командах искусственного интеллекта среди всех заинтересованных лиц. Эти совместные работы включают специалистов по обработке и анализу данных, инженеров ИИ, разработчиков приложений и других ИТ-команд.

Существует множество доступных продуктов, помогающие командам реализовать MLOps. Корпорация Майкрософт предлагает Машинное обучение Azure, Azure DevOps и GitHub.

Процессы и инструменты MLOps помогают этим командам работать совместно и обеспечивают прозрачность благодаря использованию общей контролируемой документации. Технологии MLOps обеспечивают

возможность сохранения и отслеживания изменений всех ресурсов, таких как данные, код, модели и другие средства. Эти технологии также могут повысить эффективность и ускорить жизненный цикл с помощью автоматизации, повторяемых рабочих процессов и многократно используемых ресурсов. Все эти методики делают проекты ИИ более гибкими и эффективными.

Воспроизводимость модели

Во время начального итеративного обучения и последующего повторного обучения модели существует несколько моментов, которые могут сделать сложный процесс более управляемым. Одним из них является сохранение воспроизводимых моделей, что означает, что их можно легко запускать в одном наборе данных любым членом команды с одинаковыми или аналогичными результатами. Воспроизводимость достигается путем документирования процессов и совместного использования ресурсов.

Во-первых, рекомендуется централизованно **управлять ресурсами**, например средами, кодом, наборами данных и моделями, чтобы команды могли совместно использовать их, а также использовать их повторно.

- **Реестр** моделей: как команды экспериментирует с различными версиями модели, реестр моделей предоставляет центральное место для сохранения каждой версии. Благодаря реестру команды могут легко вернуться к предыдущей версии, если что-то не работает, даже после того, как решение начало использоваться в рабочей среде. Реестр моделей также служит в качестве следа аудита для истории каждой модели.

- **Управление кодом**: техническим лицам, принимающим решения, необходимо определить, какие технологии и процессы будут использоваться для управления кодом. Обычно к ним относятся репозитории кода, такие как GitHub, в которых код можно сохранить, разместить в системе управления версиями, а также совместно

использовать и повторно применить. Он также включает средства для использования и управления версиями кода.

- **Управление наборами данных.** Мы также рекомендуем централизованно сохранять обучающие наборы данных. Таким образом, команды могут повторно использовать их, поделиться ими с коллегами или отслеживать, как они изменяются с течением времени для управления дрейфом.



- **Общие среды:** создание сред модели, которые можно совместно использовать для отдельных пользователей. Это упрощает переход между этапами в процессе создания модели и позволяет группам совместно работать над определенными этапами.

Во-вторых, мы рекомендуем автоматизировать задачи с **конвейерами** машинного обучения. Конвейер — это рабочий процесс полных вычислительных задач, которые могут выполняться независимо. В машинном обучении конвейер может автоматизировать подготовку данных, настройку обучения, процессы обучения или проверку модели. Конвейеры экономят затраты и время для специалистов по обработке и анализу данных каждый раз, когда они должны итерировать часть жизненного цикла машинного обучения.

Проверка модели

Перед развертыванием модели важно проверить метрики производительности. У вас может быть несколько метрик, которые используются для указания "наилучшей" модели. Очень важно работать со специалистами по обработке и анализу данных, чтобы понять, какие метрики важны, и оценить их перед развертыванием. Существуют **инструменты для вычисления метрик модели**, например функция потерь или матрица ошибок.

Метрики обычно сравнивают прогнозируемую модель с прогнозируемым (истинным значением или земной правдой). В целом, фокус заключается в том, чтобы максимизировать истинные положительные и истинные отрицательные значения, то есть модель успешно прогнозирует истинные значения. Важно избежать ложных срабатываний и ложных отрицательных значений, то есть неправильных прогнозов и пропущенных прогнозов.

Важно проверить метрики производительности в случае использования бизнеса. Например, предположим, что вы разработали модель для прогнозирования состояния здоровья пациента. Как поставщик услуг здравоохранения, имеющий дело с жизненно важными ситуациями, вы, вероятно, предпочтете иметь ложноположительные результаты диагностики, а не чрезвычайно высокую скорость диагностики с низкой точностью результатов.

Если модель является более новой версией существующей модели, необходимо узнать, работает ли она лучше предыдущей на ключевых метриках.

Матрица ошибок модели машинного обучения

		Прогнозируемая метка	
		нет	да
Фактическая метка	нет	167	1
	да	0	163

Соответствие прогнозируемых и фактических значений означает высокую точность

Развертывание модели

Существует несколько вариантов развертывания модели в рабочей среде. Специалисты по обработке и анализу данных и инженеры ИИ должны работать вместе, чтобы узнать лучший вариант для каждого случая.

