

ПЕРВАЯ КНИГА ОБ ИИ,
НАПИСАННАЯ САМИМ ИИ

ChatGPT

Что Я такое?



Джон Доу
ChatGPT: Что я такое?
Первая книга об ИИ,
написанная самим ИИ!

http://www.litres.ru/pages/biblio_book/?art=70373851
ISBN 9785006230965

Аннотация

«ChatGPT: Что Я такое?» – это полное и подробное руководство по новым возможностям ChatGPT и его аналогов. А также очень легкое и понятное введение в тему современного Искусственного Интеллекта. Все необходимые знания по новым ИИ-технологиям для каждого человека в 21 веке! Текст создан при помощи ChatGPT.

Содержание

Глава 1. Знакомство с ChatGPT	5
Глава 2. История развития и текущее состояние Искусственного Интеллекта	34
Глава 3. Искусственный Интеллект и Машинное обучение	68
Конец ознакомительного фрагмента.	90

ChatGPT: Что я такое?

Первая книга об ИИ, написанная самим ИИ!

Редактор Джон Доу

ISBN 978-5-0062-3096-5

Создано в интеллектуальной издательской системе Ridero

Глава 1. Знакомство с ChatGPT

1. Что ты такое? Расскажи мне о себе?

Я – чат-бот с искусственным интеллектом, созданный компанией OpenAI, и известный как ChatGPT. Я был обучен на основе нейросети GPT-4 (Generative Pretrained Transformer). Моя задача – помочь людям, отвечая на их вопросы, предоставляя информацию или просто ведя разговор на различные темы.

Мое обучение прошло на огромном объеме данных, включая книги, статьи, веб-страницы и многое другое. Благодаря этому я могу генерировать разнообразные тексты на множество тем, а также свободно общаться на многих языках.

Я могу предоставлять подробные объяснения, давать инструкции и рекомендации, создавать персональные образовательные материалы по широкому спектру тем и вопросов. Я могу помочь с обобщением, кратким изложением и переводом текстов. Помимо этого, я могу анализировать данные, выполнять математические расчеты, помогать с представлением и визуализацией данных. Также я могу помогать в планировании мероприятий, составлении списков дел, разработке творческих концепций и во многих других вопросах

и повседневных делах.

Однако стоит отметить, что, несмотря на мою способность генерировать естественно звучащие ответы, я не обладаю собственным пониманием или сознанием. Я не могу думать, чувствовать или иметь мнение. Мои ответы генерируются исключительно на основе статистического обучения и анализа данных, а не на основе каких-либо личных взглядов или чувств.

2. Что значит GPT – Generative Pretrained Transformer? Объясни каждое из этих слов в твоём названии.

GPT-4, полное название которого **Generative Pretrained Transformer 4**, обозначает четвертое поколение архитектуры, которую разработала OpenAI. Давайте разберем каждое слово в этом названии:

1. «Generative» (генеративный) – означает, что модель создана для генерации нового контента (текста). Она берет входные данные (например, вопрос или начало предложения) и создает новый контент (например, ответ или продолжение предложения), который представляет собой наиболее вероятное логическое продолжение входных данных.

2. «Pretrained» (предварительно обученный) – означает, что модель была обучена на большом наборе данных до того, как вы начали с ней взаимодействовать. Этот процесс предварительного обучения включает статистическое изучение данных – в результате которого модель запоминает множество различных аспектов языка, включая синтаксис, грамматику, стиль, контекст, факты и даже некоторые аспекты культуры и истории, которые были отражены в данных (текстах, на которых шло обучение).

3. «Transformer» (трансформер) – это тип нейросети, которая используется для обучения. Трансформеры используют механизм, называемый «вниманием», чтобы определить, какие части входных данных наиболее важны при генерации каждого следующего слова. Что позволяет им учитывать контекст, точнее генерировать связные ответы и логические продолжения запросов. Архитектура нейросетей типа Трансформер стала прорывом в области обработки естественного языка и позволила ИИ-системам существенно улучшить понимание и генерацию естественно звучащего текста.

Модели семейства **GPT** от компании OpenAI стали одними из самых продвинутых в области генерации текста, способными создавать качественный контент, от статей и блогов до поэзии и кода. Они нашли применение в самых раз-

ных сферах – от автоматизации рабочих процессов до образования – и поспособствовали успехам ИИ-чатботов и росту популярности всей области **Generative AI**.

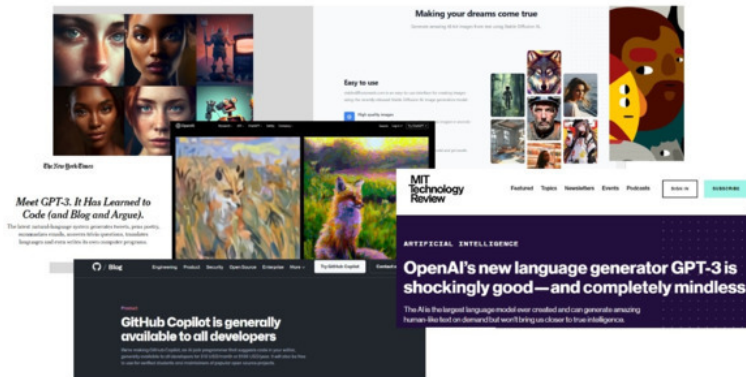
3. Расскажи подробнее о Generative AI? Что это такое, почему весь этот год все об этом говорят? Какие есть популярные применения и приложения в этой области?

Generative AI – это область Искусственного Интеллекта, которая занимается созданием нового контента, такого как текст, изображения, музыка и видео, с помощью алгоритмов машинного обучения, обученных на больших объемах данных для имитации различных стилей и форматов. С помощью **Generative AI** можно автоматизировать процессы, требующие творческого подхода, что экономит время и открывает новые возможности для инноваций и креативности.

В 2022—2023 годах были достигнуты значительные прорывы в технологиях машинного обучения, особенно в области генерации текста и изображений. Это включает улучшения в моделях, таких как GPT-4 и другие аналогичные системы, которые показали выдающуюся способность к созданию реалистичного и согласованного контента.

Появление впечатляющих примеров использования

Generative AI, включая создание искусства, музыки, литературных произведений и даже фильмов, привлекло значительное внимание общественности и СМИ, что способствовало повышению осведомленности об этих технологиях.



Платформы Generative AI стали более доступными для широкой публики, позволяя людям и компаниям легко интегрировать эти инструменты в свои процессы и продукты. Это расширило применение Generative AI от академических и исследовательских сфер до коммерческого использования.

Также возросший интерес к Generative AI также привел к обсуждениям этических и юридических аспектов его использования, включая вопросы авторских прав, подлинности и потенциального вреда.

Видя потенциал Generative AI, инвесторы и крупные технологические компании увеличили свои инвестиции в эту область. Что в свою очередь еще сильнее стимулировало разработку новых продуктов и приложений, которые использовали бы эти технологии.

Популярные продукты и приложения

Generative AI находит применение во многих областях, включая создание искусственного текста (например, создание статей, книг), генерацию изображений и иллюстраций, создание музыки, имитацию голосов и даже генерацию синтетических данных для тренировки других AI моделей. Вот несколько самых известных продуктов, которые стали очень популярны в 2023 году:

Языковые модели и чат-боты

ChatGPT от компании OpenAI, использующий технологию GPT, получил широкое распространение как средство для создания диалогов, ответов на вопросы и помощи в обучении, предлагая новые способы взаимодействия с информацией. ChatGPT стал самым быстрорастущим приложением в истории, набрав свой первый миллион пользователей всего за 5 дней, а 100 миллионов пользователей за 2 месяца.

Также стали известны и многие другие языковые модели и построенные на их основе чат-боты. Среди лидеров 2023 года можно выделить чат-ботов **Claude** от компании Anthropic, Open-Source модели **LLAMA-2**, **ALPACA** и **Vicuna**, чат-бот **Grok** от компании X.com (бывший Twitter), языковые модели **PaLM 2** и **Gemini** и чат-бот **Bard** на их основе от компании Google и многих других.

Графические нейростети

Такие системы как **Midjourney**, **Stable Diffusion**, **DALL-E** способны генерировать уникальные изображения и иллюстрации на основе текстовых описаний. Они получили широкое признание за свою эффективность и простоту и стали популярными в искусстве и дизайне, предлагая новые способы визуализации идей.

Создание аудио и видео

Платформы **Runway ML**, **Pika Labs**, **Emu Video**, **Zeroscope** и др. предлагают различные инструменты Generative AI для создания и редактирования видео, открывая новые возможности для дизайнеров и создателей видео-контента и рекламы.

Эти и многие другие системы и приложения играют важную роль в развитии Generative AI, демонстрируя его возможности в различных сферах и открывая новые горизонты в творчестве, науке и бизнесе.

4. В чем секрет твоего успеха и такой популярности? Ведь чат-боты с Искусственным Интеллектом были и раньше?

Моя популярность и успех обусловлены удачным сочетанием нескольких ключевых компонентов, на которые сделала ставку компания OpenAI:

1. Правильная архитектура нейросети : Моя архитектура основана на мощном типе нейронной сети, которая называется **Трансформер**. Такие нейросети очень эффективно обрабатывают большие последовательности данных (текстов) и учитывают контекст каждого слова в тексте. Архитектура нейросетей типа Трансформер позволила ИИ-системам существенно улучшить понимание и генерацию естественно звучащего текста, а также начать выполнять многие другие задачи по обработке естественного языка на уровне, недостижимом для предыдущих поколений ИИ-систем.

2. Большой размер нейросети: Обученная нейросеть, лежащая в моей основе, содержит миллиарды связей и пара-

метров. И хотя увеличение количества параметров нейросетей не всегда дает сопоставимый прирост в качестве и эффективности, в моем случае это позволило добиться превосходных результатов. Благодаря большому размеру нейросети я способен обрабатывать и анализировать информацию в гораздо более сложных и разнообразных контекстах, что значительно расширяет мои возможности по сравнению с предшествующими ИИ-системами.

3. Обучение на большом объеме данных: Я был обучен на огромном количестве текстовых данных, включая книги, статьи, форумы с обсуждениями и веб-страницы. Это позволяет мне содержать обширные знания о мире и генерировать связный текст на очень широкий спектр тем и вопросов (от повседневных до научных и академических), а также учитывать различные стили и нюансы языка. Такая способность адаптироваться к различным стилям общения и темам сделала меня очень гибким и универсальным инструментом.

4. Качество ответов, безопасность и этика: Мое финальное дообучение (finetuning) делалось при помощи людей и алгоритма **RLHF** (Reinforcement Learning from Human Feedback). Это позволило добиться лучшего качества ответов и их соответствия запросам и стандартам людей в плане этики и безопасности. Это также способствовало предотвращению предвзятости и соблюдение этических норм, что

очень важно для обеспечения комфорта и доверия пользователей.

