

КАРИНА ПОМАЗАН

КАК МЫСЛИТ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

— И ПОЧЕМУ ОШИБАЕТСЯ —

СОЗНАНИЕ
ТВОРЧЕСТВО
ЧЕЛОВЕК

АЛГОРИТМЫ
ЛОГИКА
ДАННЫЕ

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$E = mc^2$$

КАК УСТРОЕН ИИ • ПОЧЕМУ ОН ОШИБАЕТСЯ •
ЧТО ДЕЛАЕТ НАС УНИКАЛЬНЫМИ •
БУДУЩЕЕ ЧЕЛОВЕКА И ТЕХНОЛОГИЙ

КНИГА
О БУДУЩЕМ,
КОТОРОЕ МЫ
СОЗДАЕМ
СЕГОДНЯ

Карина Помазан

Как мыслит искусственный интеллект и почему ошибается

<https://litres.ru/74209827>

SelfPub; 2026

Аннотация

Что происходит внутри искусственного интеллекта, когда он отвечает на ваши вопросы? Почему он способен создавать картины, писать книги, анализировать сложные задачи — и при этом иногда совершает удивительно простые ошибки?

Эта книга приглашает читателя в увлекательное путешествие по миру искусственного интеллекта — без сложных технических терминов, но с глубоким пониманием того, как работают современные интеллектуальные системы.

Вы узнаете:

- как «мыслит» искусственный интеллект и чем его мышление отличается от человеческого;
- почему ИИ может ошибаться, придумывать несуществующие факты и делать неверные выводы;
- что такое нейронные сети, большие языковые модели и машинное обучение;

способен ли искусственный интеллект творить, испытывать эмоции и обладать сознанием;

- как ИИ уже меняет искусство, психологию, образование, бизнес и повседневную жизнь;

Содержание

Глава 1. Самое большое заблуждение об искусственном интеллекте	8
Нам кажется, что по другую сторону экрана кто-то думает	10
Попробуем представить другую картину	11
Представьте пианиста	13
Почему это трудно принять	14
Иллюзия понимания	15
Почему ответы кажутся такими естественными	16
Главная идея этой книги	17
Глава 2. Почему люди захотели создать мыслящую машину	18
Мечта, которая старше электричества	19
От силы мышц к силе разума	20
Что значит «думать»?	21
Почему первые ожидания были слишком оптимистичными	22
Почему ребенок учится иначе	23
Два пути к искусственному интеллекту	24
Почему это похоже на обучение музыке	26
Почему ИИ — это не электронный человек	27
Настоящий вопрос	28
Глава 3. От механических автоматов до	29

нейросетей	
Первая мечта	30
Первые настоящие автоматы	31
Машина, которая всех обманула	32
Когда появились компьютеры	33
Первая ошибка ученых	34
Мир оказался слишком сложным	35
Новая идея	36
Почему это стало революцией	37
Почему нейросети называют нейросетями	38
Самый неожиданный поворот	39
История продолжается	40
Глава 4. Почему ИИ — это не человеческий мозг	41
Представьте самолет	42
Мы часто путаем результат со способом	43
Как думает человек	44
А что происходит у ИИ?	46
Человек постоянно проверяет мир	47
Наш мозг экономит силы	48
Почему ИИ иногда кажется живым	49
Что умеет человек, но пока не умеет ИИ	50
Тогда почему ИИ настолько полезен?	51
Главная мысль этой главы	52
Глава 5. Что значит «обучение» для ИИ	53
Обучение человека и обучение машины	54

Учитель для ИИ выглядит иначе	56
Представьте пианино	57
ИИ не учится как ребенок	58
Почему тогда ИИ знает так много?	59
Библиотека без читателя	60
Что именно запоминает модель?	61
Почему обучение требует огромных ресурсов	62
После обучения начинается другая жизнь	63
Самая большая загадка	64
Глава 6. Как ИИ превращает текст в числа	65
Компьютер не видит слова	66
Что такое токены	67
Представьте конструктор	68
Но где здесь смысл?	69
Карта смыслов	71
Человек видит мир, ИИ видит связи	72
Почему это всё равно работает	73
Аналогия с художником	74
Почему перевод текста в числа — это революция	75
Самый важный вывод	76
Глава 7. Мир токенов: как ИИ видит язык	77
Слово для человека и слово для ИИ	79
Язык как огромная сеть	81
Почему одно слово может иметь разные значения	83
Контекст меняет всё	84

Как слова превращаются в координаты	85
Почему метафоры работают	86
Но где граница?	87
Почему иногда ИИ создает удивительные идеи	88
Опасность красивых связей	89
Главная идея главы	90
Глава 8. Почему слова становятся математикой	91
Числа тоже могут описывать мир	92
Представьте библиотеку без алфавита	93
Слова получают «координаты смысла»	94
Аналогия с художником	95
Почему слово связано с другими словами	96
Самое удивительное: отношения важнее самих слов	97
Конец ознакомительного фрагмента.	98

Как мыслит искусственный интеллект и почему ошибается

Глава 1. Самое большое заблуждение об искусственном интеллекте

Представьте, что вы открываете искусственный интеллект и пишете:

«Придумай сказку о девочке, которая нашла дверь в другой мир.»

Через несколько секунд появляется красивая история. Герои, диалоги, неожиданный финал. Кажется, будто по ту сторону экрана сидит настоящий рассказчик, который на ходу придумывает сюжет.

На следующий день вы просите:

«Объясни мне квантовую физику так, чтобы понял десятилетний ребенок.»

И снова получаете понятный ответ.

Потом спрашиваете:

«Помоги написать письмо директору.»

ИИ помогает и с этим.

Через неделю он сочиняет стихотворение, переводит текст, придумывает название книги, помогает с кодом и даже поддерживает разговор.

Невольно возникает мысль:

«Он действительно понимает, о чем говорит».

Именно здесь рождается самое большое заблуждение нашего времени.

Нам кажется, что по другую сторону экрана кто-то думает

Человеческий мозг устроен удивительно.

