



Станислав Петровский

# Темная сторона искусственного интеллекта

Станислав Петровский

**Темная сторона искусственного  
интеллекта. Почему  
ученые бьют тревогу?**

«Издательские решения»

## **Петровский С.**

Темная сторона искусственного интеллекта. Почему ученые бьют тревогу? / С. Петровский — «Издательские решения»,

ISBN 978-5-00-533300-1

В 2023 году более тысячи ученых и разработчиков всего мира подписали открытое письмо с призывом приостановить развитие искусственного интеллекта. Билл Гейтс вовсе сравнивает искусственный интеллект с атомной бомбой. Какие из угроз надуманы, а какие реальны? Книга рассчитана на широкий круг читателей, интересующихся влиянием на общество современных информационных технологий.

ISBN 978-5-00-533300-1

© Петровский С.  
© Издательские решения

# Содержание

Введение	6
Естественный интеллект	8
Что называют искусственным интеллектом?	13
Конец ознакомительного фрагмента.	17

# **Темная сторона искусственного интеллекта Почему ученые бьют тревогу?**

**Станислав Петровский**

© Станислав Петровский, 2024

ISBN 978-5-0053-3300-1

Создано в интеллектуальной издательской системе Ridero

## Введение

Сегодня смартфон превосходит суперкомпьютеры, которые работали тридцать лет назад. Вычислительные мощности, которые раньше требовали устройство размером с комнату и ценой в десятки миллионов, теперь помещаются в кармане и стоят тысячи. Скорость передачи данных также выросла в тысячи раз. Современный уровень скорости вычислений и передачи данных позволяет широко применять новый вид компьютерных программ и устройств на базе технологий искусственного интеллекта (ИИ).

Сочетание гениальных человеческих умов и ИИ может дать фантастические возможности – лечение ранее неизлечимых болезней, продление жизни, значительное снижение количества аварий на транспорте и производствах, большой комфорт, экологичность и общественную безопасность.

Однако, как у любой технологии, есть и обратная, темная сторона огромной силы ИИ. Современная реальность все больше похожа на сюжеты Стивена Кинга. Некоторые обыденные вещи кажутся безобидными на первый взгляд, но таят в себе смертельную опасность. Это абсолютно новый вызов, не похожий на чуму или атомную бомбу. Интуиция и разум, сформировавшиеся за миллионы лет для выживания человека в природе, теперь подводят нас. Когда речь идет о компьютере – с виду безобидной коробке с разноцветными лампочками, трудно осознать скрытые в ней угрозы. У машины нет собственных эмоций, эволюционной истории с реакцией типа «бей или беги». Зачем ей завоевывать мир или причинять вред человеку?

ИИ может и не нужно, а вот в руках алчных корпораций или злонамеренных спецслужб или хакеров ИИ вполне способен принести ущерб, сравнимый с ущербом от мировой войны – массовую безработицу, усиление разрыва между бедными и богатыми, диктатуру полицейского государства, манипуляции общественным мнением в социальных сетях, техногенные катастрофы, гонку вооружений и войны.

В 2017 году на конференции в США были приняты Азиломарские принципы ИИ. Данные принципы отражают общие подходы к созданию и использованию систем ИИ. Ведущие мировые разработчики и эксперты по ИИ, включая Илона Маска, Стивена Хокинга и Рэя Курцвейла, а также представители Google, Apple, IBM, Microsoft и другие согласились с тем, что инвестиции в ИИ должны сопровождаться исследованиями, позволяющими развивать правовые системы с учетом рисков, связанных с ИИ.

В соответствии с Азиломарскими принципами, в частности:

– системы ИИ должны быть совместимы с идеалами человеческого достоинства, его прав и свобод, многообразия культур;

– экономическое процветание, достигнутое благодаря ИИ, должно широко использоваться в интересах всего человечества;

– следует избегать гонки вооружений в разработке смертельного автономного оружия<sup>1</sup> и так далее.