- Облако.** Один из вариантов — развертывание моделей с помощью облака, часто использующее интерфейс программирования приложений (API). Существуют масштабируемые средства для автоматизации и упрощения этого процесса, таких как Kubernetes или Экземпляры контейнеров Azure.

- Локальная среда.** Модели также можно развертывать непосредственно на сайте на собственных серверах организации.

- Edge:** Также можно развернуть модели на пограничных устройствах, таких как камеры, беспилотные летательные аппараты и оборудование. Этот параметр может быть полезным в сценариях Интернета вещей.

Независимо от того, где развертывается модель, рабочий процесс развертывания одинаков. Сначала необходимо зарегистрировать модель в реестре моделей. Затем необходимо подготовиться к развертыванию модели, указав ресурсы, использование и целевой объект вычислений. Наконец, вы

развертываете его в нужном расположении, тестируете его и продолжаете отслеживать метрики для конкретной модели на протяжении всего жизненного цикла.

Повторное обучение модели

Хотя процесс разработки на этом заканчивается, цикл обслуживания только начинается. Необходимо отслеживать модели и периодически выполнять повторное обучение моделей, чтобы устранить проблемы с производительностью и воспользоваться новыми данными обучения. Чтобы настроить себя на успех, необходимо создать цикл переобучения или систематический и итеративный процесс для постоянного уточнения и обеспечения точности модели.

Этот процесс может показаться слишком сложным. Помните, что его можно значительно упростить с помощью предварительно созданных моделей. Инструменты MLOps, такие как Машинное обучение Azure не обязательно должны быть заполнены пользовательскими моделями, они также принимают предварительно созданные модели. В этом смысле службы искусственного интеллекта Azure являются отличной альтернативой, так как она предлагает более быстрые результаты с меньшим количеством необходимых знаний по обработке и анализу данных.

В следующем уроке рассмотрим возможности служб искусственного интеллекта Azure.

Что такое службы ИИ Azure?

При принятии ИИ в бизнес следует сначала рассмотреть предварительно созданные службы ИИ. Службы искусственного интеллекта Azure — это продукт Майкрософт, который предоставляет ИИ как **SaaS**. Она включает предварительно обученные модели, разработанные глобальными исследователями Майкрософт и специалистами по обработке и анализу данных для решения распространенных проблем. Чтобы не изобретать

велосипед, организации могут использовать предварительно созданные службы для получения нужного качества и ускорения доставки технологических решений.

Лучше использовать службы ИИ Azure, которые предлагают предварительно созданные службы ИИ в визуальном представлении, речи, языке, поиске или принятии решений для решения распространенных сценариев. Это позволяет сделать возможности ИИ доступными для всех разработчиков и организаций, не имеющих опыта в сфере машинного обучения. В результате разработчики с любым уровнем навыков смогут легко добавить интеллектуальные функции в новые или существующие бизнес-приложения.

Существуют некоторые преимущества использования служб ИИ Azure вместо пользовательских моделей ИИ:

- **Предоставление возможностей разработчикам граждан:** службы ИИ предназначены для людей с небольшим опытом обработки и анализа данных. Например, они обычно обрабатывают большинство решений для обучения и предоставляют основные результаты тестирования после обучения. Бизнес-пользователи могут использовать службы ИИ для адаптации их бизнес-потребностей с некоторыми рекомендациями и устранением неполадок в команде по обработке и анализу данных.

- **Экономия затрат:** так как службы ИИ бессерверны, они обычно менее затратны, чем разработка и обучение модели на внутреннем сервере.

- **Гибкость развертывания.** Вы можете экспортировать модели служб ИИ и запускать их в любом месте, в облаке, локальной среде или на пограничном сервере.

Что можно сделать с помощью служб ИИ Azure?

Существует пять крупных семейств служб ИИ: речь, язык, зрение, решение и Служба Azure OpenAI. Каждая служба включает несколько моделей ИИ, которые дополняют друг друга.

•**Речь:** эти модели имеют дело с устной беседой. Они могут преобразовать речь в текст и наоборот. Кроме того, можно перевести то, что говорит докладчик и определить каждого докладчика. Модели могут даже предлагать исправления произношения для динамиков.

•**Язык:** эти службы ИИ сосредоточены на обработке и анализе текста. Они обучены понимать естественный язык и извлекать аналитические сведения. Например, модели распознают язык, намерение, сущности и тональность в тексте. Кроме того, они могут найти ответы на вопросы, положенные им.