Мы постоянно ищем разум вокруг себя.

Мы разговариваем с автомобилем:

— Ну давай же, заводишься!

Мы ругаем компьютер:

— Почему ты опять завис?

Мы благодарим навигатор.

Мы можем даже извиниться перед роботом-пылесосом, случайно ударив его ногой.

Так работает наше мышление.

Если что-то ведет себя достаточно сложно, мы начинаем видеть в этом личность.

С искусственным интеллектом происходит то же самое.

Он отвечает настолько естественно, что мозг автоматически делает вывод:

«Если он разговаривает как человек — значит, он и думает как человек.»

Но это не так.

Попробуем представить другую картину

Вообразите огромную библиотеку.

Не просто большую.

Самую большую библиотеку в истории человечества.

В ней находятся миллиарды книг.

Романы.

Газеты.

Научные статьи.

Форумы.

Энциклопедии.

Сценарии фильмов.

Учебники.

Поэзия.

Разговоры людей.

Рецепты.

Истории.

Все это не разложено по полкам.

Это превращено в невероятно сложную карту языка.

Когда вы задаете вопрос, ИИ не идет искать готовый ответ

в одной книге.

Он не вспоминает страницу.

Он не открывает энциклопедию.

Вместо этого он пытается продолжить ваш текст наиболее подходящим образом, опираясь на закономерности, которые усвоил во время обучения.

Это очень важное различие.

Представьте пианиста

Допустим, человек с детства слушал музыку.

Двадцать лет.

Каждый день.

Он знает тысячи произведений.

Потом вы просите его:

— Сыграй что-нибудь в стиле Шопена.

Он не вспоминает одну конкретную пьесу.

Он начинает создавать новую музыку, используя огромный опыт.

ИИ делает нечто похожее с языком.

Только вместо нот у него слова.

Вместо мелодий — предложения.

Вместо музыкальных произведений — тексты.

Но здесь аналогия заканчивается.

Пианист переживает музыку.

Он чувствует настроение.

Он понимает, что хочет выразить.

Современный ИИ не обладает таким внутренним опытом.

Он не испытывает эмоций и не осознает смысл так, как это делает человек. Он строит ответ, опираясь на статистические закономерности языка.

Почему это трудно принять

Потому что язык всегда был признаком разума.

На протяжении тысячелетий только человек мог:

рассказывать истории;

спорить;

объяснять;

шутить;

писать книги.

Теперь это умеет делать машина.

Поэтому нам кажется, что она должна обладать тем же внутренним миром.

Но способность создавать убедительный текст сама по себе еще не доказывает наличие сознания или человеческого понимания.

Иллюзия понимания

Представьте актера.

Он играет врача.

Во время фильма он произносит медицинские термины.

Уверенно ставит диагноз.

Объясняет лечение.

Для зрителя он выглядит настоящим специалистом.

Но после съемок выясняется, что актер никогда не изучал медицину.

Он великолепно сыграл роль.

ИИ тоже способен создавать впечатление глубокого понимания.

Но впечатление — не доказательство.

Это одна из главных идей всей книги.

Почему ответы кажутся такими естественными

Потому что ИИ учился на огромном количестве человеческих текстов.

За годы обучения он «увидел» бесчисленное количество примеров того, как люди:

задают вопросы;

спорят;

рассказывают истории;

пишут научные статьи;

объясняют сложные идеи;

признают ошибки;

убеждают друг друга.

Из этих примеров модель научилась воспроизводить характерные закономерности языка.

Она не копирует готовые тексты, а создает новые ответы, комбинируя усвоенные закономерности.

Главная идея этой книги

Большинство людей задают неправильный вопрос:

«Умеет ли ИИ думать?»

Но гораздо полезнее спросить иначе:

«Как именно ИИ создает ответы, которые кажутся результатом мышления?»

Это совершенно другой вопрос.

Именно ему посвящена эта книга.

Мы шаг за шагом разберем:

почему ИИ иногда отвечает блестяще;

почему иногда уверенно ошибается;

почему один вопрос может дать совершенно разные ответы;

почему ИИ может написать роман, но ошибиться в простом факте;

почему он не устает, но не обладает человеческим опытом;

и почему понимание принципов его работы делает нас не только более грамотными пользователями технологий, но и помогает лучше понять особенности собственного мышления.

Глава 2. Почему люди захотели создать мыслящую машину

Когда люди впервые услышали об искусственном интеллекте, многим показалось, что это технология XXI века. Но на самом деле мечта о мыслящей машине появилась задолго до первых компьютеров.

Чтобы понять, откуда взялся ИИ, нужно отправиться на тысячи лет назад.

Мечта, которая старше электричества

Задолго до появления программирования люди рассказывали истории о существах, созданных человеческими руками.

В древних мифах встречались оживающие статуи, механические слуги и искусственные создания, способные выполнять поручения. Эти истории не были описанием реальных технологий. Они отражали древнее желание человека: создать помощника, который мог бы действовать самостоятельно.

Позже, в эпоху Возрождения и Нового времени, появились сложные механические автоматы — устройства, которые могли двигаться, играть на музыкальных инструментах или имитировать действия человека. Они не думали, но производили сильное впечатление.

Люди снова задавались вопросом:

«А что, если однажды машина сможет не только двигаться, но и рассуждать?»

От силы мышц к силе разума

Каждая крупная технология освобождала человека от какого-то вида труда.

Плуг уменьшил физические усилия земледельца.

Паровоз заменил долгие пешие путешествия.

Автомобиль сократил расстояния.

Подъемный кран заменил тяжелый физический труд.

Компьютер ускорил вычисления.

Но оставалась одна область, которую долго считали исключительно человеческой, — мышление.

Если двигатель усиливает мышцы, то почему нельзя создать инструмент, который усиливает способность работать с информацией?

Этот вопрос и стал отправной точкой для развития искусственного интеллекта.

Что значит «думать»?

Кажется, что ответ очевиден.

Но попробуйте сами определить, что такое мышление.

Это память?

Логика?