В 2020 году официальные лица Ватикана совместно с компаниями Микрософт и IBM утвердили этические принципы, по которым ИИ должен действовать прозрачно, работать надежно и беспристрастно, учитывать права человека, включая право на неприкосновенность частной жизни.

---

<sup>1</sup> Автономное оружие – это система, которая сама находит и поражает цель. Например, в Израиле применяется беспилотный летательный аппарат, запрограммированный атаковать радиолокационные системы противника. В ходе применения таких систем остро встает вопрос отличия автономным оружием военных от гражданских лиц и распознавания своих и чужих войск, полицейских и охранных подразделений.

Папа Римский Франциск призвал молиться, чтобы достижения в области ИИ и робототехники служили на благо человечества и не привели «к новой форме варварства, где общее благо сброшено верховенством слова сильнейшего».

К сожалению, следует констатировать: вопреки предостережениям, многие угрозы ИИ стали реальностью.

Стартовала гонка автономных вооружений. Невиданное ранее совершенство технологий подделки голоса, видео и фотографий используется для того, чтобы опорочить или шантажировать не только политиков или звезд, но и обычных людей, включая детей. В крупных городах каждый находится под полицейским надзором камер, подключенных к системе распознавания лиц. Растут технические возможности преступников для мошенничества, краж, шантажа, угроз, а также похищения людей.

Данные факты показывают необходимость совершенствования регулирования применения новых технологий.

По словам Эльвиры Талапиной, главного научного сотрудника Института государства и права Российской академии наук, доктора юридических наук, доктора права (Франция):

«Амбициозные планы человечества по использованию цифровых технологий во всех сферах жизнедеятельности впечатляют масштабами и футуристическими зарисовками, но не стоит забывать, что все это будет происходить непосредственно в нашем обществе. В обществе со сложившейся правовой культурой, традициями и ожиданиями. Было бы несомненной ошибкой концентрироваться лишь на радужных перспективах, игнорируя риски, которые таят цифровые технологии. Перед публичным правом стоит задача осмысления технологического обновления, с тем чтобы уберечь при этом устоявшиеся правовые ценности, одну из которых представляют собой фундаментальные права человека».<sup>2</sup>

Эта книга написана, чтобы стучаться в каждую дверь. Эта книга – приглашение задуматься над социальными и правовыми проблемами в сфере ИИ. Их решение видится не только в технологиях, но и в законах. Только все вместе мы способны повлиять на принятие новых законов, ограничивающих корпорации и правительства во имя процветания и самой жизни человечества.

---

<sup>2</sup> Талапина Э. Алгоритмы и искусственный интеллект сквозь призму прав человека // Журнал российского права. 2020. №10.

## Естественный интеллект

Говоря об искусственном интеллекте, мы неизбежно сталкиваемся с более широким вопросом: что такое интеллект?

Ученые придумали сотни, если не тысячи вариантов ответов на этот вопрос и не пришли к точному определению. В общем, мы понимаем, что интеллект включает память, самосознание, целеполагание, умение делать обобщения, обучение, прогнозирование последствий поступков, планирование, управление поведением, оценку результатов действий и решений. Некоторые также относят к интеллекту мышление по аналогии (распознавание проблем или поведения людей, похожих на виденное или прочитанное в прошлом), а также креативность, интуицию, озарение. Также выделяют эмоциональный интеллект, т. е. способность понимать взаимоотношения людей, предвидеть эмоциональную реакцию на слова, интонации, жесты и поступки. Все эти элементы интеллекта взаимосвязаны.

Так, память нужна, чтобы хранить сведения о себе, знания, необходимые для прогнозов, планирования, деятельности и оценки результатов. Соответственно, от наших воспоминаний зависят самооценка, ценности, которые мы разделяем, профессиональные знания и навыки, умение обучаться. То, что мы помним, во многом определяет наш образ действий и личность.

Память человека не является раз и навсегда записанным списком событий, мыслей и чувств. Каждый раз, когда мы вспоминаем какую-то информацию, мы как будто достаем папку с полки, выбрасываем из нее ненужные документы и добавляем новые.