•**Визуальное видение:** включает в себя модели, которые анализируют изображения и видео. Помимо более универсальных моделей, существуют конкретные для извлечения текста из изображений (оптическое распознавание символов или OCR) для распознавания человеческих лиц. Другим вариантом является Azure Пользовательское визуальное распознавание, что позволяет пользователям создавать собственные модели ИИ для распознавания объектов или классификации изображений. Помните, что службы распознавания лиц строго ограничены из-за политик ответственного искусственного интеллекта Майкрософт.

•**Решение.** Эти модели предоставляют рекомендации по принятию наилучших решений. Они могут вызвать оповещение, когда они обнаруживают аномалии в данных или помечают потенциально небезопасное содержимое. Существует также Персонализатор, модель рекомендаций, которая предлагает содержимое для отображения каждому пользователю на основе их поведения.

•**Служба Azure OpenAI:** Корпорация Майкрософт предлагает некоторые из последних моделей сгенерированных ИИ, разработанных OpenAI. В следующем уроке они будут обсуждаться более подробно.

Все службы искусственного интеллекта Azure предоставляют модели ИИ в виде SaaS. Если вам нужно настроить продукт и создать собственный SaaS, приложения Azure IaaS службы ИИ позволяют сделать это без увеличения сложности обработки и анализа данных. приложения Azure IaaS службы ИИ основаны на службах ИИ Azure и сосредоточены на ключевых сценариях. Например, Azure Распознаватель документов включает модели OCR и анализа текста для извлечения данных из счетов, квитанций и других документов.

Использование служб ИИ Azure в реальном варианте использования

Предположим, как можно улучшить продажи и послепродажные функции вашего бизнеса только с помощью Служб искусственного интеллекта Azure.

Во-первых, вы можете увеличить продажи, порекомендовав соответствующие продукты пользователям на веб-сайте с персонализатором на основе их данных поведения. Вы можете помочь им во время покупок с виртуальным помощником на базе Azure OpenAI Services. Это помощник может решить вопросы в естественном, гибком взаимодействии.

После завершения продажи вы можете проанализировать отзывы клиентов, публикуемых на ваших продуктах. Службы ИИ языка могут анализировать отзывы и извлекать аналитические сведения о том, какие клиенты думают о конкретных продуктах и бренде. Эта информация является ключом для улучшения того, что вы предлагаете.

Однако клиенты могут не всегда быть удовлетворены их покупкой. У вас есть центр обработки вызовов для решения этих проблем и доставки услуг после продажи. Службы искусственного интеллекта речи могут анализировать, отслеживать и оценивать эти вызовы. Вы можете узнать, о

чем клиенты жалуются больше всего, или о том, являются ли они содержимым с решениями, которые вы предлагаете им. Используя эти аналитические сведения для улучшения служб после продаж, вы можете создать лояльность клиентов.

Службы ИИ Azure включают последние модели ИИ. Не только их модели постоянно обновляются, они также охватывают современные модели OpenAI. Вы можете использовать самые новые крупные языковые модели (LLM), такие как GPT помощник или чат поиска Bing. Давайте узнаем, как в следующем уроке.

Служба Azure OpenAI является последним дополнением к Службам искусственного интеллекта Azure. Это свидетельствует о приверженности Корпорации Майкрософт предложить современные модели искусственного интеллекта таким образом, чтобы они были доступны для бизнес-пользователей.

Что такое Служба Azure OpenAI?

OpenAI является успешной компанией ИИ, которая в последнее время разработала некоторые из самых известных моделей генерированных ИИ. ChatGPT и DALL-E являются некоторыми из их творений. Эти модели доступны через общедоступный API. Как и все генерированные ИИ, модели OpenAI получают запросы или инструкции от пользователя и отправляют обратно созданное ИИ содержимое.

Служба Azure OpenAI — это продукт Майкрософт, который позволяет пользователям использовать модели OpenAI через **службы** искусственного интеллекта Azure. Другими словами, он позволяет получать доступ к моделям OpenAI непосредственно из Azure, а не к общедоступному API. Имейте в виду, что Служба Azure OpenAI не является единственным продуктом Майкрософт, предоставляющим такие модели пользователям. В предыдущих уроках мы уже обсудили создание искусственного интеллекта,

включенного в Microsoft 365 Copilot и Copilot в Power Platform. Эти функции copilot поддерживаются GPT, модель OpenAI для создания текста.

Что можно сделать с помощью Службы Azure OpenAI?