Воображение?

Способность учиться?

Язык?

Сознание?

Чем больше ученые исследовали человеческий разум, тем яснее становилось: мышление состоит из множества разных процессов.

Человек узнает лица.

Понимает речь.

Строит планы.

Принимает решения.

Фантазирует.

Запоминает.

Ошибается.

Исправляет ошибки.

Все это разные способности, которые работают вместе.

Поэтому создать «мыслящую машину» оказалось гораздо сложнее, чем представлялось сначала.

Почему первые ожидания были слишком оптимистичными

Когда появились первые компьютеры, многие ученые были уверены, что настоящий искусственный интеллект появится очень быстро.

Логика казалась простой:

Если компьютер умеет считать быстрее человека, значит скоро научится и думать лучше человека.

Но оказалось, что считать намного легче, чем понимать окружающий мир.

Компьютер легко перемножает огромные числа.

Но долгое время ему было чрезвычайно трудно распознать кошку на фотографии или понять шутку.

Для человека такие задачи кажутся простыми.

Для машины они оказались невероятно сложными.

Почему ребенок учится иначе

Представьте двух существ.

Первое — суперкомпьютер.

Второе — трехлетний ребенок.

Положите перед ними игрушечного медведя.

Ребенок сразу поймет:

— Это игрушка.

Он сможет взять ее.

Назовет цвет.

Придумает историю.

Представит, что медведь разговаривает.

Компьютеру же потребуется множество примеров, чтобы научиться распознавать такой объект на изображениях.

Это показывает важную особенность человеческого интеллекта: мы учимся не только по словам, но и через тело, ощущения, действия и опыт взаимодействия с миром.

Два пути к искусственному интеллекту

Со временем исследователи поняли, что существуют разные подходы к созданию ИИ.

Первый путь можно сравнить с написанием очень подробной инструкции.

Представьте, что вы учите робота готовить чай.

Можно записать:

Возьми чайник.

Налей воду.

Включи чайник.

Подожди.

Возьми чашку.

И так далее.

Это подход, основанный на правилах.

Второй путь похож на обучение ребенка.

Вместо длинной инструкции вы показываете тысячи примеров.

Со временем ребенок начинает понимать закономерности сам.

Именно по этому пути развивались современные нейросети.

Им не объясняют каждое правило языка по отдельности.

Они учатся на огромном количестве примеров.

Почему это похоже на обучение музыке

Представьте человека, который никогда не слышал музыки.

Теперь дайте ему учебник по теории музыки.

Он узнает, что такое ноты.

Ритм.

Гаммы.

Аккорды.

Но вряд ли сразу станет музыкантом.

Другой человек просто слушает музыку каждый день много лет.

Со временем он начинает чувствовать стиль, ритм и гармонию.

Современные языковые модели учатся скорее вторым способом.

Они не получают полный список правил языка заранее.

Они постепенно выявляют закономерности, анализируя огромное количество текстов.

Почему ИИ — это не электронный человек

Одна из самых распространенных ошибок — представлять искусственный интеллект как цифровую копию человеческого мозга.

Но это неверная аналогия.

Самолет летает.

Птица тоже летает.

Однако самолет не машет крыльями.

Он использует совсем другие принципы.

Точно так же современный ИИ способен выполнять некоторые интеллектуальные задачи, но делает это иначе, чем человек.

Иногда результаты похожи.

Но путь к ним совершенно другой.

Настоящий вопрос

Когда мы спрашиваем:

«Может ли машина думать?»

мы часто представляем себе человека внутри компьютера.

Но современный искусственный интеллект устроен иначе.

Он не стал цифровой копией мозга.

Он нашел собственный способ работать с языком и информацией.

Именно поэтому иногда он удивляет нас своими возможностями, а иногда допускает ошибки, которые человеку кажутся странными.

Глава 3. От механических автоматов до нейросетей

Представьте, что вы заходите в музей.

В первом зале стоит деревянная механическая утка XVIII века. Она машет крыльями, поворачивает голову и кажется почти живой.

Во втором — шахматный компьютер, который однажды победил чемпиона мира.

В третьем — современный чат-бот, способный написать рассказ, объяснить квантовую физику или помочь создать бизнес-план.

На первый взгляд кажется, что между ними нет ничего общего.

Но на самом деле все эти изобретения объединяет одна идея:

Человек всегда пытался создать машину, которая сможет делать то, что раньше умел только человек.

История искусственного интеллекта — это не только история технологий. Это история человеческой мечты.

Первая мечта

Еще тысячи лет назад люди представляли искусственных существ.

В древнегреческих мифах кузнец богов Гефест создавал золотых помощников.

В еврейской традиции существовала легенда о Големе — создании из глины, которое оживляли с помощью священных слов.

В разных культурах появлялись рассказы о механических слугах, говорящих статуях и искусственных людях.

Конечно, это были легенды.

Но они показывают одну важную вещь.

Человек всегда мечтал создать разум вне самого себя.

Первые настоящие автоматы

В XVIII веке мастера начали создавать удивительные механизмы.

Появлялись часы, которые играли музыку.

Куклы, умеющие писать.

Механические музыканты.

Автоматические птицы.

Для своего времени они казались настоящим чудом.

Однако все эти устройства имели одну особенность.

Они **не принимали решений**.

Каждое движение заранее было запрограммировано шестеренками и пружинами.

Если открыть механизм, внутри не окажется никакого разума.

Только сложная последовательность деталей.

Машина, которая всех обманула

В 1770 году появился знаменитый автомат под названием «Турок».

Он играл в шахматы настолько хорошо, что побеждал известных игроков Европы.

Им восхищались императоры.

Его изучали ученые.

Люди были уверены:

— Наконец-то создана мыслящая машина!

Но спустя десятилетия раскрылась тайна.

Внутри автомата прятался настоящий шахматист.

Машина лишь создавала иллюзию интеллекта.

История «Турка» напоминает нам важную вещь.

Люди очень легко принимают внешние признаки интеллекта за сам интеллект.