Данный эффект используется адвокатами, которые беседуют со свидетелями перед судом. Задавая наводящие вопросы, адвокат упорядочивает картину произошедшего в голове свидетеля. В итоге свидетель твердо помнит и описывает события... иногда не совсем соответствующие действительности.

Интереснейший эксперимент на эту тему провели в 1974 году Джон Палмер и Элизабет Лофтус в Калифорнийском университете (США). Двум группам испытуемых показали одинаковые фотографии с места автокатастрофы. Через некоторое время одну группу спросили: «Как быстро машина ехала в момент, когда столкнулась с другой машиной?», а второй группе задали похожий вопрос, но с глаголом «разбила». Участники второй группы в своих ответах указали большую скорость автомобиля при столкновении. Кроме того, при повторном опросе через несколько дней участники второй группы согласились с утверждениями экспериментаторов, что видели осколки стекла на фотографиях.

Таким образом, вопрос о разбитой машине вызвал из памяти образ фотографии автокатастрофы и дополнил его описанием действия – машина разбита. А раз авто разбито, значит, должны быть стекла – воображение достроило недостающие детали.

Это еще одно подтверждение огромного значения исторических книг и фильмов, ведь они, без преувеличения, способны кардинально поменять память об истории даже очевидцев событий. Конечно, одномоментное существенное искажение фактов будет замечено и отмечено людьми с критическим мышлением. Но незначительная корректировка, смещение точек зрения, скорее всего, пройдут незамеченными и через какое-то время станут привычными. После этого возможно следующее и следующее небольшое смещение в описаниях фактов и их оценке, что в итоге может привести к кардинальной подмене коллективной памяти. Подмене как сознательно продвигаемой определенными людьми, так и произвольно сложившейся в силу особенностей человеческой памяти.

Хорошим примером распространенной и, по сути, удивительно благотворной корректировки памяти являются воспоминания о детстве и молодости. Обсуждая с родителями и друзьями интересные события из прошлого, мы придаем им дополнительный шарм. Плохие события бледнеют на этом фоне, поскольку обычно принято сглаживать негатив. Через десятки

и сотни таких повторений детство помнится многим из нас как каждодневный праздник, который иногда портили мелкие досады. Так миф о беззаботности детства создается особенностями работы человеческой памяти. Сталкиваясь с реальностью, мы удивляемся и переживаем, когда выясняется, что у детей немало всяких тревог и неприятностей. Это нормально, ведь память работает так, что изменяет воспоминания каждый раз, когда к ним обращаются.

Огромная сила социальных сетей и СМИ заключается в возможности вызывать в нашей памяти определенные воспоминания, много раз обращаясь к тому или иному факту или оценочному суждению. Таким образом, четвертая власть способна сохранить или изменить коллективную память о событиях, людях и целых странах.

Мы называем себя Homo Sapiens, т. е. относим к Человеку Разумному. Конечно, такое самоназвание можно отнести к разряду комплиментов или даже лестии. Наши возможные оппоненты возразить не могут: другие представители вида людей вымерли, смешались с нами или истреблены. Но все же, если отбросить иронию, разум является основным признаком, определяющим и отделяющим человека от всего остального мира – ни у животных, ни у вещей разума нет (пока). При этом и первые, а теперь и вторые могут демонстрировать поведение, которое называют умным или интеллектуальным (например, сейчас, помимо умного пса, мы говорим об умном доме, интеллектуальной системе парковки).

Чем же отличается человеческий разум? В замечательной книге «Sapiens. Краткая история человечества» Юваль Ной Харари пишет: «...животные, весьма схожие с современными людьми, впервые появились 2,5 миллиона лет тому назад... древние люди умели играть и любить, между ними складывались прочные отношения, они боролись за власть и статус, но так же себя вели и шимпанзе, бабуины и слоны». Несмотря на большой мозг, умение пользоваться орудиями труда и сложную социальную организацию, люди боролись за выживание миллионы лет, довольствуясь обедками после львов и других хищников. Поколение за поколением люди повторяли примерно одно и то же, их жизнь не менялась, как в фильме «День сурка».