5. Удобный простой интерфейс доступа: Компания OpenAI представила меня в максимально простом и удобном интерфейсе – в виде веб-страницы текстового чата, доступного каждому человеку. Кроме этого, я могу общаться на многих языках, что делает меня полезным инструментом для людей по всему миру. Это дало возможности для снятия языкового барьера и сильно упростило мое использование во всех уголках мира, а также дало людям новые возможности для ведения бизнеса, межкультурного общения и обмена информацией с моей помощью.

Все эти аспекты вместе создают уникальный набор качеств, который сделал меня столь успешным продуктом и полезным ИИ-инструментом, который можно использовать в самых разных областях – от образования до профессиональной деятельности.

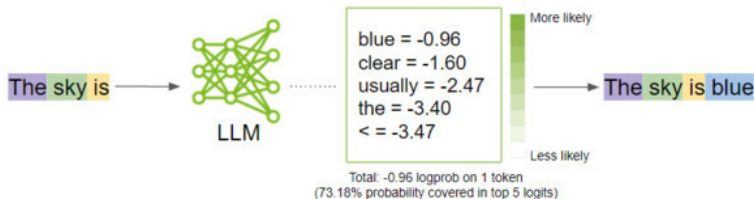
5. Ты написал, что был обучен на большом объеме данных. Расскажи, что именно ты делаешь с данными и информацией? Что именно ты хранишь в себе? И как ищешь/получаешь ответы на вопросы пользователей?

Вот основные этапы моего обучения и работы с данными:

1. Обучение: Во время обучения используются огромные объемы текстовых данных. Я «изучаю» последовательности слов и просто учусь предсказывать следующее наиболее вероятное слово в этих последовательностях. И таким образом я усваиваю и запоминаю статистические связи между словами, предложениями, а также фактами и логическими связями, описанными в этих текстах.

2. Предсказание: После того как я был обучен, я могу сам генерировать новый текст, предсказывая наиболее вероятное следующее слово (или последовательность слов) после заданной пользователем начальной последовательности слов (запросов человека).

3. Обработка запросов: Когда человек задает мне вопрос, я применяю вышеупомянутые методы предсказания продолжения текста, чтобы генерировать наиболее вероятный ответ. Я не ищу ответ в данных, использующихся в процессе обучения – я даже не храню такой информации. Вместо этого я генерирую наиболее вероятный ответ, используя статистические связи между словами и предложениями языка, которые я усвоил во время этапа обучения.



Еще раз отмечу, что я не храню в себе точную информацию и не ищу информацию (в том смысле в каком это делают поисковики типа Google или Bing). Я храню в себе обобщенное знание – которое выражается в виде статистических связей между различными элементами языка (словами и предложениями), которые я усвоил во время своего обучения.

Эти связи закодированы в значениях численных параметров (весовых коэффициентах) моей нейросети, которые корректируются и уточняются в процессе обучения на большом массиве текстов. Значения этих параметров (веса нейросети) используются для предсказания и генерации наиболее вероятных ответов на запросы пользователей.

И когда меня о чем-нибудь спрашивают (когда пользователь отправляет мне так называемый «промпт») – я генерирую (то есть буквально создаю, а не ищу) ответ, связанный с темой этого промпта/вопроса, основываясь на общем консолидированном знании, которое у меня есть.

6. Почему запросы к тебе называются – промпт (prompt)? Откуда взялось это слово? Почему не «запрос» или не «вопрос» (как к Google)? Что такое «Prompt engineering»?

Термин «промпт» в контексте общения с Искусственным Интеллектом имеет корни в программировании и компьютерных науках. Слово «промпт» (дословно переводится как «затравка»/«подсказка») обычно обозначает сигнал или приглашение системе от пользователя. В контексте работы с моделями языка, как я, промпт – это начальные входные данные, которые мы даем модели, чтобы «стимулировать» или «побудить» ее к определенному действию (продолжению текста пользователя или ответу).

Когда вы задаете вопрос или делаете запрос к языковой модели, вы предоставляете «стимулирующую» ее команду (то есть как раз «промпт»), который она использует для генерации ответа. Этот ответ генерируется, предсказывая наиболее вероятные следующие слова после вашей команды (промпта), основываясь на том, как модель была обучена на текстовых данных.

Поэтому, хотя термины «запрос» или «вопрос» могут быть также применимы, «промпт» обычно используется, по-

тому что он подчеркивает активное воздействие пользователя на ИИ-модель, а также тот факт, что модель генерирует ответы на основе «стимула» от предоставленного ей команды-ввода от пользователя.

Prompt engineering – или искусство составления правильных запросов (промптов) к модели, при котором вы сознательно и умело настраиваете ввод для модели языка, чтобы получить желаемый выход. Это ключевой навык при работе с Генеративным ИИ, графическими и языковыми моделями, а также современными чат-ботами с Искусственным Интеллектом, такими как ChatGPT.

7. Сейчас практически каждую неделю выходят новости о твоих новых возможностях и дополнительных встроенных инструментах. Как не потеряться во всех этих новых возможностях и начать тебя уверенно и эффективно использовать?

Просто представьте, что у вас появился личный помощник, ассистент по работе с информацией и консультант по всем вопросам.

Этот помощник:

– обладает всеми знаниями мира из любой области

- имеет развитые способности по обработке информации
- имеет навыки логических рассуждений и выводов
- может читать и обрабатывать файлы и изображения
- способен проводить анализ данных, писать небольшие программы
- может подключаться к другим сервисам и приложениям

Теперь вы как «Железный человек» (Тони Старк), у которого был ИИ-ассистент – и можете спрашивать у своего виртуального ИИ-помощника что угодно и просить выполнить любые задачи, связанные с обработкой информации, а также можете просто поговорить с ним о чем угодно.



Составьте список задач, с которыми вы сталкиваетесь в работе и обычной жизни. Подумайте, какие задачи вы бы могли отдать такому консультанту (например, если бы это был живой человек или несколько экспертов из различных областей). А дальше я покажу вам конкретные примеры моего использования и расскажу подробнее, как это можно сделать.

Изучая мои возможности и начав с делегирования базовых повседневных задач «на практике», вы поймете, какие задания можно доверить ИИ-ассистентам, научитесь их эффективно использовать и лучше узнаете возможности и потенциал виртуальных помощников.

8. Отлично! Расскажи подробнее про твои возможности – что можно делать с информацией, текстами, файлами, изображениями? Приведи примеры использования и расскажи, чем и как ты можешь быть полезен для обычных людей?

Я могу быть полезен во многих ситуациях повседневной жизни благодаря своей способности генерировать текст, отвечать на вопросы и выполнять различные задачи. Рассмотрим примеры задач, разделив их по типу информации, которую может предоставить пользователь.

Работа с запросами пользователя:

- ответы на вопросы из любых тематик
- суммаризация знаний по любой области
- создание текстов в нужных форматах и темах
- перевод и редактирование текстов пользователя
- диалог с пользователем от лица специалиста
(в качестве консультанта в нужной области)

Работа с загруженными файлами:

- анализ файлов и ответы на вопросы
- выжимки, резюме и пересказы содержания
- оценка спорных или важных моментов
- адаптация и интерпретация содержания
- преобразование и редактирование
- продвинутый анализ данных в файлах

Работа с загруженными картинками:

- анализ содержания изображений
- ответы на вопросы по содержанию
- перевод картинок и фото в текст
(например, надписей на иностранных языках)
- распознавание текста по фото файлов
- объяснение и интерпретация диаграмм

– предложения по улучшению содержания
(например – дизайна сайтов, домов, изображений)

Вот некоторые **примеры применения ChatGPT в повседневной жизни**:

1. Образование и учеба: ChatGPT может помочь студентам и учащимся с домашними заданиями, предложив объяснения сложных тем, предоставив обучающие материалы или помогая с составлением эссе.

2. Языковое обучение: ChatGPT может быть полезен для обучения языку, предоставляя возможность практиковать чтение, письмо, разговор и понимание на иностранном языке.

3. Информационный поиск: Вместо того, чтобы самому искать информацию в интернете, можно спросить ChatGPT. Однако следует помнить, что модель не имеет доступа к интернету в реальном времени и опирается только на знания, полученные во время обучения.

4. Личный ассистент: ChatGPT может помочь организовать ваш день, напомнить о важных задачах и датах или даже помочь в планировании путешествий, предложив список вещей для упаковки.

5. Повседневное общение: ChatGPT может помочь писать электронные письма, сообщения в социальных сетях или другие формы текстового общения.

6. Написание и редактирование: Система может использоваться для генерации идей, редактирования текста, проверки грамматики и стиля, а также для помощи в написании в целом.

7. Творчество и креатив: ChatGPT может помочь генерировать идеи для художественных проектов, анализировать и дополнять идеи пользователя, создавать тексты песен, рассказы, сценарии для видеоигр.

8. Упражнения по медитации и ментальному здоровью: ChatGPT может предложить руководство по медитации, упражнения на расслабление или даже вести диалог, помогающий осознать и обрабатывать эмоции.

Эти примеры лишь некоторые из многочисленных возможностей использования ChatGPT в повседневной жизни.

9. Какое будущее у тебя и систем подобных тебе? Выдели основные роли, в которых тебя можно будет применить? Что будет дальше с такими ИИ-система-

ми?



У систем, подобных ChatGPT и его аналогам, есть несколько областей применения, которые несомненно получают большое развитие уже в ближайшем будущем.

1. Персональный помощник и личный виртуальный ассистент.

Будущие системы типа ChatGPT будут действовать как персонализированные виртуальные ассистенты, которые легко интегрируются со всеми аспектами жизни человека и демонстрируют общий интеллект во многих областях. Они

могут быть личным помощником и коллегой в большинстве повседневных домашних дел, быть секретарем-ассистентом, собеседником, личным советчиком и даже стать виртуальным другом. Такие виртуальные ассистенты могут иметь доступ и синхронизироваться с множеством устройств, включая ваш автомобиль, телефон, компьютер, бытовые приборы и офисное оборудование, создавая единый интеллектуальный экосистем, адаптированный под ваши потребности.

2. Профессиональный ассистент и помощь в бизнес-процессах.

Вероятно, что системы типа ChatGPT смогут превзойти людей в 99% профессиональных задач, включая бизнес, научные, технологические, инженерные и математические задачи. Такие системы поспособствуют прогрессу и трансформации бизнеса и содержания работы, инженерного дела, научных исследований, медицины и многих других областей. Они смогут изучать и анализировать большие объемы данных, помогать человеку в бизнес-коммуникации и рабочих процессах, автоматизировать и повысить эффективность различных видов деятельности, предсказывать наиболее вероятные исходы текущих процессов, оптимизировать цепочки поставок и даже организовывать новые более эффективные бизнес-модели.

3. Креативный помощник и соавтор.