Эта мысль окажется важной и для понимания современных языковых моделей.

Когда появились компьютеры

В XX веке произошло событие, изменившее историю.

Появились электронные вычислительные машины.

Они считали быстрее любого человека.

Если раньше сложные вычисления занимали недели, то компьютер справлялся за минуты.

Казалось, следующий шаг очевиден.

Если машина умеет считать...

Может быть, она научится рассуждать?

Первая ошибка ученых

Многие исследователи считали, что достаточно записать все правила мышления.

Например:

Если человек говорит:

— Все люди смертны.

И:

— Сократ — человек.

То компьютер должен сделать вывод:

— Сократ смертен.

Логично?

Конечно.

Так появились первые интеллектуальные программы.

Они действительно умели делать логические выводы.

Но быстро выяснилось неожиданное.

Реальная жизнь почти никогда не состоит только из четких правил.

Мир оказался слишком сложным

Представьте, что вы хотите научить компьютер отличать кошку от собаки.

Кажется просто.

Можно написать правило:

Если уши острые — это кошка.

Но некоторые собаки тоже имеют острые уши.

Хорошо.

Добавим хвост.

Но хвост есть почти у всех.

Добавим усы.

Размер.

Форму морды.

Через некоторое время становится ясно:

Правил потребуется миллионы.

И все равно они будут ошибаться.

Новая идея

Тогда ученые решили отказаться от огромного списка правил.

Вместо этого машине начали показывать примеры.

Тысячи фотографий кошек.

Тысячи фотографий собак.

Компьютер стал искать закономерности самостоятельно.

Так появились современные методы машинного обучения.

Это было похоже на обучение ребенка.

Мы редко объясняем малышу научное определение кошки.

Мы просто много раз показываем ему разных кошек.

И постепенно он начинает узнавать их сам.

Почему это стало революцией

Старые программы были похожи на повара, который готовит только строго по рецепту.

Современные нейросети больше напоминают ученика.

Они не получают готовый список всех правил.

Они постепенно замечают закономерности.

Например, ребенок не знает правил грамматики.

Но к пяти годам уже говорит почти без ошибок.

Он никогда не изучал учебник русского языка.

Он просто много слушал взрослых.

Современные языковые модели учатся похожим образом — не через заучивание правил, а через анализ огромного количества текстов.

Почему нейросети называют нейросетями

Многих вводит в заблуждение само название.

Кажется, будто внутри компьютера находится настоящий искусственный мозг.

На самом деле это не так.

Название появилось потому, что разработчики вдохновлялись очень общей идеей работы нервной системы.

Однако современные искусственные нейронные сети устроены намного проще и работают иначе, чем человеческий мозг.

Это скорее инженерная конструкция, чем цифровая копия человека.

Самый неожиданный поворот

Долгое время ученые считали, что сначала компьютер научится понимать мир, а потом уже сможет красиво говорить.

Произошло почти наоборот.

Сначала появились модели, которые научились удивительно хорошо работать с языком.

И только потом стало ясно, что умение обращаться с языком открывает путь к решению множества других задач.

Это стало одним из самых неожиданных открытий последних лет.

История продолжается

От деревянных автоматов до современных нейросетей прошло несколько столетий.

Но цель оставалась одной и той же.

Создать машину, которая сможет выполнять интеллектуальную работу.

Каждое поколение ученых думало, что до этой цели осталось совсем немного.

На практике каждый новый шаг открывал еще больше вопросов.

И главный из них звучит так:

Если современные модели не думают так, как человек, то почему они умеют разговаривать настолько убедительно?

Ответ начинается с неожиданного факта.

ИИ вообще не видит слова так, как видим их мы.

Для него слова сначала исчезают.

А затем превращаются... в числа.

Глава 4. Почему ИИ — это не человеческий мозг

Когда люди впервые начинают пользоваться ChatGPT или другими современными системами искусственного интеллекта, почти у всех возникает одна и та же мысль:

«Наверное, внутри компьютера находится что-то похожее на человеческий мозг».

Это кажется логичным.

Мы задаем вопрос.

ИИ отвечает.

Мы продолжаем разговор.

Он помнит контекст.

Иногда шутит.

Иногда ошибается.

Иногда удивляет.

Все это очень напоминает общение с человеком.

Но здесь скрывается одно из самых больших заблуждений.

Современный искусственный интеллект не является цифровой копией человеческого мозга.

И это хорошая новость.

Представьте самолет

Если спросить ребенка:

— Что общего у птицы и самолета?

Он ответит:

— Они летают.

Это правда.

Но если спросить инженера, он скажет совсем другое.

Птица машет крыльями.

Она чувствует ветер.

Меняет положение каждого пера.

Использует мышцы.

Самолет не делает ничего подобного.

Он вообще не пытается стать птицей.

Он использует законы аэродинамики совершенно другим способом.

И тем не менее прекрасно летает.

То же самое произошло и с искусственным интеллектом.

Он не стал электронным человеком.

Он нашел собственный путь к решению некоторых интеллектуальных задач.

Мы часто путаем результат со способом

Представьте двух людей.

Один умеет складывать огромные числа в уме.

Другой пользуется калькулятором.

Ответ будет одинаковым.

Но способы совершенно разные.

Или представьте фотографа.

Один рисует картину маслом.

Другой делает снимок камерой.

Получается изображение.

Но процесс создания изображения абсолютно разный.

Так и с ИИ.

Если человек и модель отвечают на один и тот же вопрос, это еще не означает, что они пришли к ответу одинаковым путем.

Как думает человек

Попробуйте вспомнить свое детство.

Первый снег.

Запах моря.

Вкус лимона.

Первую влюбленность.

Падение с велосипеда.

Объятия родителей.

Все это стало частью вашего понимания мира.

Человек познает реальность через тело.

Через чувства.

Через боль.

Через радость.

Через прикосновения.

Через запахи.

Через собственные ошибки.

Когда вы слышите слово «яблоко», мозг связывает его не только с буквами.

Он может вызвать вкус.

Запах.