Что же вознесло современного человека на вершину эволюции? Есть много жарких споров на эту тему. Большинство исследователей считают, что причиной взлета человека стала революция интеллектуальных способностей, которая произошла примерно 70 000 лет назад. Революция заключалась в обретении людьми возможности думать и общаться, используя сложную устную речь. Под сложной речью мы здесь понимаем потенциальную возможность строить из ограниченного набора простых звуков слова, из слов – предложения, а из предложений – неограниченное по сложности повествование (иерархичность, вложенность речи).

Причины произошедшего обновления программы, по которой работал мозг наших предков, пока точно неизвестны. Известно только, что у человека развился гораздо более сложный и большой головной мозг, чем у других животных. Для сравнения, у кошки объемы спинного и головного мозга примерно равны, у собаки головной мозг в 3 раза больше спинного, у человека головной мозг больше спинного в 50 раз.

Мозг содержит более 85 миллиардов нейронов. Каждый нейрон может иметь до 10 тысяч связей. Удивительно, но факт: мозг каждого из нас – самая сложная вещь из известных во Вселенной!

Ученые высказывают различные предположения о причинах такого развития головного мозга человека от случайной мутации до естественного или искусственного отбора либо комбинаций этих причин. Большой мозг требует много энергии – составляя всего 2% массы тела, он потребляет 20% всей энергии организма. Наличие такого энергоемкого органа должно быть оправдано какой-то эволюционной необходимостью, потребностями выживания или размножения.

Одна из любопытных теорий выдвинута исследователями из Университета Сент-Эндрюс. Они полагают, что мозг человека увеличивался из-за необходимости решать проблемы с добычей и сохранением пищи, а также с сотрудничеством или борьбой с другими людьми.<sup>3</sup>

Была построена математическая модель, которая учитывает энергетические затраты на растущий мозг, а также оценку того, насколько более крупный мозг позволяет решать те или иные проблемы. Эта модель показала, что доля проблем с добычей пищи в увеличении мозга составила 60%, 30% – доля необходимости общаться с другими людьми и 10% – необходимость конкурировать. Авторы пришли к выводу, что развитая социальная жизнь, напротив, способствует уменьшению мозга, поскольку позволяет полагаться на других при решении проблем.

Есть другая гипотеза, согласно которой мозг развивался по мере расселения людей по планете. Самые активные и интеллектуальные люди отторгались консервативным обществом, уходили с насиженных мест и создавали новые поселения. В новых поселениях выживали и оставляли большое потомство те, у кого получалось лучше приспособиться к новым условиям, т. е. более умные особи. Затем в этих поселениях также выделялись наиболее активные и уходили дальше. Вспомните 60% в развитии мозга, необходимые на адаптацию к внешней среде – выживанию, добыче и сохранению пищи, из предыдущей теории. Этот искусственный отбор значительно увеличил размер мозга, поскольку у более адаптированных, умных людей выживало больше потомков. Кроме того, дети таких людей получали больше ресурсов для развития в плане питания и обучения.

Когда расселение человека по планете завершилось, механизм искусственного отбора перестал работать. В больших обществах преимущество отдается способностям к общению и коллективным действиям, личный интеллект имеет уже не такое большое значение.

Результат не заставил себя ждать – вот уже 30 000 лет объем мозга человека постепенно сокращается. При этом растет сложность нашей культуры и цивилизации. Такой рост происходит за счет все более сильного разделения труда. Каждый человек по отдельности знает и умеет меньше, чем раньше, но вместе мы знаем гораздо больше прежних людей. Специализации отдельных людей, их обмен результатами труда и знаниями позволяет наращивать сложность деятельности без усложнения деятельности конкретного человека. Добавлю от себя, что мобильный интернет, объединяя еще больше, снижает значение индивидуального интеллекта. Похоже на то, что дальнейшее развитие мозга человека естественным биологическим путем, скорее всего, невозможно.