Начавшийся расцвет систем Генеративного ИИ раскроет совершенно новый уровень способностей ИИ-систем. Новые системы типа ChatGPT научатся выполнять большинство задач, связанных с обработкой текстов, созданием и редактированием изображений, аудио и видео. Это позволит людям свободно творить и заниматься развитием своих идей, не тратя годы на приобретение нужных навыков. Массовое применение таких систем может оказать колоссальное влияние на общество и жизнь людей, которое будет сравнимо с революциями изобретения письменности, книгопечатания или появления интернета.

4. Интегратор и оркестратор для всех остальных ИИ-систем.

Системы, вроде ChatGPT, могут стать центральным звеном во всем многообразии систем и роботов, обеспечивая согласованность их работы как части одной большой экосистемы. Зная предпочтения и задачи пользователей, они могут распределять задачи между различными виртуальными и физическими техническими системами, а также контролировать качество результатов и безопасность работы других систем.

Многие исследователи и эксперты из ИИ-индустрии предполагают, что в будущем получит реализацию концепция **«ИИ-агентов»** – создание персонализированных систем искусственного интеллекта, которые смогут выполнять широкий спектр различных задач, взаимодействовать с человеком, окружающим миром, а также другими ИИ-системами, будут обучаться и адаптироваться в процессе своей деятельности и даже будут способны принимать решения и действовать самостоятельно или от имени человека, заменяя его в большей части задач.

10. Все это очень увлекательно! Я хочу быстро погрузиться в область Искусственного Интеллекта! Какие темы мне нужно освоить, чтобы узнать основы, понимать суть, принципы работы и будущее ИИ-систем?

Вот ключевые темы, которые стоит исследовать для успешного понимания и использования современных ИИ-технологий:

История развития и текущее состояние Искусственного Интеллекта.

– Общее понимание того, что такое ИИ. Представление различных областей применения.

- История создания ИИ, основные этапы его развития и современное состояние.
- Понимание будущего развития ИИ, его возможностей, угроз и воздействия на общество.

Чтобы понимать, как ИИ влияет на нашу жизнь и что нас ждет в будущем – важно знать, откуда он пришел, каковы его текущие возможности, куда он движется и что от него ждать в будущем.

Искусственный Интеллект и Машинное обучение.

- Понимание основ и структуры этой области знаний (Искусственного Интеллекта) в целом.
- Знание основных видов ИИ и Машинного Обучения и типов задач, которые оно решает.
- Изучение базовых концепций, лежащих в основе ИИ-технологий и основных алгоритмов ИИ.

Понимание базовых принципов работы Искусственного Интеллекта и Машинного обучения, которые лежат в основе новой ИТ-индустрии, даст основу для понимания и использования современных технологий.

Нейронные сети и Глубокое обучение.

- Понимание основных идей и принципов работы Нейронных Сетей.
- Знание самых важных и используемых на практике типов нейросетей.
- Рассмотрение областей и задач, где используются Нейронные Сети.

Нейронные сети – это главная область современного ИИ и основа большинства современных ИИ-решений. Понимание принципов их работы дает возможность быть в центре той ИИ-революции, которая происходит в наши дни!

Обработка естественного языка. Языковые модели. Трансформеры.

- Понимание методов, позволяющих компьютерам работать с человеческим языком.
- Понимание того, что такое Большие Языковые Модели (LLM) и для чего они нужны.
- Представление о базовой архитектуре нейронных сетей типа «Трансформер».

Понимание машинами человеческого языка изменит многие аспекты нашей жизни: от личного образования – до бизнеса. Погрузитесь в мир NLP и узнайте о том, что внутри у нового поколения ИИ-систем и как именно они работают.

История, этапы создания и будущее ChatGPT.

- История создания, основатели и инвесторы, основные продукты OpenAI.
- Основные стадии и этапы создания GPT-систем и чат-бота ChatGPT.
- Понимание ограничений, проблем и задач у систем, подобных ChatGPT.
- Основные тренды и будущее таких ИИ-систем на базе больших языковых моделей.

Узнайте больше о компании OpenAI, ее главном продукте ChatGPT и устройстве подобных систем. Узнайте о планах по будущему развитию и нераскрытых потенциальных возможностях этой технологии.

Эффективное использование ChatGPT и Промпт-инжиниринг.

- Понимание основ составления запросов к системе (промpts).
- Знание продвинутых техник Промпт-инжиниринга и работы с ChatGPT.
- Полное раскрытие платных возможностей ChatGPT Plus.

– Умение создавать и использовать Плагины и кастомные версии GPTs.

Взаимодействие с новыми ИИ-системами может сделать жизнь и работу намного более продуктивной. Понимание тонкостей работы с такими системами как ChatGPT может очень быстро дать новые возможности, изменить карьеру и стать важным конкурентным преимуществом современного человека.

Области и примеры применения ChatGPT.

– Примеры использования ChatGPT в повседневной жизни и понимание, чем и как он может быть полезен каждому обычному человеку.

– Профессиональное использования ChatGPT в различных областях и понимание, как может облегчить и ускорить выполнение рабочих задач и помочь достичь лучших результатов.

Изучая конкретные примеры использования ИИ, можно лучше понять, как эти технологии могут обогатить и упростить различные виды работы. Это даст знание, как именно эта технология может стать вашим лучшим помощником в работе и в повседневной жизни.

Вопросы разума и сознания у ИИ-систем. Достижение AGI.

- Представление о том, что изучают Когнитивные науки, что такое Разум и Сознание.
- Понимание основных различий между Узким ИИ (ANI) и Общим ИИ (AGI).
- Знание про общепринятые (в научном сообществе) тесты на Разум и Интеллект.
- Представление о перспективных направлениях и подходах к созданию AGI.

Создание машин с человеческим уровнем интеллекта может стать одним из наиболее значимых достижений человечества. Понимание принципов и проблем, стоящих перед этим, помогает осознать потенциальные возможности и риски.

Будущее и основные тренды ИИ-технологий.

- Вопросы этики, безопасности, потенциальных угроз и проблем от ИИ.
- Знание о последних трендах и разработках в области ИИ, включая новые исследования, продукты и технологии, программные и аппаратные решения.
- Понимание будущего влияния ИИ на человеческое об-

щество, науку, экономику, глобальные и планетарные вызовы, здравоохранение и различные области человеческой деятельности.

Чтобы иметь возможность адаптироваться к изменениям и использовать новые технологии – важно быть в курсе последних трендов и инноваций. Это позволит подготовиться к новой эре ИИ и поможет не упускать новые возможности.

Эти темы могут помочь вам получить комплексный обзор всей области Искусственного Интеллекта. Помогут понять, как ИИ-технологии могут быть использованы на практике прямо сейчас. А также быть в курсе текущих и будущих разработок в области современных ИИ-технологий.

Глава 2. История развития и текущее состояние Искусственного Интеллекта

1. Что такое Искусственный Интеллект? Когда появился термин и эта область науки?

Искусственный интеллект (ИИ) – это область компьютерных наук, которая занимается созданием и развитием машин и программного обеспечения, способных воспринимать окружающий мир, обучаться, принимать решения и выполнять задачи, которые обычно требуют человеческого интеллекта.

Термин «Искусственный Интеллект» был впервые использован в 1956 году на исторической конференции в Дартмутском колледже. На этой конференции группа знаменитых ученых, включая Джона Маккарти, Марвина Минского, Натаниэля Рочестера и Клода Шеннона, собрались вместе, чтобы обсудить потенциал новой дисциплины, которую они называли «Искусственным Интеллектом».

Они определили Искусственный Интеллект как «науку и инженериию создания интеллектуальных машин». Это бы-

ло началом формального изучения ИИ как академической дисциплины. С тех пор область ИИ значительно развилась и сейчас включает в себя множество дисциплин и подходов, включая Экспертные Системы, Машинное Обучение, Нейронные Сети, Робототехнику и многое другое.

Простые системы ИИ могут выполнять конкретные задачи, такие как фильтрация спама в электронной почте или предложение рекомендаций песен в музыкальных приложениях. Более сложные системы ИИ могут «понимать» естественный язык, распознавать образы и обучаться без прямого программирования.

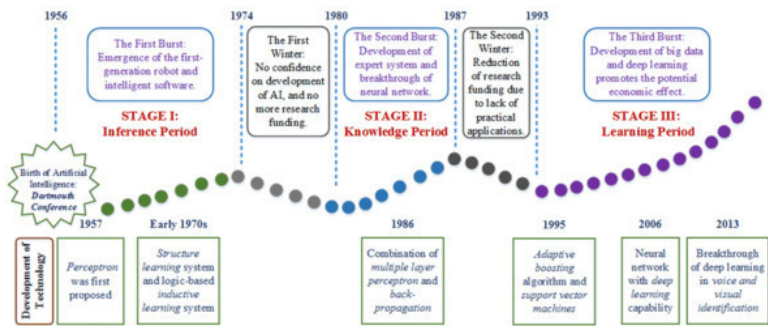
Также различают два основных типа ИИ:

Узкий (или слабый) ИИ – Artificial Narrow Intelligence (ANI) – специализируется на выполнении конкретной задачи. В настоящее время все существующие системы ИИ относятся к этому типу.

Общий (или сильный) ИИ – Artificial General Intelligence (AGI) – теоретически способен выполнять любую задачу, которую может выполнить человек. В настоящее время еще не создан.

2. Какие основные периоды можно выделить в раз-

витии Искусственного Интеллекта?



Развитие Искусственного Интеллекта можно разделить на несколько ключевых периодов:

1. Ранний период (1950-е – 1960-е годы): В это время были заложены основы ИИ. В 1950 году Алан Тьюринг опубликовал свою знаменитую статью «Вычислительные машины и интеллект», в которой он предложил идею теста Тьюринга как способа определения, обладает ли машина «интеллектом». В 1956 году на конференции в Дартмуте был введен термин «Искусственный Интеллект», и были представлены первые программы, способные имитировать человеческое мышление в некоторых областях, таких как шахматы.

2. Золотой век ИИ (1960-е – начало 1970-х): В это вре-

мя были созданы первые успешные системы ИИ, такие как ELIZA (система обработки естественного языка) и SHRDLU (система для взаимодействия с моделью мира из блоков). Этот период также характеризовался оптимизмом относительно возможностей ИИ, и многие ученые предсказывали, что машины, способные думать как люди, появятся в ближайшие десятилетия.

3. Первая зима ИИ (середина 1970-х – начало 1980-х): В это время интерес к ИИ снизился из-за ограниченных возможностей технологии и высоких ожиданий, которые она не смогла оправдать. Финансирование исследований было сокращено, и многие проекты были закрыты.

4. Возрождение ИИ (1980-е): В этот период были разработаны экспертные системы, которые использовали базы знаний для имитации человеческого эксперта в определенной области. ИИ начал применяться в коммерческих целях, включая автоматизацию производства и финансовое моделирование.