Цвет.

Воспоминания.

Эмоции.

Даже звук, с которым яблоко падает на землю.

Одно слово соединяет множество видов опыта.

А что происходит у ИИ?

Для современного языкового ИИ слово «яблоко» не связано с личным опытом.

Он не держал яблоко в руках.

Не чувствовал его запах.

Не пробовал его на вкус.

Не видел, как оно падает с дерева.

Его представление строится иначе.

Во время обучения модель многократно встречала слово «яблоко» рядом с другими словами:

дерево;

фрукт;

красный;

зеленый;

сад;

вкус;

сок;

пирог.

Постепенно она научилась учитывать такие закономерности и использовать их при создании ответов.

Важно не представлять это как набор карточек или словарь в человеческом смысле. Это результат обучения на огромном количестве примеров, а не хранение отдельных определений.

Человек постоянно проверяет мир

Представьте ситуацию.

Вы смотрите в окно.

Кажется, идет дождь.

Вы открываете окно.

Нет.

Это поливальная машина.

Вы сразу меняете свое мнение.

Ваш мозг непрерывно сверяет предположения с окружающим миром.

Современная языковая модель во время генерации ответа обычно не наблюдает мир напрямую.

Она работает с тем текстом, который получила в запросе, и с теми закономерностями языка, которым научилась ранее. Поэтому она может ошибаться в фактах или уверенно продолжать неверное предположение, если не получает новой информации.

Наш мозг экономит силы

Интересно, что между человеком и ИИ все же есть кое-что общее.

Ни человек, ни модель не анализируют каждый новый текст «с нуля».

Когда вы слышите начало фразы:

«Жили-были...»

Вы почти мгновенно ожидаете сказку.

Когда друг говорит:

«Ты не поверишь, что произошло сегодня...»

Вы готовитесь услышать необычную историю.

Мозг постоянно строит ожидания.

Языковые модели тоже используют закономерности, чтобы предсказывать наиболее вероятное продолжение текста.

Однако сходство заканчивается именно здесь.

Для человека такие ожидания связаны с опытом жизни.

Для модели — с закономерностями в данных обучения.

Почему ИИ иногда кажется живым

Есть простой эксперимент.

Представьте, что вам приходит сообщение:

«Мне сегодня грустно.»

Если это написал ваш друг, вы представляете его настроение.

Если то же самое написал ИИ, многие люди все равно начинают воспринимать сообщение как выражение настоящего чувства.

Почему?

Потому что наш мозг устроен так, чтобы видеть личность за словами.

Это качество помогает людям понимать друг друга.

Но оно же заставляет нас иногда приписывать человеческие свойства системам, которые их не имеют.

Психологи называют такую склонность **антропоморфизмом** — стремлением объяснять поведение нечеловеческих объектов через человеческие качества.

Что умеет человек, но пока не умеет ИИ

Человек способен:

проживать события;

чувствовать боль и радость;

формировать долговременные личные воспоминания;

самостоятельно ставить цели;

менять свои убеждения под влиянием жизненного опыта.

Современная языковая модель не обладает такими свойствами только потому, что умеет вести убедительный диалог. Она не имеет субъективного опыта, эмоций или сознания в человеческом смысле.

Тогда почему ИИ настолько полезен?

Потому что человеческий интеллект — это не одна способность.

Это целый набор разных навыков.

Некоторые из них современные модели освоили очень хорошо:

- работать с текстами;

- замечать языковые закономерности;

- объяснять;

- обобщать;

- переводить;

- помогать структурировать информацию.

Другие способности остаются принципиально человеческими или пока не реализованы в современных языковых моделях.

Главная мысль этой главы

Самая большая ошибка — пытаться понять ИИ, представляя внутри компьютера маленького человека.

Его там нет.

Есть совершенно иной способ обработки информации.

Именно поэтому искусственный интеллект одновременно способен поражать нас своими ответами и допускать ошибки, которые человеку кажутся невероятно странными.

Глава 5. Что значит «обучение» для ИИ

Когда человек слышит слово «обучение», он обычно представляет знакомую картину.

Ребенок идет в школу.

Студент читает книги.

Музыкант годами тренируется.

Художник делает тысячи набросков.

Человек учится через опыт.

Но когда мы говорим:

«Искусственный интеллект обучается»

— мы используем то же слово для совершенно другого процесса.

ИИ не сидит за партой.

Он не понимает объяснения учителя.

Он не запоминает уроки так, как человек.

Тогда что же происходит?

Обучение человека и обучение машины

Представьте ребенка, который впервые видит собаку.

Мама говорит:

— Это собака.

Ребенок смотрит.

Он видит четыре лапы.

Шерсть.

Хвост.

Слышит лай.

Потом встречает другую собаку.

Другую породу.

Другой размер.

Через некоторое время ребенок формирует понятие:

«Собака — это не конкретное животное, а целый

класс объектов».

Он не просто запоминает картинку.

Он создает внутреннюю модель мира.

Теперь представьте искусственный интеллект.

Ему показывают миллионы примеров:

Фотографии собак.

Описания собак.

Тексты о собаках.

Разные контексты, где встречается слово «собака».

Но модель не формирует человеческое понятие.

У нее нет внутреннего образа животного, которое можно погладить.

Она выявляет закономерности:

какие элементы часто встречаются вместе;

какие слова связаны между собой;

какие структуры повторяются.

Учитель для ИИ выглядит иначе

У человека есть учитель.

Он может сказать:

— Ты ошибся.

— Попробуй иначе.

— Обрати внимание на это.

У модели обучения другой принцип.

Ей дают большое количество примеров.

Она пытается выполнить задачу.

Потом получает сигнал:

ответ был ближе к правильному или дальше от него.

И постепенно внутренние параметры модели изменяются.

Это похоже на настройку огромного музыкального инструмента.

Представьте пианино

Перед вами пианино.

Оно настроено неправильно.

Некоторые ноты звучат слишком высоко.

Другие слишком низко.

Музыкант играет.

Слушает.

Настраивает.