Приведенные исследования и соображения о причинах возникновения сложного мозга и устной речи у человека не отвечают на главные вопросы. Почему при схожих стартовых условиях жизни у человекоподобных обезьян не появились речь и цивилизация?

Для меня лично это одна из самых больших и поразительных загадок Вселенной.

По сравнению с мозгом животных, человеческий мозг очень хорошо способен «понимать строение форм и рисунков, состоящих из различных элементов, представлять эти структуры в виде символов и использовать эти символы в еще более сложных структурах».<sup>4</sup> Символы, условные обозначения каких-либо предметов или явлений, позволяют людям легко и быстро обмениваться идеями и обсуждать их, передавая и объединяя опыт и знания. Один символ может содержать в себе несколько вложенных.

К примеру, когда вы слышите, что знакомый хочет открыть кафе, вы сразу представляете ряд символов, который ассоциируется у вас с кафе – чашечку ароматного кофе, столики, стулья или кресла, диваны, официантов, меню, витрину и тому подобные вещи. Чем более раз-

---

<sup>3</sup> Tough times made our brains grow large. URL: <https://news.st-andrews.ac.uk/archive/tough-times-made-our-brains-grow-large/>

<sup>4</sup> Курцвейл Р. Эволюция разума, или бесконечные возможности человеческого мозга, основанные на распознавании образов. – М.: Эксмо, 2018. – С. 11.

вита культура, тем более сложные символы она порождает. К примеру, в некоторых племенах аборигенов нет слова «дерево», а есть только конкретные названия деревьев, обычно дающих съедобные или несъедобные плоды либо пригодных для возведения построек.

Более развитые культуры оперируют более сложными символами, например, такими как государство. Оно включает в себя массу других символов, начиная с независимых от других государств органов власти, своей границы, армии и заканчивая отдельными людьми.

Само по себе общение не отличает человека от животных. Все животные общаются друг с другом, передавая жестами, мимикой, позами и звуковыми сигналами конкретную информацию.

К примеру, воздушный танец пчел перед ульем содержит в себе информацию о направлении и длительности полета до ближайшего источника пищи. Дельфины даже различают сорочичей по именам, данным при рождении. Предполагается, что их словарный запас составляет до 14 000 звуковых сигналов! Исследовательская группа SpeakDolphin.com недавно совершила научный прорыв в понимании языка дельфинов, установив, что они обмениваются графическими образами с помощью звуков. Образно говоря, дельфины могут отправлять друг другу фотографии того, что они видят.

Однако даже такая удивительная и поразительно сложная система общения не привела дельфинов или других животных к созданию цивилизации.

Только человеческий язык способен передать абстрактную информацию, сообщить о том, что нельзя пощупать. Мифы, сказки и сплетни позволяют передавать информацию о том, кто и как себя вел, на кого можно положиться, а на кого не стоит, как правильно поступать, а как неверно. С появлением языка люди смогли учиться не только путем подражания старшему или на своих ошибках, они смогли объединять опыт и знания всех членов племени, они смогли начать накапливать информацию. Накапливались не только факты, но и мифы – попытки объяснения того, как появились мир и человек в нем, что объединяет племя или народ, почему обязательно соблюдать обычаи и так далее.

Давайте отметим для себя, что вот эти новые способности к объединению коллектива с помощью речи позволили ранее прозябавшему виду человека полностью извести или растворить в своей культуре все другие виды людей, а также уничтожить или резко сократить численность всех хищников, которые властвовали на Земле миллионы лет до этого. Эта способность позволяла менять поведение всего племени, например, перенимать удачный прием одного охотника, накапливать и передавать из поколения в поколение эффективные способы охоты и выживания в природе, а также мифы и обычаи.