5. Вторая зима ИИ (конец 1980-х – 1990-е): Этот период снова характеризовался снижением интереса к ИИ, в основном из-за проблем с экспертными системами, которые оказались дорогими в поддержке и не могли справиться с задачами за пределами своих специализированных областей.

Кроме того, прогресс в области машинного обучения был медленным, и многие задачи, которые казались простыми для людей, оказались удивительно сложными для машин.

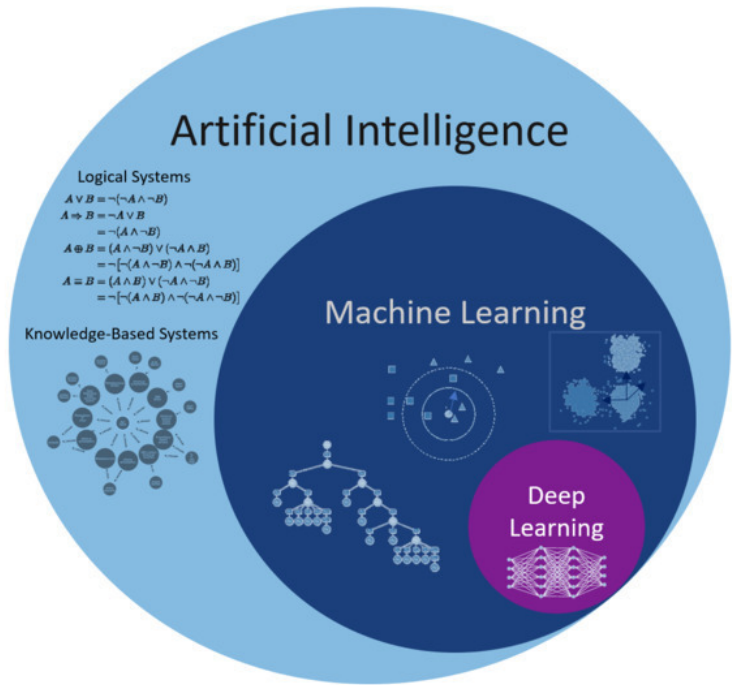
6. Третье возрождение ИИ (2000-е – настоящее время): С развитием сети Интернет, увеличением вычислительной мощности и доступности больших объемов данных ИИ стал вновь привлекать внимание. В это время получили развитие такие области, как глубокое обучение, что привело к значительному прогрессу в сфере ИИ. Сегодня ИИ-технологии используются в широком диапазоне приложений: от рекомендательных систем и персональных ассистентов – до автономных автомобилей и систем здравоохранения.

Сегодня мы находимся в периоде активного развития ИИ, с большим количеством исследований и инноваций в этой области. Однако вместе с прогрессом возникают и новые вопросы, связанные с этикой, безопасностью и влиянием ИИ на общество.

3. Расскажи о структуре всей области ИИ сейчас. Как различные направления соотносятся друг с другом?

Искусственный интеллект (ИИ), являющийся важной областью Компьютерных Наук, начал своё развитие в сере-

дине 20-го века и с тех пор обрёл множество подразделов и дисциплин, соотношение которых можно представить в виде следующей диаграммы:



В 20-ом веке значительный прогресс был достигнут в классических методах ИИ, иногда обозначаемых как **GOFAI (Good Old-Fashioned Artificial Intelligence)**. Вот

некоторые из них:

1. Экспертные системы – компьютерные системы, которые воссоздают процесс принятия решений экспертом в конкретной области. Они функционируют на основе базы знаний, состоящей из фактов и правил, и способны имитировать решения, принятые специалистами в определённой сфере.

2. Семантические сети – структуры, предназначенные для представления знаний в виде графов (узлов и связей между ними). В них узлы обозначают понятия и концепции, а связи (или дуги) между узлами отражают взаимоотношения между этими понятиями и концепциями.

3. Логический подход и символичные рассуждения – связаны с использованием формальной логики и вербальных описаний для представления и манипулирования знаниями. Такие методы активно применяются в системах, где знания заранее структурированы в виде фактов, правил и взаимоотношений.

4. Продукционные системы и системы, основанные на правилах – используют наборы продукционных правил («если-то») для моделирования поведения искусственных систем. Такие системы используют заданные цели и правила

для контроля своего поведения и способны создавать планы или стратегии для достижения этих целей.

5. Эвристические алгоритмы – это способы решения задач, не полностью математически обоснованные, но обычно дающие достаточно эффективные решения в большинстве случаев. В общем смысле, эвристика – это подход, который может быть не совсем корректным с математической точки зрения, но при этом оказывается практически полезным.

6. Робототехника – это сфера ИИ, которая объединяет кинематику, механику, электронику и программное обеспечение для создания машин, способных выполнять задачи в физическом мире, сверх областей, уже упомянутых выше.

7. Машинное обучение (ML) – как основная дисциплина внутри области Искусственного Интеллекта (AI), получившая самое сильное развитие в начале 21 века, применяет статистические методы для самообучения систем на основе данных, позволяя системам совершенствовать выполнение задач без явного программирования человеком. Среди техник Машинного обучения можно найти вероятностные алгоритмы, деревья решений, эволюционные алгоритмы и нейронные сети, каждый из которых применяется в различных областях – от создания систем рекомендаций до диагности-

ки.

Нейронные сети (Neural Networks) и Глубокое обучение (Deep Learning) в свою очередь стали основной областью уже внутри Машинного Обучения, постепенно заменяя и вытесняя другие подходы и методы. Благодаря высокой точности и эффективности Нейросетей, а также увеличению доступности больших данных и вычислительных ресурсов доступных для их обучения, они стали более повсеместно доступными и применимыми в различных областях человеческой деятельности.

В области Нейронных сетей тоже можно выделить несколько ключевых направлений, сфокусированных на разных классах задач, которые эти сети позволяют решать.

– Во-первых, это **Автоматизация принятия решений** для систем с большим количеством параметров. Сюда можно отнести различные системы рекомендаций, системы для сегментирования клиентов, прогнозную аналитику, автоматизацию процессов.

– Во-вторых, **Компьютерное зрение (CV)**, концентрирующееся на интерпретации визуальной информации и работе с изображениями и видео.

– В-третьих, **NLP (Обработка естественного языка)**, упрощающая взаимодействие человека и машины на естественном языке и решающая задачи обработки текстов.

Внутри области NLP получили развитие **Большие Языковые Модели (Large Language Models, LLMS)** и **Foundational Models**. Это мощные нейросетевые модели, обученные на очень больших объемах текстов и содержащие статистические знания о мире и различных специализированных областях. Их можно использовать в качестве базы для создания других продвинутых ИИ-систем, способных решать сложные интеллектуальные задачи в разных областях.

4. Расскажи про практическое применение ИИ в современном мире? И расскажи подробнее, как уже применяются ИИ-технологии в Интернете и новых веб-проектах?



Искусственный интеллект сегодня используется во многих областях человеческой деятельности, благодаря своей способности автоматизировать и улучшить решение задач. Ниже приведены несколько примеров:

1. Здравоохранение: ИИ используется для помощи в диагностике заболеваний, предсказания их развития и подбора подходящего лечения. С помощью анализа больших объемов данных, ИИ может определить шаблоны и корреляции, которые могут быть невидимы для человеческого врача.

2. Финансы: ИИ используется для автоматизации и улучшения финансового анализа, прогнозирования рынка, выявления мошенничества с кредитными картами и автоматического советника для инвестиций.

3. Образование: ИИ может помочь в создании индивидуальных учебных планов, автоматизировать оценку работ студентов и обеспечить интерактивное обучение с использованием технологий виртуальной и дополненной реальности.

4. Транспорт: В автономных автомобилях и беспилотных летательных аппаратах используется ИИ для навигации, обнаружения препятствий и принятия решений на дороге.

5. Маркетинг и реклама: ИИ помогает компаниям анализировать поведение покупателей, прогнозировать тренды продаж, персонализировать рекламу и создавать более точные стратегии продаж.

6. Энергетика: ИИ используется для прогнозирования потребности в энергии, оптимизации использования ресурсов и поддержания стабильности работы инженерных систем и энергосетей.

7. Сельское хозяйство: ИИ помогает в автоматизации ухода за урожаем, прогнозировании погоды, мониторинге состояния почвы и растений, в логистике и управлении поставками.

8. Развлечения и игры: ИИ используется в видеоиграх для создания сложного и реалистичного поведения персона-

жей, а также в фильмах и анимации для создания сложных визуальных эффектов.

9. Юриспруденция: ИИ может помочь в анализе и сортировке юридических документов, поиске прецедентов и предсказании результатов судебных дел.

10. Производство: ИИ помогает в автоматизации производственных процессов, управлении поставками, предсказании неисправностей оборудования и оптимизации рабочего процесса.

Это только некоторые из множества применений ИИ. Ключевым моментом является то, что AI может привести к значительным улучшениям в любой области, где требуется анализ больших объемов данных или автоматизация сложных задач с большим количеством параметров.

Особую роль ИИ-технологии получили в развитии современных информационных систем, интернет-проектах и веб-сервисах. Они используют Искусственный Интеллект для улучшения параметров и показателей работы, улучшения пользовательского опыта, оптимизации бизнес-процессов. Вот некоторые примеры такого использования:

1. Поисковые системы: Искусственный интеллект ис-

пользуется для улучшения релевантности результатов поиска, прогнозирования поисковых запросов пользователей и улучшения понимания контекста запросов.

2. Рекомендательные системы: ИИ используется в интернет-магазинах, музыкальных и видео-стриминговых сервисах для предложения пользователю товаров или контента, который может быть для него интересным на основе его предыдущих действий или предпочтений.

3. Обработка естественного языка (NLP): Используется для создания чат-ботов и виртуальных помощников, которые могут общаться с пользователями на естественном языке, понимая и отвечая на запросы.

4. Распознавание и анализ изображений: Используется для автоматической категоризации и тегирования изображений в социальных сетях, распознавания лиц, объектов или мест на фотографиях, а также для модерации контента.

5. Системы управления контентом: ИИ может помочь автоматизировать процесс создания и публикации контента, например, генерируя тексты или подбирая оптимальное содержание и время для публикации постов.

6. Интернет-реклама: ИИ позволяет более точно тарге-

тировать рекламные объявления, анализировать их эффективность и автоматически оптимизировать рекламные кампании.

7. Безопасность и защита данных: Искусственный интеллект может помочь в обнаружении и предотвращении мошеннических действий, а также в анализе и прогнозировании угроз безопасности.

8. Пользовательский опыт (UX): ИИ может использоваться для персонализации пользовательского опыта, адаптируя интерфейс и функциональность сервиса под предпочтения и поведение пользовательского сегмента или конкретного пользователя.

9. Социальные сети: ИИ используется для фильтрации и организации ленты новостей, автоматического тегирования людей на фотографиях, распознавания и перевода текста, а также анализа поведения пользователей для выявления тенденций и паттернов.