Снова играет.

Постепенно инструмент начинает звучать гармонично.

Обучение нейросети похоже на такую настройку.

Только вместо нескольких струн — огромное количество внутренних связей.

ИИ не учится как ребенок

Это очень важный момент.

Когда ребенок узнает слово «мама», за ним стоит огромный опыт:

голос;

лицо;

забота;

эмоции;

отношения;

память.

Для ребенка слово становится частью жизни.

Когда модель встречает слово «мама», она изучает его связи с другими словами и контекстами.

Она может прекрасно написать:

«Мама — самый близкий человек для многих людей»

Но у нее нет собственной матери.

Нет детства.

Нет воспоминаний.

Нет личной истории.

Почему тогда ИИ знает так много?

Потому что количество примеров имеет огромное значение.

Человек за всю жизнь прочитает тысячи книг.

Современные модели обучаются на объемах текстов, которые невозможно прочитать одному человеку за множество жизней.

Это дает им удивительную способность:

находить связи;

продолжать тексты;

создавать объяснения;

видеть повторяющиеся структуры.

Но количество информации не равно пониманию.

Библиотека без читателя

Представьте гигантскую библиотеку.

В ней собраны все книги мира.

Но внутри нет ни одного читателя.

Нет человека, который бы переживал истории.

Нет того, кто бы задавался вопросами:

— Почему этот герой поступил так?

— Что хотел сказать автор?

— Как это связано с моей жизнью?

Такая библиотека содержит огромное количество информации.

Но сама по себе она ничего не переживает.

Современный ИИ похож на очень сложную систему работы с информацией, а не на человека внутри библиотеки.

Что именно запоминает модель?

Здесь возникает интересный вопрос.

Многие думают:

«Если ИИ обучился на миллиардах текстов, значит внутри лежат все эти книги».

Это не совсем так.

Модель не является обычной базой данных.

Она не хранит все тексты в виде списка.

Во время обучения она изменяет свои внутренние настройки так, чтобы лучше работать с языком.

Можно представить это как художника.

После тысячи картин в его голове остается не каждая картина отдельно.

У него формируется чувство композиции.

Цвета.

Стиля.

Пропорций.

Способов изображения.

Так же модель не хранит мир как коллекцию страниц, а формирует сложную систему закономерностей.

Почему обучение требует огромных ресурсов

Чтобы обучить современную большую модель, нужны:
мощные компьютеры;
огромное количество данных;
много времени;
специалисты разных областей.

Это похоже на строительство огромного города.

Нужно создать фундамент.

Дороги.

Электричество.

Инфраструктуру.

Только вместо домов строится математическая система, способная работать с языком.

После обучения начинается другая жизнь

Интересно, что после завершения обучения модель уже не продолжает учиться так же, как человек.

Обычный человек каждый день получает новый опыт.
Он меняется.

У него появляются новые воспоминания.

Он пересматривает взгляды.

Модель без специального процесса обновления остается такой, какой ее создали.

Она может отвечать на новые вопросы, но это не означает, что у нее автоматически появляется жизненный опыт.

Самая большая загадка

Чем больше ученые изучают искусственный интеллект, тем больше возникает философских вопросов.

Если система может писать книги...

Создавать музыку...

Вести разговор...

Объяснять сложные идеи...

Что именно происходит внутри?

Это понимание?

Или очень сложное распознавание закономерностей?

Ответ на этот вопрос связан с тем, как ИИ вообще видит язык.

Для человека слово — это смысл.

Для машины прежде всего — структура.

И чтобы понять это, нам нужно отправиться в самое неожиданное место:

в мир токенов.

Глава 6. Как ИИ превращает текст в числа

Когда человек читает предложение, он видит слова.

Например:

«Художник пишет картину маслом.»

Для нас это не просто набор букв.

Мы сразу создаем образ.

Мы представляем художника.

Мастерскую.

Кисть.

Цвета.

Запах краски.

Движение руки.

Одно короткое предложение вызывает целый мир ассоциаций.

Но для искусственного интеллекта всё начинается иначе.

Прежде чем модель сможет работать с текстом, ей нужно перевести человеческий язык в форму, которую может обрабатывать компьютер.

И этой формой становятся числа.

Компьютер не видит слова

Это одна из самых важных идей для понимания ИИ.

Мы привыкли думать:

«Компьютер читает текст».

Но компьютер не читает так, как человек.

Он не видит буквы и не слышит внутренний голос.

Для него текст должен быть преобразован в математическое представление.

Можно представить это как перевод с одного языка на другой.

Человек говорит на языке слов.

Компьютер работает с языком чисел.

Что такое токены

Чтобы превратить текст в числа, современные языковые модели используют специальные единицы — **токены**.

Токен — это небольшой фрагмент текста.

Это может быть:

целое слово;

часть слова;

знак препинания;

сочетание букв.

Например, слово:

«искусственный»

может быть разделено на несколько частей.

Почему?

Потому что язык содержит огромное количество слов.

Если создать отдельный «значок» для каждого слова, система станет слишком громоздкой.

Кроме того, люди постоянно создают новые слова.

Появляются:

новые технологии;

новые имена;

новые термины;

новые выражения.

Поэтому удобнее работать не только со словами целиком, а с их частями.

Представьте конструктор

У ребенка есть коробка с деталями LEGO.

Он может построить тысячи разных объектов.

Ему не нужны отдельные детали для каждого возможного дома.

Есть:

кубики;

окна;

двери;

крыши.

Из небольшого набора элементов можно создавать бесконечное количество комбинаций.

Токены работают похожим образом.

ИИ не обязательно должен иметь отдельный элемент для каждого слова.

Он использует небольшие части, из которых можно собрать огромное количество выражений.

Но где здесь смысл?

Вот здесь возникает самый интересный вопрос.

Если слово превращается в число, не исчезает ли смысл?

На первый взгляд кажется:

«Ну хорошо, компьютер получил набор цифр. Но как он понимает, что это означает?»

Ответ сложнее, чем кажется.