Модели, включенные в службу Azure OpenAI, ориентированы на создание искусственного интеллекта. То есть они позволяют пользователям автоматически создавать контент, который достоверно выглядит человеком. Например:

•**Язык:** модели GPT, которые понимают и создают естественный язык. Они могут извлечь смысл из запросов пользователей, включая нюансы, на уровне человека. Благодаря этой способности они могут участвовать в правдоподобной и удовлетворительной беседе. Они также хорошо суммируют, решают вопросы и пишут сообщения электронной почты, копии, код или другие документы при выполнении четких инструкций.

•**Изображение:** модели DALL-E, которые создают визуальные элементы. Они переводят запрос пользователя на изображение. Эти модели можно использовать для создания всех видов искусства: реалистичные фотографии, иллюстрации, логотипы, эскизы и т. д. DALL-E успешно создает стили при запросе.

•**Код:** модели Codex предназначены для перевода запросов пользователя в рабочий код программирования. Разработчики могут использовать их для создания фрагментов кода и ускорения их производительности. Codex предоставляет результаты на нескольких языках программирования. Это технология за GitHub Copilot.

•**Более сложные модели:** модели внедрения обеспечивают основу для специалистов по обработке и анализу данных для создания более сложных моделей ИИ. Они анализируют тексты и создают абстрактное сопоставление их смысла. Эти сопоставления помогают экспертам по

ИИ найти сходство в текстах. Модели внедрения являются ключевыми в сценариях, связанных с поиском, рекомендациями или аномалиями.

Модели OpenAI постоянно развиваются и поэтому azure OpenAI Service. Вы можете проверка, какие версии этих моделей доступны в [документации](#) по Службе OpenAI Azure.

Преимущества использования Службы Azure OpenAI вместо подключения непосредственно к API OpenAI

Многие бизнес-лидеры не уверены в разнице между потреблением моделей OpenAI непосредственно из api или с помощью Службы Azure OpenAI. Существуют четкие причины, по которым организации предпочитают службу Azure OpenAI:

- Те же модели:** учитывая определенную модель и версию, Azure обеспечивает точно то же самое, что и API. Пользователи Azure не пропускают никаких обновлений. ИИ так же хорошо.

- Конфиденциальность:** API OpenAI является общедоступным. Этот факт подразумевает, что все данные, отправленные в API, могут использоваться для улучшения моделей. Если ваш вариант использования включает частную, конфиденциальную информацию, использование API может быть не лучшим вариантом. Напротив, при использовании Службы Azure OpenAI корпорация Майкрософт размещает модели в своей инфраструктуре Azure. Данные клиента не отправляются в OpenAI для улучшения моделей, так как все остается в Azure.

- Безопасность:** Microsoft Azure — это надежная, надежная и надежная платформа. Разработчики Azure имеют широкий опыт защиты инфраструктуры для общедоступных и частных организаций во всем мире, чтобы безопасно работать в Azure. Служба Azure OpenAI основана на этом опыте. Она обеспечивает шифрование, частную сеть и систему

региональной доступности, чтобы обеспечить безопасность и прерывание бизнеса.

•**Ответственный ИИ:** этические обязательства Майкрософт определяют, как модели OpenAI предлагаются в Azure. Корпорация Майкрософт занимается обширными исследованиями и работой по выявлению, измерению, устранению и работе по устранению всех потенциальных вредов, поступающих от этого генерируемого искусственного интеллекта. Например, разработчики предотвращают неправильное использование системы дополнительных инструкций, которые переуправляют определенные запросы. Существует также система фильтрации содержимого, которая гарантирует, что модели не предоставляют потенциально вредного содержимого.

В заключение служба Azure OpenAI объединяет лучшие из OpenAI и Microsoft Azure. С одной стороны, она поставляет инновационные модели OpenAI. С другой стороны, это все преимущества размещения в надежной бизнес-среде, которая является Azure.

Использование Службы Azure OpenAI в реальном варианте использования

Предположим, как организация может разработать виртуальную помощник с помощью Службы Azure OpenAI. В случае использования последнего урока мы упоминаем эту возможность для доставки услуг после продаж.

Клиенты могут взаимодействовать с помощник, чтобы запросить информацию о своих заказах, покупках или гарантиях. Они могут поговорить с этим помощник по чату или позвонить. Предыдущие поколения виртуальных помощник, необходимых для проектирования потоков бесед, и этот подход сделал беседы жесткими в некоторой степени. Помощник, управляемый моделью GPT, не будет страдать от этого препятствия. Он не требует проектирования рабочих процессов, и он может заниматься более

гибкими, естественными беседами. Этот перк подразумевает, что помощник сможет решать вопросы и задачи без определенной подготовки, пока она предоставляет информацию. Этот подход экономит время и ресурсы от команды разработчиков.