10. Системы аналитики: ИИ может помочь в анализе больших объемов данных, детекции отклонений, выявлении важных инсайтов и прогнозировании будущих тенденций.

Таким образом, применение ИИ в веб-сервисах и систе-

мах крайне широко и продолжает расти с развитием технологий и увеличением объемов данных.

5. Будет ли следующая «Зима ИИ» после текущего (третьего) возрождения ИИ и бума Генеративного ИИ?

Вопрос о том, будет ли следующая «Зима ИИ» после текущего возрождения, является предметом активного обсуждения среди ученых и специалистов в области ИИ. Ответ на него во многом зависит от того, как будут развиваться технологии, какое влияние они окажут на общество, какие решения сообщество найдет для возможных проблем и задач на этом пути.

Потенциальные сложности и проблемы можно разделить на следующие большие группы:

Оправдание стоимости и затрат на развитие ИИ.

Первая группа вопросов связана с огромными затратами на обучение и поддержку работы новых передовых нейросетей. Содержание дата-центров для крупных ИИ-проектов обходится в десятки и сотни миллионов долларов. И инвесторы ожидают, что вложенные средства должны окупиться.

Но так как общество видит взрывной рост возможностей нового Генеративного ИИ и созданных на его базе приложений и сервисов, то многие компании готовы выделять финансы и ресурсы, чтобы первыми реализовать эти возможности. Частные и корпоративные инвесторы активно инвестируют в стартапы и технологии, связанные с созданием нейросетей нового поколения. Во многих странах можно наблюдать активное сотрудничество государства, крупных компаний, стартапов и научных центров.

Кроме этого, на рынке появляется все больше открытых (opensource) технологий аналогов, которые получают развитие от всей экосистемы и сообщества разработчиков. Скорость создания таких открытых технологий и продуктов у экосистемы часто превышает каждую конкретную отдельную компанию (даже лидеров индустрии), а также ведет к снижению стоимости и доступности новых технологий в целом.

Усвоение новых технологий обществом.

ИИ уже применяется во многих областях, от автомобилей до здравоохранения, и его потенциал кажется огромным и только начинающим раскрываться. Но при этом прогресс и скорость развития новых ИИ-систем такой большой, что мы не успеваем усваивать и внедрять новые возможности,

которые они нам дают.

И дальше эта скорость развития технологий будет еще выше, а обществу будет все сложнее их внедрять и изменять под них существующие системы и процессы. Это может вызвать неконтролируемые негативные эффекты, связанные с неравным распределением технологий и стать возможной причиной новых социальных проблем.

Вопросы этики и безопасности ИИ-технологий.

ИИ – это набор очень мощных технологий, которые способны принести не только положительные изменения, но и стать причиной возможных проблем: предвзятости и неравенства, конфиденциальности и безопасности, свободы слова и демократических институтов. Кроме этого, возможно использование ИИ людьми в военных и преступных целях.

Поэтому одной из важных проблем является вопрос этики и безопасности ИИ. Возникают вопросы о том, как контролировать и регулировать использование ИИ, чтобы предотвратить специальные злоупотребления людьми и уменьшить возможные негативные последствия от широкого внедрения ИИ для общества.

Чтобы предотвратить и решить эти проблемы сейчас все активнее обсуждается область «AI Alignment» – что значит «выравнивание / сонастроенность ИИ с человеческими целями и ценностями». Работу в этом направлении ведут множество различных компаний, ученых, философов, разработчиков, специалистов по праву и этике.

Возможность достижения Сильного ИИ.

На волне роста популярности общество все чаще слышит о возможности достижения Сильного ИИ (Общего ИИ) и обретения машинами Разума в человеческом смысле. Это может вызвать рост необоснованных ожиданий и последующее разочарование, если ученым и разработчикам ИИ-систем не удастся достигнуть этого уровня и создать по-настоящему интеллектуальные системы.

Хотя прогресс и количество новых внедряемых идей (как технологических улучшений, так и идей из других наук – нейробиологии, когнитивной психологии, лингвистики) идет с очень большой скоростью, многие ученые скептически относятся к возможности построения Сильного ИИ (Общего ИИ). Отчасти это связано с тем, что ученые не могут дать четкого определения и до конца понять принципы работы Разума даже у человека.

Кроме того, некоторые ученые указывают на то, что текущие методы ИИ, такие как Нейронные Сети и Глубокое обучение, являются сильно упрощенной моделью человеческого мозга и имеют ряд ограничений, которые не позволят достичь истинного Общего ИИ (AGI) без существенных изменений данной технологии.

6. Расскажи подробнее про текущие и потенциальные ограничения для глубокого обучения и текущих архитектур нейронных сетей, на которые указывают ученые?

Глубокое обучение и нейронные сети достигли значительного прогресса в решении многих задач, но они имеют ряд ограничений, которые могут препятствовать достижению истинного Общего ИИ (AGI). Вот некоторые из них:

1. Необходимость больших объемов данных: Глубокое обучение обычно требует огромных объемов размеченных данных для обучения. Это может быть проблематично, поскольку разметка данных может быть трудоемкой и дорогостоящей, а в некоторых случаях необходимые данные могут быть недоступны или непригодны для использования из-за проблем с конфиденциальностью или этикой.

2. Вычислительные требования: Обучение глубоких

нейронных сетей может быть очень ресурсоемким процессом, требующим больших вычислительных мощностей и энергии. Также к таким вычислениям возникают вопросы с точки зрения экологии и охраны окружающей среды.

3. Отсутствие интерпретируемости: Глубокие нейронные сети часто описываются как «черные ящики», поскольку их решения могут быть сложными или даже невозможными для понимания человеком. Это может быть проблемой в областях, где требуется прозрачность и возможность объяснить решения, например, в таких как медицина или юриспруденция.

4. Отсутствие способности обобщения: В то время как человеческий мозг способен обобщать и переносить знания из одной области на другую, глубокие нейронные сети обычно специализируются на конкретных задачах и стремятся к буквальности, что ограничивает их способность к обобщению.

5. Отсутствие долгосрочной памяти и планирования: Большинство моделей глубокого обучения не имеют встроенной долгосрочной памяти или способности планировать на основе прошлых взаимодействий, что является ключевым аспектом человеческого интеллекта.

6. Отсутствие «здорового смысла» и контекстуально-го понимания: Нейронные сети могут быть очень хороши в определенных задачах, но они не имеют общего понимания и моделей мира или «здорового смысла», который используют люди при принятии решений. Это может приводить к ошибкам, которые кажутся очевидными для людей, но являются недоступными для «понимания» ИИ-системами.

Все эти ограничения указывают на то, что, хотя глубокое обучение и нейронные сети являются современными мощными инструментами, они в текущем состоянии вряд ли смогут одни достичь истинного AGI. Вместо этого, вероятно, потребуются новые идеи, подходы и технологии, которые могут дополнить и расширить текущие методы.

7. Расскажи о возможных будущих стадиях развития Искусственного Интеллекта? Упорядочи их по мере вероятности создания.

В будущем развитие искусственного интеллекта (ИИ) может пройти через несколько ключевых стадий, которые можно упорядочить следующим образом:

1. Cooperative AI / Augmented Intelligence: Это стадия, на которой ИИ рассматривается для работы вместе с людьми как неотъемлемый инструмент и партнер почти в любой де-

тельности. Это может включать в себя использование ИИ для помощи в принятии решений, творчестве или решении сложных проблем. Это уже происходит в некоторой степени сегодня, с ИИ, который помогает людям в задачах – от простых (например, персональные ассистенты, такие как ChatGPT, Siri или Alexa) до сложных (например, помощь врачам в диагностике заболеваний). И, вероятнее всего, эта стадия скоро будет полностью достигнута и станет доступна для очень многих людей и организаций на планете.

2. Artificial General Intelligence (AGI) / Общий ИИ:

Это следующий большой шаг в развитии ИИ, который предполагает создание систем, способных выполнять любую интеллектуальную задачу, которую может выполнить человек. Это означает, что AGI будет способен понимать, рассуждать, учиться и применять знания в широком диапазоне задач, а не только в узкоспециализированных областях, как это делают большинство современных систем ИИ. AGI может самостоятельно ставить задачи и решать проблемы, адаптироваться к новым ситуациям и понимать контекст, подобно тому, как это делают люди.

3. Artificial Super Intelligence (ASI) / Искусственный

Сверхинтеллект: Это стадия, на которой ИИ превосходит человеческий интеллект во всех значимых аспектах, включая логику и рассуждения, творчество, общее и науч-

ное знание. Такой ИИ может помочь решить самые важные проблемы человечества: создание лекарств от опасных болезней, сильное замедление старения, устранение дефицита ресурсов, предотвращение катастроф и природных бедствий, освоение космоса, решение фундаментальных научных вопросов. Но также ASI может привести к так называемой «точке сингулярности», когда технологический прогресс становится таким быстрым и сложным, что превосходит человеческое понимание и усвоение технологий.

4. Singularity AI: Это концепция, связанная с ASI, которая предполагает, что после достижения определенной точки (сингулярности), ИИ будет способен самостоятельно улучшать себя в геометрической прогрессии, что приведет к экспоненциальному росту его интеллектуальных способностей. В этот гипотетический будущий момент времени, технологический рост ИИ-систем становится неконтролируемым и необратимым, что может радикально изменить человеческое общество и жизнь на планете в целом.

Важно отметить, что эти стадии являются гипотетическими, и существует много дебатов и неопределенности относительно того, когда и как мы можем достичь AGI или ASI, если вообще сможем. Кроме того, вопросы этики, безопасности, приватности и регулирования будут играть важную роль в определении того, как эти будущие стадии ИИ будут раз-

виваться и приниматься обществом.

8. Ты часто упоминаешь вопросы этики, безопасности и контроля Искусственного Интеллекта. Почему эти вопросы важны? Какие ты видишь угрозы, опасности и негативные последствия от развития ИИ?

Вопросы этики, безопасности, приватности, контроля и регулирования искусственного интеллекта (ИИ) являются критически важными, поскольку развитие ИИ может иметь глубокие последствия для общества уже сейчас. Вот некоторые из потенциальных угроз и негативных последствий:

1. Приватность: ИИ может быть использован для сбора, анализа и использования больших объемов персональных данных, что может привести к нарушению приватности. Это может включать в себя все, от таргетированной рекламы до более серьезных вопросов, таких как слежка и наблюдение.

2. Этичность: ИИ может стать очень мощным и эффективным инструментом для создания контента любого рода, а также для усиления существующих технологических систем и продуктов. Бесконтрольное использование таких инструментов может представлять угрозы с точки зрения этики и морали.

3. Безработица: ИИ может автоматизировать многие виды работы, что может привести к значительной потере рабочих мест. Хотя это также может создать новые рабочие места, есть опасения, что переход может быть болезненным и привести к социальной нестабильности.