ИИ не хранит значение слова как определение в словаре.

Например, у человека:

Солнце:

звезда;

тепло;

свет;

лето;

рассвет;

детские воспоминания.

У модели нет личных ощущений.

Но во время обучения она видит огромное количество примеров:

«солнце светит»;

«солнце восходит»;

«солнечный день»;

«солнце находится в центре Солнечной системы».

Постепенно она формирует математические связи между

ПОНЯТИЯМИ.

Карта смыслов

Представьте огромную карту.

В одном районе находятся слова:

море;

волна;

корабль;

берег;

путешествие.

В другом:

компьютер;

программа;

данные;

алгоритм.

Слова, которые часто встречаются вместе, оказываются ближе друг к другу.

Так появляется своеобразная карта языка.

Но важно понимать:

это не карта мира.

Это карта закономерностей языка.

Человек видит мир, ИИ видит связи

Представьте два человека.

Первый говорит:

— Я люблю море.

Для него море может означать:

детство;

отпуск;

свободу;

романтику;

воспоминания.

Второй человек никогда не был у моря.

Он прочитал миллионы описаний моря.

Он знает, какие слова обычно связаны с морем.

Он может прекрасно написать:

«Море символизирует бесконечность и свободу».

Но его отношение к морю построено не на опыте.

Так отличается человеческий смысл от статистической
связи.

Почему это всё равно работает

Может возникнуть вопрос:

«Если ИИ не понимает слова как человек, почему его ответы настолько хорошие?»

Потому что язык сам по себе содержит огромное количество информации о человеческом мышлении.

Когда люди пишут:

книги;

статьи;

письма;

исследования;

рассказы;

они оставляют в языке следы своего опыта.

В языке отражаются:

логика;

культура;

эмоции;

знания;

способы объяснения мира.

ИИ учится видеть эти закономерности.

Аналогия с художником

Представьте художника, который никогда не видел настоящего леса.

Но у него есть миллион картин лесов.

Он изучает:

какие цвета обычно используют;

как изображают деревья;

как передают свет;

какие композиции встречаются чаще.

Он сможет создать картину леса.

Возможно, очень красивую.

Но его лес будет создан из изученных образов, а не из личного путешествия среди деревьев.

Почему перевод текста в числа — это революция

Потому что числа позволяют компьютеру работать с языком.

Модель может:

сравнивать слова;

находить закономерности;

учитывать контекст;

строить связи;

создавать новые комбинации.

То, что для человека является потоком смыслов, для машины становится сложной системой отношений между элементами.

Самый важный вывод

Когда вы пишете вопрос ИИ, происходит удивительный процесс.

Вы вводите:

мысль слова токены числа

А затем происходит обратное превращение:

числа закономерности токены слова ответ

Весь разговор с искусственным интеллектом — это путешествие туда и обратно между человеческим языком и математической системой.

Но возникает новый вопрос.

Если ИИ работает не со словами, а с числовыми представлениями, то **как он понимает, какое слово должно появиться следующим?**

Почему после одного слова появляется именно другое?

Почему ответ строится именно так?

Глава 7. Мир токенов: как ИИ видит язык

Когда человек смотрит на страницу книги, он видит слова.

Он видит буквы.

Предложения.

Абзацы.

Знаки препинания.

Но за этими символами у человека мгновенно возникает целый мир.

Фраза:

«Осенний парк после дождя»

может вызвать у одного человека воспоминание о детстве, у другого — образ будущей картины, у третьего — чувство спокойствия.

Человек воспринимает язык не просто как последовательность знаков.

Он воспринимает его как дверь в мир опыта.

ИИ смотрит иначе.

Для него предложение — это сложная структура взаимосвязанных элементов.

Чтобы понять, как работает искусственный интеллект, нужно ненадолго забыть привычное человеческое восприятие языка и попробовать посмотреть на него глазами маши-

ны.

Слово для человека и слово для ИИ

Возьмем слово:

«дом»

Для человека оно может означать:

место, где он вырос;

чувство безопасности;

семью;

воспоминания;

архитектуру;

пространство жизни.

Но для модели слово «дом» не является объектом опыта.

Она не представляет комнату.

Не чувствует запах дерева.

Не знает ощущения возвращения домой после долгой дороги.

Она видит огромное количество контекстов:

«строить дом»;

«родной дом»;

«дом культуры»;

«дом у моря»;

«дом как символ».

Из этих связей формируется математическое представле-

ние.

Язык как огромная сеть

Представьте город.

В нем есть миллионы дорог.

Одни улицы соединяют центр с окраинами.

Другие ведут к паркам.

Третьи — к промышленным районам.

Если вы знаете карту города, вы можете предсказывать:

куда ведет дорога;

какие места находятся рядом;

какой маршрут выбрать.

Язык тоже похож на город.

Слова соединены между собой огромным количеством связей.

Например:

слово «художник» связано с:

картиной;

кистью;

цветом;

мастерской;

творчеством;

выставкой.

Но оно также может быть связано с:

историей искусства;

стилем;

техникой;
вдохновением.

ИИ изучает не отдельные слова, а всю систему дорог между ними.

Почему одно слово может иметь разные значения

Возьмем слово:

«лук»

Для человека понятно:

это может быть овощ.

А может быть оружие.

Мы определяем значение благодаря контексту.

Фраза:

«Я добавил лук в суп»

и:

«Он натянул лук перед выстрелом»

имеют совершенно разные смыслы.

ИИ тоже должен научиться учитывать окружение слова.

Именно поэтому современные модели называются **языковыми моделями**.

Они работают не только с отдельными словами, а с отношениями между ними.

Контекст меняет всё

Представьте человека, который слышит только одно слово:

«ключ»

Что это?

Ключ от двери?

Родник?

Инструмент?

Музыкальный ключ?

Без контекста невозможно понять.

Но если сказать:

«Я потерял ключ от квартиры»

значение становится очевидным.

ИИ работает похожим образом.

Он анализирует не только отдельный элемент, но и его окружение.