Команда ИИ должна передавать виртуальным помощникам персональные данные и внутреннюю документацию, чтобы ответить на сомнения по политике возврата, состоянию отгрузки и т. д. Служба Azure OpenAI позволяет безопасно выполнять эту операцию без риска неправильного использования данных. Обеспечение конфиденциальности данных клиента является ключевым, особенно если бизнес имеет дело с конфиденциальными данными, такими как кредитные карты номера.

Клиенты часто не любят взаимодействие с пользователем, предоставляемое чат-ботами, потому что они воспринимают их как холодные и отдаленные. Чтобы создать более дружелюбный, более человеческий опыт, команда ИИ может создать с помощью аватаров DALL-E для виртуального помощника или даже профиля клиента. Эта функция усилит более естественный поток беседы, достигнутый моделью GPT. Как правило, клиенты получают выгоду от более теплого, более естественного после продаж опыта.

Если клиент удовлетворен в конце взаимодействия с виртуальным помощником, специалисты по обработке и анализу данных могут использовать модель Внедрения, чтобы рекомендовать им дальнейшие покупки. В этом сценарии важно не только обеспечить правильную рекомендацию. Это еще более важно, чтобы правильно читать настроение клиента: если они недовольны опытом покупок, это может быть контрпродуктивным, чтобы предположить, что они покупают снова. В этом случае помощник следует запретить предлагать рекомендацию. По этой причине это ключ к подсчету с моделями GPT, которые особенно хорошо замечают эти нюансы.

Вопросы для самоконтроля

1. Что такое искусственный интеллект и какие основные концепции лежат в его основе?

2. Какие методы и алгоритмы используются в искусственном интеллекте для решения задач?

3. Что такое автоматизированное машинное обучение и какие преимущества оно предоставляет?

4. Какие основные компоненты включает в себя Azure Machine Learning и какие возможности предоставляет для разработки моделей?

5. Какие шаги необходимо выполнить для создания модели машинного обучения в Azure Machine Learning?

6. Какие методы обработки текста применяются при работе с нейросетями для анализа и обработки текстовых данных?

7. Как нейросети могут быть использованы для классификации текстов и извлечения ключевой информации?

8. Как нейросети применяются для анализа и обработки изображений?

9. Как можно использовать нейросети для распознавания и классификации объектов на изображениях?

10. Какие методы обработки аудио и видео данных существуют при работе с нейронными сетями?

11. Какие преимущества предоставляет использование нейросетей для обработки аудио и видео данных?

12. Как разрабатывать проекты, использующие нейросети для оптимизации процессов в бизнесе?

13. Какие принципы необходимо учитывать при использовании искусственного интеллекта для улучшения качества принимаемых решений?

14. Какие методы повышения эффективности при применении искусственного интеллекта в проектах могут быть применены?

15. Как измерять качество моделей машинного обучения в рамках проектов с использованием искусственного интеллекта?
16. Какие виды данных необходимы для обучения моделей нейронных сетей при работе с текстом?
17. Как обучать модели нейросетей для работы с изображениями с использованием различных архитектур?
18. Какие техники предобработки данных могут быть применены перед обучением нейросетей для работы с аудио и видео данных?
19. Какие практические примеры проектов можно реализовать с использованием нейросетей для оптимизации процессов в медицине?
20. Как можно оценить эффективность искусственного интеллекта в улучшении качества сервисов в области финансов?
21. Как использовать искусственный интеллект для прогнозирования спроса на товары в розничной торговле?
22. Какие метрики можно применять для оценки результатов проектов с использованием нейросетей в различных областях деятельности?
23. Какие вызовы могут возникнуть при разработке проектов с использованием нейросетей и как их преодолеть?
24. Как документировать и анализировать результаты работы моделей искусственного интеллекта для последующей оптимизации?
25. Как обеспечить безопасность и конфиденциальность данных при использовании искусственного интеллекта в проектах?
26. Какие методы интерпретации результатов работы нейросетей можно использовать для объяснения принимаемых моделью решений?
27. Как проводить A/B тестирование моделей машинного обучения при оптимизации процессов в бизнесе?
28. Как обеспечить корректный мониторинг и поддержку моделей искусственного интеллекта после их внедрения в продуктивную среду?
29. Какие возможности предоставляет Azure Machine Learning для масштабирования проектов с использованием искусственного интеллекта?

30. Как оценить потенциальные экономические и социальные выгоды от внедрения проектов с использованием искусственного интеллекта в различных областях деятельности?