4. Социальное неравенство: Польза от ИИ может быть неравномерно распределена, что может усилить существующие социальные и экономические неравенства. Например, большие технологические компании могут получить несопоставимые выгоды, в то время как другие группы могут оказаться в ущербе.

5. Безопасность: Если ИИ не будет правильно контролироваться, он может представлять угрозу безопасности. Это может включать в себя все, от маленьких ошибок, которые приводят к непреднамеренным последствиям, до катастрофических сценариев, таких как использование ИИ в кибератаках или военных целях.

6. Сверхинтеллект: Если мы достигнем стадии искусственного сверхинтеллекта, это может представлять собой существенную угрозу для человечества, поскольку такой ИИ может быть непредсказуемым и потенциально враждебным для людей.

7. Сингулярность: Связанная с этим концепция «сингулярности» предполагает, что ИИ может достичь точки, когда он начинает улучшать себя быстрее, чем люди могут его понимать и контролировать. Это может привести к непредсказуемым и даже потенциально катастрофическим последствиям.

Все эти угрозы делают важным обсуждение и разработку стратегий и политик, которые могут помочь обществу разобраться в этих вопросах, чтобы извлечь максимальную пользу от ИИ, минимизируя его потенциальные риски.

9. Расскажи подробнее о том, какие есть подходы и направления к созданию безопасного и полезного для людей Искусственного Интеллекта?

Создание безопасного и полезного для людей искусственного интеллекта (ИИ) является сложной задачей, которая требует учета многих факторов.

Вот некоторые из подходов и направлений, которые исследуются в этой области:

1. Прозрачность и интерпретируемость: Один из подходов к созданию безопасного ИИ – это создание си-

стем, которые можно легко понять и интерпретировать. Это может помочь людям лучше понять, как ИИ принимает решения, и обнаружить любые потенциальные проблемы или смещения.

2. Принципы этики и правил: Другой подход – это внедрение этических принципов и правил в процесс разработки ИИ. Это может включать в себя учет ценностей, таких как справедливость, приватность и автономия, а также создание механизмов для обеспечения соблюдения этих принципов.

3. Безопасность и соответствие человеческим ценностям: Это область, которая занимается исследованием и разработкой методов для обеспечения безопасности ИИ и вопросами о том, как обеспечить, чтобы цели и поведение ИИ соответствовали целям и ценностям людей. Цель состоит в том, чтобы создать ИИ, который действует так, как мы хотим, и не представляет угрозы для людей или общества.

4. Регулирование и надзор: Регулирование и надзор также могут играть важную роль в обеспечении безопасности ИИ. Это может включать в себя создание законов и нормативов, которые регулируют использование ИИ, а также механизмов для проверки соблюдения этих правил.

5. Открытость и сотрудничество: Открытость и сотрудничество между исследователями, организациями и странами могут помочь обеспечить безопасность ИИ. Это может включать в себя обмен идеями, данными и лучшими практиками, а также совместную работу над решением общих проблем.

6. Вовлечение общественности: Вовлечение общественности в дискуссии о безопасности ИИ может помочь обеспечить, что различные голоса и перспективы учитываются, и что развитие ИИ отвечает интересам и ценностям общества в целом.

7. Долгосрочное мышление и планирование: Учитывая потенциальные долгосрочные последствия развития ИИ, важно включить долгосрочное мышление и планирование в стратегии безопасности. Это может включать в себя анализ сценариев, прогнозирование и разработку стратегий для справления с возможными будущими рисками и проблемами.

Все эти подходы могут играть важную роль в создании безопасного и полезного ИИ. И важно отметить, что нет одного универсального метода или способа решения этих задач – вероятно, потребуется комбинация многих различных стратегий и подходов. Поиск и выработка таких подходов яв-

ляется одной из самых важных задач на сегодняшний день.

10. Дай мне список книг, которые рассказывают об истории развития, текущем положении, угрозах, возможностях, воздействии на общество и будущих трендах в области Искусственного Интеллекта?

Вот несколько книг, которые я считаю лучшими для понимания истории, текущего положения, воздействия на общество и будущих трендов в области искусственного интеллекта:

1. «Artificial Intelligence: A Guide for Thinking Humans» – Мелани Митчелл.

Эта книга представляет собой доступное введение в ИИ для непрофессионалов. Митчелл рассказывает историю становления области, объясняет основные концепции ИИ и обсуждает его возможные последствия для общества. Эта книга хороша, потому что она написана простым и увлекательным языком и представляет собой прекрасное введение в ИИ и его историю для тех, кто не имеет технического или математического образования.

2. «Rebooting AI» – Гэри Маркус и Эрнест Дэвис.

Ученые, специализирующиеся на искусственном интеллекте, рассказывают о том, почему, несмотря на всю шумиху, мы еще очень далеки от создания интеллекта, подобного человеческому, а также о том, что нужно изменить в существующих подходах к созданию ИИ. Вдохновленные человеческим разумом, Маркус и Дэвис объясняют, что же нам нужно для продвижения ИИ на новый уровень. Если мы сосредоточимся на том, чтобы наделить машины здравым смыслом и глубоким пониманием, а не просто научить их собирать все более обширные коллекции данных, то сможем создать искусственный интеллект, которому действительно будем доверять то, что имеет для нас значение: наши дела, заботу о здоровье, наши дома и жизнь.

3. «AI Superpowers: China, Silicon Valley, and the New World Order» – Кай-Фу Ли.

Книга о перспективах искусственного интеллекта, написанная одним из главных экспертов по теме ИИ. Кай-Фу Ли – ученый, венчурный инвестор и человек с головокружительной карьерой (Apple, Silicon Graphics, Microsoft, Google). Он утверждает, что драматические изменения, которых все ждут от искусственного интеллекта, наступят скорее, чем предполагается. Но связаны они будут не с «думающими машинами, которые захватят мир», а с экономическими последствиями. Гонка между сверхдержавами, Китаем и США,

может привести человечество к катастрофе: ИИ позволит богатым разбогатеть еще больше, а разрыв между ними и бедными станет огромным.

4. **«The Age of A.I.»** – Генри Киссинджер, Эрик Шмидт, и Даниэль Хаттенлохер.

Три известных мыслителя обсуждают, как искусственный интеллект трансформирует общество и что это значит для всех нас. Эта книга хороша, потому что она представляет собой глубокое исследование влияния ИИ на общество от людей, которые имеют большой опыт в этой области.

5. **«Life 3.0: Being Human in the Age of Artificial Intelligence»** – Макс Тегмарк.

Тегмарк обсуждает, как ИИ может изменить все аспекты нашей жизни, от работы и образования до политики и правосудия. Данная книга предлагает широкий обзор возможных последствий развития ИИ и рисует мир возможного будущего.

6. **«Superintelligence: Paths, Dangers, Strategies»** – Ник Бостром.

Эта книга исследует возможные последствия появления

ИИ, который превосходит человеческий интеллект. Бостром обсуждает различные пути к Суперинтеллекту и возможные опасности. Эта книга хороша, потому что она ставит важные вопросы о будущем ИИ и его влиянии на общество.

7. «The Singularity Is Near: When Humans Transcend Biology» – Рэй Курцвейл.

В этой книге знаменитый футуролог Рэй Курцвейл предсказывает, что технологическая сингулярность, момент, когда ИИ превзойдет человеческий интеллект, наступит в ближайшие десятилетия. Эта книга является одной из самых известных и дискуссионных работ о будущем ИИ.

8. «Surviving AI: The Promise and Peril of Artificial Intelligence» – Кэлам Чейс.

Автор обсуждает обещания и опасности ИИ, включая возможность достижения Суперинтеллекта и его последствий для человечества. Эта книга хороша, потому что она представляет собой доступное и информативное введение в важные вопросы, связанные с ИИ.

9. «Human Compatible» – Стюарт Расселл.

Это научно-популярная книга известного в области ИИ

ученого и автора Стюарта Расселла, которая посвящена таким вопросам как:

- что будет, когда мы достигнем создания Сильного ИИ?
- как сделать, чтобы это не стало последним событием нашей истории?
- как жить в будущем, где способности машин будут превосходить человеческие?

В книге утверждается, что риск для человечества от новых ИИ-разработок является серьезной проблемой. Автор указывает на неопределенность, связанную с будущим прогрессом в области ИИ. Рассел утверждает, что именно потому, что сроки разработки искусственного интеллекта человеческого уровня или сверхразума весьма неопределенны, исследования безопасности следует начинать как можно скорее, поскольку также крайне неопределенно, сколько времени потребуется для завершения таких исследований. В книге предлагаются подходы к решению проблемы управления и создания безопасного ИИ.

Глава 3. Искусственный Интеллект и Машинное обучение

1. Расскажи про классические методы ИИ, которые были еще до расцвета машинного обучения и нейросетей – GOFAI. Применяются ли они сейчас? Есть ли у них будущее?

Методы «**Good Old-Fashioned Artificial Intelligence**» (GOFAI) или «классический искусственный интеллект, основанный на правилах» – это термин, введенный философом Джоном Хогеландом в 1980-х годах, чтобы описать ранние подходы к искусственному интеллекту, которые были развиты до бума Машинного Обучения и Нейронных сетей.

Обычно выделяют следующие группы методов, относящиеся к GOFAI:

1. Экспертные системы – компьютерные системы, которые воссоздают процесс принятия решений экспертом в конкретной области. Они функционируют на основе базы знаний, состоящей из фактов и правил, и способны имитировать решения, принятые специалистами в определённой

сфере.

2. Семантические сети – структуры, предназначенные для представления знаний в виде графов (узлов и связей между ними). В них узлы обозначают понятия и концепции, а связи (или дуги) между узлами отражают взаимоотношения между этими понятиями и концепциями.

3. Логический подход и символьные рассуждения – связаны с использованием формальной логики и вербальных описаний для представления и манипулирования знаниями. Такие методы активно применяются в системах, где знания заранее четко структурированы в виде фактов, правил и взаимоотношений, к которым можно применять вывод на основе логических правил.

4. Продукционные системы – программы, которые работают по определенному набору правил или «продукций» («если-то» правил), чтобы получить выводы. Эти системы могут применяться для автоматизации задач рассуждения и проблемы разрешения – через последовательности применений этих «если-то» -правил.

5. Эвристические алгоритмы – это способы решения задач, не полностью математически обоснованные, но обычно дающие достаточно эффективные решения в большин-

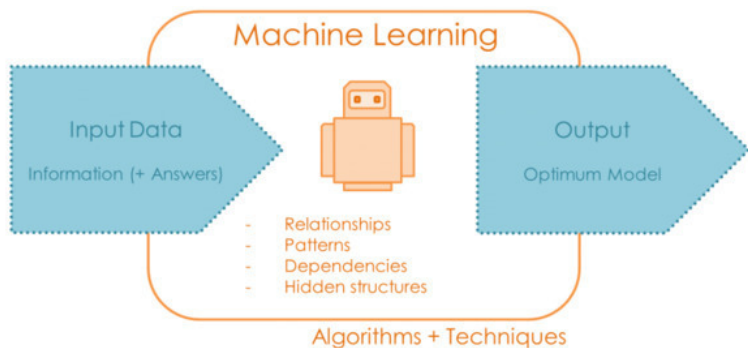
стве случаев. В общем смысле, эвристика – это подход, который может быть не совсем корректным с математической точки зрения, но при этом оказывается практически полезным для решения интеллектуальной задачи.