Каждое новое слово меняет понимание предыдущих.

Как слова превращаются в координаты

Можно представить, что каждое слово получает свое место в огромном пространстве.

Не физическом.

Математическом.

У каждого понятия появляется своеобразный «адрес».

Слова с похожими значениями оказываются ближе.

Например:

радость;

счастье;

удовольствие;

восторг

будут находиться ближе друг к другу, чем:

радость;

трактор;

вулкан.

Но это не значит, что ИИ «чувствует» радость.

Он видит структуру отношений между словами.

Почему метафоры работают

Удивительно, но ИИ хорошо работает с метафорами.

Например:

«Время — это река».

Почему?

Потому что люди веками создавали такие связи:

время течет;

река движется;

жизнь имеет направление.

В огромном количестве текстов подобные сравнения повторяются.

Модель учится замечать эти закономерности.

Она может создавать новые метафоры, комбинируя уже известные языковые структуры.

Но где граница?

Возьмем фразу:

«Сердце города».

Человек понимает.

Это не настоящий орган.

Это центр, главная часть, место жизни.

ИИ тоже сможет объяснить эту метафору.

Но важно понять разницу.

Человек понимает ее через опыт тела:

сердце действительно существует;

оно связано с жизнью;

оно ощущается внутри нас.

ИИ работает через связи между выражениями.

Почему иногда ИИ создает удивительные идеи

Многие люди замечают:

«Я попросил ИИ придумать идею, и он предложил то, чего я сам не ожидал».

Почему это происходит?

Потому что человек часто думает привычными маршрутами.

У него есть:

личный опыт;

любимые темы;

профессиональные привычки;

устоявшиеся взгляды.

ИИ не имеет такой личной истории.

Он может соединять области, которые человек обычно не соединяет.

Например:

искусство + биология;

психология + архитектура;

история + технологии.

Это делает его полезным творческим партнером.

Опасность красивых связей

Но здесь есть важная ловушка.

Красивое соединение слов еще не означает истину.

ИИ может создать:

убедительную теорию;

красивое объяснение;

интересную гипотезу.

Но это может оказаться ошибочным.

Почему?

Потому что язык и реальность — не одно и то же.

Карта города не является самим городом.

Так же карта языка не является самим миром.

Главная идея главы

ИИ видит язык не так, как человек.

Человек видит:

слово опыт смысл

ИИ работает иначе:

слово связи вероятности новый текст

Именно поэтому он может создавать удивительные ответы, не имея человеческого понимания.

Он не путешествует по миру.

Он путешествует по огромной карте человеческого языка.

Но возникает следующий вопрос:

Если ИИ работает со связями между словами, то что именно хранится внутри него? Где находятся все эти знания?

Глава 8. Почему слова становятся математикой

Одна из самых загадочных особенностей искусственного интеллекта заключается в следующем:

Как могут обычные человеческие слова превратиться в числа и при этом сохранить смысл?

На первый взгляд это кажется невозможным.

Слово «любовь» — это чувство.

Слово «красота» — это идея.

Слово «свобода» — это философское понятие.

Как можно превратить такие сложные вещи в математические значения?

Чтобы понять это, нужно представить, что математика в ИИ — это не холодные цифры, а своеобразный язык отношений.

Числа тоже могут описывать мир

Когда мы слышим слово «число», мы часто думаем о простом счете:

1, 2, 3, 4.

Но математика гораздо шире.

Она может описывать:

расстояние;

форму;

движение;

вероятность;

связи между объектами.

Например, карта города тоже состоит из чисел.

Координаты показывают, где находится место.

Расстояния показывают, насколько объекты близки друг к другу.

Но сама карта не является городом.

Она лишь помогает описывать его структуру.

Так же ИИ использует числа не как «значения слов», а как способ описать отношения между понятиями.

Представьте библиотеку без алфавита

Допустим, существует огромная библиотека.

Но вместо названий книг на обложках стоят только числа.

Одна книга имеет номер:

Другая:

Сначала кажется, что это бессмысленно.

Но если рядом появляется информация:

какие книги часто читают вместе;

какие темы встречаются рядом;

какие идеи повторяются;

постепенно можно восстановить структуру библиотеки.

Можно понять:

какие книги похожи;

какие темы связаны;

какие направления существуют.

ИИ работает похожим образом.

Слова получают «координаты смысла»

Внутри модели слова и понятия представлены как математические объекты.

Можно представить их как точки в огромном многомерном пространстве.

Это не пространство, которое можно увидеть глазами.

В нем могут быть тысячи измерений.

Например, одна группа координат может отражать:

связь с эмоциями;

связь с действиями;

связь с предметами;

связь с определенными темами.

Важно понимать:

это не значит, что внутри ИИ существует маленький словарь с определением каждого слова.

Там нет карточки:

«Любовь — это сильное чувство привязанности».

Вместо этого существует сложная система связей.

Аналогия с художником

Представьте художника.

Он смотрит на картину великого мастера.

Он не записывает:

красный цвет = страсть;

синий цвет = спокойствие;

вертикальная линия = сила.

Но после многих лет практики у него формируется художественное мышление.

Он начинает чувствовать:

какие цвета сочетаются;

какая композиция работает;

какое настроение создается.

Его знание находится не в виде списка правил.

Оно стало частью его навыка.

У ИИ происходит нечто похожее, но в другой форме.

Почему слово связано с другими словами

Возьмем слово:

«океан»

Если человек слышит его, он может представить:

волны;

берег;

путешествие;

бесконечность.

Но почему именно такие ассоциации?

Потому что в человеческом опыте они уже связаны.

ИИ обнаруживает похожие связи через тексты.

Он видит, что:

«океан» часто появляется рядом с:

водой;

глубиной;

кораблями;

штормом;

горизонтом.

Так формируется математическая структура понятия.

Самое удивительное: отношения важнее самих слов

Для ИИ важно не только слово.

Важно место слова среди других слов.

Например:

Если изменить отношения между понятиями, меняется вся структура.

Можно провести аналогию с живописью.

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.