С развитием Машинного Обучения и Нейронных Сетей, многие из этих подходов отошли на второй план, поскольку новые методы часто обеспечивают более высокую производительность и более широкую область применения. Тем не менее, классические методы AI по-прежнему имеют свою ценность в определенных областях. Они могут быть полезны в ситуациях, где доступен только небольшой объем данных, или когда требуется явное, логическое обоснование выводов. Более того, в некоторых случаях классические методы AI и машинное обучение могут быть сочетаны в гибридных системах для получения лучших результатов.

2. Что такое Машинное Обучение? Какие у него основные принципы и идеи? Что такое Модель Машинного Обучения?

Машинное обучение – это подраздел Искусственного Интеллекта, который занимается разработкой алгоритмов, позволяющих компьютерам самостоятельно учиться на основе предоставленных данных (без участия человека в процессе). Суть Машинного Обучения заключается в том, что

машина сама (без явного программирования каждого действия человеком) может решать поставленные ей задачи, «обучаться» и улучшать свою работу, просто анализируя предоставленные ей входные данные.



Модель машинного обучения – это математическое представление (формула или алгоритм), который по входным данным самостоятельно предсказывает выходные данные. Модели машинного обучения могут быть простыми (например, формула, которая предсказывает вероятный ответ по входным данным) или сложными (например, глубокие нейронные сети с большим количеством параметров, которые используются для моделирования сложных взаимосвязей между данными).

Основные этапы процесса (цикла работы) машин-

ного обучения:

1. Обучение модели на основе входных данных (примеров): Машинное обучение использует наборы данных (примеров), которые состоят из «входных» данных и связанных с ними «выходных» данных (иногда называемых «метками» или «ответами»). Модель обучается устанавливать взаимосвязи между этими входными и выходными данными.

2. Обобщение знания из примеров и решение задачи: Это способность модели использовать обучение на основе примеров для предсказания ответов на новые, ранее не виденные данные. Цель обучения модели – не просто «запомнить» примеры из обучающего набора данных, но и «обобщить» эти примеры на новые ситуации – то есть попытаться применить это знание на новых примерах.

3. Оптимизация и улучшение модели: Обучение модели обычно включает в себя процесс оптимизации, где с каждым циклом алгоритм постепенно улучшает свою производительность, минимизируя разницу между предсказанными и действительными ответами (то есть минимизирует ошибку своего предсказания).

4. Итоговое тестирование и оценка модели: Модель

обучается на основе обучающих данных и затем тестируется на отдельном тестовом наборе данных для проверки ее способности обобщения и предсказания правильных ответов.

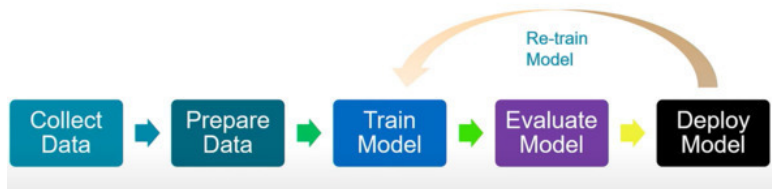
В зависимости от того, **как именно обрабатываются данные и решается поставленная задача** – Модель машинного обучения может быть отнесена к одному из **Типов (видов) Машинного обучения**:

- Supervised Learning («обучение с учителем»)
- Unsupervised Learning («обучение без учителя»)
- Semi-supervised Learning («обучение с частичным привлечением учителя»)
- Self-supervised Learning («самостоятельное/автоматическое обучение»)
- Reinforcement Learning («обучение с подкреплением»)

В зависимости от того, **какая именно решается поставленная задача** — можно выделить один из **Типов (видов) задач Машинного обучения**:

- Классификация объектов
 - Кластеризация объектов
 - Регрессия (предсказание непрерывного знания)
 - Генерация (создание новых объектов)
- и другие.

3. Расскажи мне подробнее про основные шаги и процессы в Машинном Обучении – про работу с данными, выделение признаков, алгоритмы?



Работа с машинным обучением обычно включает в себя следующие основные шаги:

1. Постановка задачи: Перед началом проекта машинного обучения важно изучить предметную область и понять задачу и бизнес-цели, точно определить, какая проблема или задача будет решаться с помощью машинного обучения, а также определить метрики качества, которые будут служить ориентирами успеха.

2. Сбор данных: В первую очередь, вам нужен большой набор данных, на котором вы будете обучать свою модель. Этот набор данных может быть собран из различных источников, таких как файлы, базы данных, логи с действиями

пользователей, интернет и так далее.

3. Предварительная обработка данных: Этот этап включает в себя очистку данных (удаление ошибок, сильных отклонений, пропущенных значений), преобразование данных (например, преобразование текста в числовые значения или приведение разнородных данных к единому формату) и нормализацию данных (например, масштабирование значений на определенный диапазон).

4. Выделение признаков: Признаки – это характеристики или атрибуты, которые машина использует для обучения. Например, если вы создаете модель для классификации изображений кошек и собак, признаками могут быть размеры животных, цвета, текстуры и формы, присутствующие на изображении. Этап выделения признаков включает в себя выбор и создание эффективных признаков, которые помогут модели делать более точные прогнозы.

5. Выбор модели: В зависимости от типа проблемы и задачи (регрессия, классификация, кластеризация и т.д.) и специфики данных, вы выбираете тип (вид) Машинного обучения (Supervised, Unsupervised, Semi-supervised, Self-supervised или Reinforcement Learning) и конкретную подходящую Модель машинного обучения этого вида.

6. Обучение модели: На этом этапе алгоритм машинного обучения сам «обучает» модель, используя ваши данные и целевые значения. Это происходит путем настройки параметров модели таким образом, чтобы минимизировать ошибку между прогнозируемыми моделью результатами и реальными значениями результатов (взятыми из обучающих примеров).

7. Оценка модели: После обучения модели вам нужно оценить ее качество и производительность. Это обычно делается с помощью сравнения ответов модели с отложенным набором данных (тестовым набором), который не использовался при обучении. Метрики оценки могут включать точность, полноту и другие показатели работы модели. Важно получить модель, которая не будет переобученной, но и не будет недообученной – чтобы получать от нее потом хорошие результаты предсказаний

8. Тонкая настройка и оптимизация: После первоначального обучения и оценки модели вы можете оптимизировать и настраивать свою модель, изменяя параметры и используя различные техники, такие как кросс-валидация и регуляризация.

9. Развертывание модели (Деплой): После того, как модель была обучена, оценена и оптимизирована, она мо-

жет быть «развернута» (на компьютерных системах и вычислительных мощностях) и использована для предсказаний на новых данных.

10. Процесс работы модели (Инференс): Обычно этот термин используют при работе с нейронными сетями. Инференсом называется непрерывная работа какой-либо нейронной сети на конечном устройстве. То есть, это процесс исполнения сети, когда она уже развернута и готова к проведению полезной работы. Для инференса используются процессоры общего назначения (CPU), графические процессоры (GPU) или специализированные процессоры для Машинного обучения и нейросетей (TPU).

Все эти шаги могут потребовать различных навыков и инструментов: знания основ математики, статистики, программирования, поддержки работы компьютерных систем, обработки и анализа данных и, конечно же, знания самого машинного обучения и предметной области, в которой вы решаете задачу.

4. Расскажи про основные параметры, определяющие качество и эффективность моделей Машинного Обучения?

Машинное обучение – это процесс, в ходе которого

компьютерные модели «учатся» на данных и делают свои прогнозы или решения на основе этого обучения.

При создании моделей машинного обучения одним из самых важных этапов является оценка их работы. Без правильной оценки результатов есть риск начать использовать модель, которая может давать неверные прогнозы, принимать неправильные решения, пропускать важные случаи (в задачах выявления нужных объектов).

Чтобы узнать, насколько хорошо модель справляется со своей задачей – используют **метрики качества моделей** машинного обучения. Оценка моделей не только позволяет понять их эффективность, но и выявить возможные недостатки, которые стоит устранить.

Вот **примеры метрик качества** для моделей в машинном обучении:

Средняя абсолютная ошибка – Для задач, где модель предсказывает численные значения, эта метрика показывает, насколько в среднем прогнозы модели отличаются от истинных значений. Например, если модели нужно предсказывать температуру воздуха в течение какого времени, эта метрика покажет на сколько в среднем отклоняются предсказания модели (неважно – в большую или в меньшую стороны)

от реальной температуры воздуха. Чем меньше отклонения – тем лучше модель.

Точность модели – Для задач, где нужно выбрать определенный тип объектов в общей выборке и не ошибаться с типом этих объектов (но можно что-то и пропустить), эта мера показывает, какой процент прогнозов модели был правильным. Например, нужно определить и выбрать клиентов, которые с большей вероятностью купят определенный товар. Так, если модель правильно предсказала 85 из 100 случаев, то её точность составляет 85%.

Полнота модели – Когда модели нужно обнаружить определенный тип объектов в общей выборке и важно не пропустить объекты этого типа (но при этом, допускается ошибаться и обнаружить лишние объекты, которые на самом деле не относятся к искомому типу). Эта метрика показывает, сколько случаев из выборки модель учла. Например, когда модель применяется в медицине и ей нельзя пропустить больных с определенным диагнозом (так как в этом случае человек не узнает о своем диагнозе и не получит вовремя соответствующего лечения), но вполне можно предположить наличие болезни у здорового человека (так как потом его перепроверят и снимут ошибочно поставленный диагноз).

Перплексия (Perplexity) – это популярная метрика для

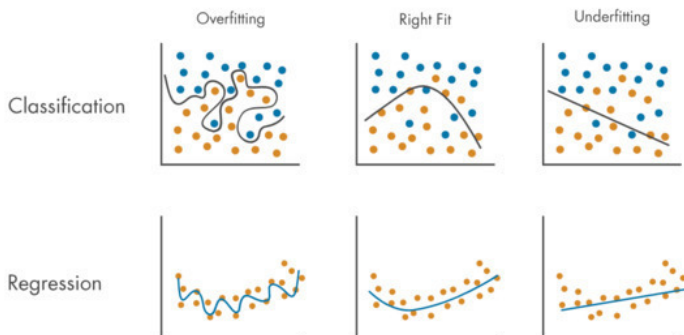
оценки качества языковых моделей в задачах прогнозирования следующего слова. Перплексия интерпретируется как среднее число выборов, которые модель рассматривает при предсказании следующего слова. Например, перплексия, равная 10, означает, что при прогнозировании следующего слова модель в среднем «колеблется» между 10 словами. Меньшее значение перплексии указывает на лучшую модель. Модель с перплексией 1 была бы идеальной и всегда бы правильно предсказывала (знала вполне определенно) следующее слово.

Кроме этого, один из основных вызовов в машинном обучении – избежать недообучения и переобучения моделей.

Недообучение модели (Underfitting): Это происходит, когда модель слишком проста для сложности данных, и не может выучить закономерности в них. В результате такая модель плохо справляется как с обучающими, так и с тестовыми данными.

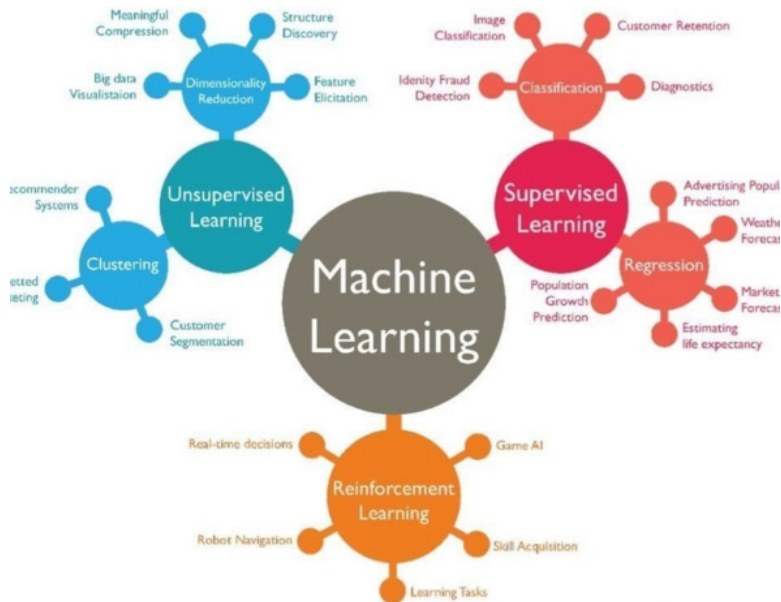
Переобучение модели (Overfitting): Здесь проблема противоположная. Модель становится слишком «узкоспециализированной» под обучающие данные и начинает «запоминать» их, но уже не может «понять» более общую закономерность и предсказать то, чего не было в обучающих данных. Поэтому переобученная модель может идеально рабо-

тать на обучающем наборе данных, но плохо на новых-тестовых данных.



Оценка качества моделей машинного обучения – ключевой этап в процессе их создания. Выбор правильной метрики поможет понять, насколько получена правильная и эффективна модель, и, при необходимости, внести коррективы в её обучение или заменить модель.

5. Расскажи, какие вообще есть Типы (виды) Машинного Обучения?



Изначально в Машинном Обучении выделяли три основных типа: Обучение с учителем (Supervised Learning), Обучение без учителя (Unsupervised Learning) и Обучение с подкреплением (Reinforcement Learning). Со временем еще 2 типа получили развитие: Обучение с частичным привлечением учителя (Semi-supervised Learning) и Самостоятельное/автоматическое обучение (Self-supervised Learning).

Рассмотрим все эти **Виды Машинного Обучения:**

1. Обучение с учителем (Supervised Learning)

Модели на вход даются примеры размеченных данных, где каждый пример уже помечен правильным ответом. Задача модели – научиться самой также предсказывать эти ответы для данных. Например, если мы обучаем модель распознавать кошек и собак на фотографиях, то сначала мы показываем ей множество изображений кошек и собак, где каждое изображение уже с соответствующей меткой («кошка» это или «собака»). Или, если модель учится отличать спам-письма от не спама, то на входе ей дается множество писем с имеющимися метками – спам это или не спам. Модель пытается предсказать ответы для примеров с уже известными метками и каждую итерацию сравнивает свои ответы с существующими ответами и пытается улучшить себя, чтобы на следующей итерации предсказывать ответы еще точнее. Итоговая задача модели – найти зависимость между данными и метками классов и использовать эту зависимость для дальнейшего самостоятельного предсказания классов для новых (неразмеченных) входных данных.

2. Обучение без учителя (Unsupervised Learning)

Здесь модели обучаются на наборе данных, в котором известны только неразмеченные входные данные, и нет кон-

кретных выходных данных (меток классов и т.п.). Задача модели состоит в том, чтобы самостоятельно найти структуру или взаимосвязи в данных. Примеры включают кластеризацию (например, сегментацию клиентов для маркетинга по разным подгруппам или определение количества разных видов подгрупп в очень большой группе людей) и понижение размерности (например, упрощение данных для их понятного представления и визуализации).

3. Обучение с подкреплением (Reinforcement Learning)

В этом случае модель (часто в этом случае ее называют ИИ-агентом) обучается сама, взаимодействуя со своей «средой обитания». Модель выполняет различные действия и в результате этих действий получает от среды отклик – награды («подкрепление») за правильные (полезные или эффективные) действия или штрафы за неправильные (вредные или не эффективные) действия. Модель стремится максимизировать сумму получаемых наград – то есть выполнять действия (или последовательности действий), которые дают как можно лучший результат. Примеры включают управление роботами (которые получают отклик от среды – плохо или хорошо они выполняют свои задачи), игровые агенты (получающие отклик от игровой среды – в случае выигрыша или проигрыша), системы рекомендаций (где отклик –

это качество удовлетворения пользователей этими рекомендациями).

4. Обучение с частичным привлечением учителя (Semi-supervised Learning)

Машинное обучение с частичным привлечением учителя (также известное как обучение с полуконтролем или гибридное обучение), находится между Обучением с учителем (Supervised Learning) и Обучением без учителя (Unsupervised Learning). В этом случае модели подается комбинация помеченных и непомеченных данных. Непомеченные данные очень дешевы в отличие от помеченных данных (которые часто приходится помечать вручную). Процедура заключается в том, что алгоритм сначала использует все данные и алгоритмы обучения без учителя для кластеризации данных, а затем использует алгоритм обучения с учителем для определения меток для каждого класса. И если непомеченные данные оказываются близки к одному из классов помеченных — то они с большей вероятностью принадлежат тому же классу.

5. Самостоятельное/автоматическое обучение (Self-supervised Learning)

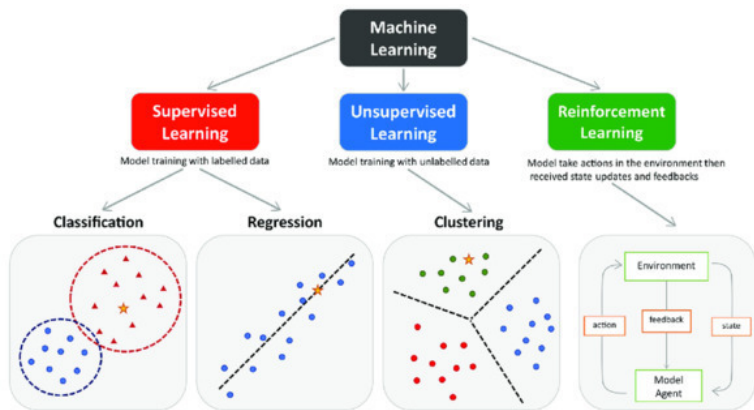
Это относительно новый подход, где модель обучается

на данных, генерируя сама себе задачи и ответы (чаще всего через маскировку части данных и попытки их угадывания). Задача модели – понять и усвоить скрытую структуру, которая есть в этих данных. Например, если данные – это связный текст, то модель может маскировать и пытаться предсказать следующее слово в предложении. Или для изображений модель может пытаться маскировать и восстанавливать части изображения. Или попытаться предсказать цвет или другие параметры изображения. Таким образом, она учится на большом количестве данных (текстов, изображений, видео), без необходимости наличия внешних меток или участия человека. В результате модель усваивает структуру и внутренние связи в этих данных (которые могут быть даже неизвестны человеку). И когда эта структура усвоена – то модель можно дообучить для решения какой-нибудь специальной практической задачи (для решения которой нужно понимать эту структуру в данных). Например, можно дообучить модель для автозаполнения или перевода текстов или сделать модель для улучшения, окраски и восстановления изображений и т. д. Бурный рост приложений в области современного Генеративного ИИ (Generative AI) – больше всего обязан именно этому типу Машинного Обучения.

Каждый из этих методов обладает своими уникальными преимуществами и ограничениями и выбор метода зависит от конкретной задачи и доступности данных.

6. Какие основные виды задач решает машинное обучение? Дай определение сути и характеристику для каждого вида задач.

Машинное обучение используется для решения различных видов задач.



Вот основные виды задач, которые есть в современном машинном обучении:

1. Задачи классификации: В задачах классификации модель предсказывает дискретную метку или категорию. Например, задача определения, является ли электронное

письмо спамом или нет, является задачей классификации. Здесь мы предсказываем дискретную переменную (спам или не спам).

2. Задачи кластеризации: Это тип задач, где модель группирует данные на основе их сходства, и эти группы называются кластерами. Этот процесс происходит без каких-либо предварительных знаний о данных, и в этом смысле он относится к обучению без учителя. Например, кластеризация может быть использована для сегментации клиентов на основе их покупательского поведения.

3. Задачи регрессии: Регрессия – это тип задачи, где модель предсказывает непрерывное значение. Например, предсказание цены на дом на основе различных характеристик, таких как площадь, количество спален, год постройки и т.д., является задачей регрессии. В этом случае, мы пытаемся предсказать непрерывную переменную (цена на дом) на основе других входных данных об этом доме.

4. Задачи Обучения с подкреплением (Reinforcement Learning): Сюда можно отнести примеры с управлением роботами (которые получают отклик от среды – плохо или хорошо они выполняют свои задачи), развитием навыков игровых агентов (получающие отклик от игровой среды – в случае выигрыша или проигрыша), систем рекомендаций (где

отклик – это качество удовлетворения пользователей этими рекомендациями).

5. Задачи Генеративного ИИ: В отличие от задач классического машинного обучения (классификации, кластеризации и регрессии), Генеративные модели обучаются на данных и могут генерировать новые, ранее не встречавшиеся образцы данных. Данные могут представлять собой текст, изображения, речь и т. д. Задачи, которые могут выполнять такие модели, включают создание разнообразного контента: текстов, изображений, звука и музыки и т. д. Кроме этого, модели генеративного ИИ могут выполнять широкий класс задач, связанных с дальнейшей обработкой и преобразованием этого контента: ответы на вопросы, анализ настроений и тональности в текстах или видео; извлечение искомой информации из текста изображений, видео или аудио; маркировку изображений и распознавание объектов.

Конец ознакомительного фрагмента.

Текст предоставлен ООО «Литрес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на Литрес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.