Человеческий интеллект позволяет мечтать, общаться на абстрактные темы, обучать, устанавливать правила поведения, договариваться о совместных действиях, а также следовать договоренностям в поведении и общении. Такие возможности кажутся настолько избыточными для эволюции, для обычного выживания, что издревле вызывали восхищение мыслителей, идеи о чудесной природе разума, о божественном сотворении человека.

Почему у религии, философии и науки пока не получилось создать стройных непротиворечивых представлений о человеческом интеллекте, сознании, личности, душе?

Думается, дело тут не только в неопределенности того, что мы понимаем под указанными категориями. Интеллект, сознание, душа – сами по себе субстанции изменчивые. Их постоянно меняют новый опыт, переживания, размышления. Иногда изменения резкие и видны сразу, иногда изменения накапливаются постепенно и незаметно.

Современные исследования показывают, что человеческий разум – конгломерат параллельно работающих центров в головном и спинном мозге, на которые влияют гормоны и сигналы от нервной системы. Гормоны удовольствия, страха, а также чувства боли, холод, тепло и т. д. активизируют или тормозят узлы, отвечающие за обработку различных ситуаций и дан-

ных. Эмоциональное состояние, привычки, утомление и другие факторы влияют на то, принимаем ли мы осмысленные решения или действуем на автомате.

Ученые знают, какие зоны мозга человека отвечают за отдельные операции, например за распознавание зрительных образов или обоняние. Но одна и та же задача может решаться с помощью разных областей мозга. Например, привычные действия выполняются быстрой системой мышления как бы на автомате, а когда нам приходится задуматься, действует другая область мозга. Подробнее об этих системах мышления написано в книге лауреата Нобелевской премии Даниэля Канемана «Думай медленно... Решай быстро!».

Что остается загадкой: есть ли главная область мозга, раздающая задания другим областям мозга, или наше сознание – это лишь видимость?

Многие процессы в мозге выполняются параллельно, к примеру, управление движением, распознавание речи и изображений, поэтому для прорыва в изучении сознания нужен инструментарий параллельных вычислений, позволяющий изучать биологические системы. Кроме того, существующая аппаратура недостаточно точно и быстро фиксирует нервные импульсы в мозге. Магнитно-резонансный томограф полчаса делает снимок области мозга с точностью в миллионы нейронов, тогда как в мозге за секунду проходит множество нервных импульсов.

В настоящее время ученые из Университета Карнеги – Меллона в Питтсбурге (США) и из университета Киото (Япония) собрали данные исследований магнитно-резонансным томографом, создав список из 240 реакций мозга на определенные изображения (людей, мест, действий). Они обучили ИИ с точностью 87% определять, о каком человеке, месте или физическом действии из списка думал человек. Потенциально продолжение исследований в этом направлении может привести через десятки лет к созданию аппаратуры для чтения мыслей как в фантастических произведениях.

Исследования случаев раздвоения личности с применением современного оборудования позволят в будущем пролить больше света на природу сознания и личности человека. В документальном романе «Таинственная история Билли Миллигана» (англ. *The Minds of Billy Milligan*) наиболее ярко описано диссоциативное расстройство личности, когда у человека множество личностей. Если регистрировать и сравнивать мозговую активность в периоды, когда у больных активизируются различные личности, то можно многое понять. Но регистрировать только активность головного мозга недостаточно, т. к. это лишь часть аппарата, реализующего сознание. Хотя функциональная магнитно-резонансная томография позволяет дальше продвинуться в изучении мозга, этого инструментария и его сегодняшней скорости явно недостаточно для исследования сознания.

Учитывая вышесказанное, мы вряд ли увидим в ближайшие десятилетия ИИ, обладающий сознанием как у человека. Ученые пока не знают, может ли в принципе возникнуть сознание без человеческого мозга, а возможно, и других органов человеческого тела.

Плохая новость: для реализации основных угроз возникновения искусственного разума не нужно. Достаточно, чтобы применялось несколько специализированных ИИ, которые выполняли бы отдельные интеллектуальные задачи хотя бы как человек. Сейчас компьютеры уже лучше человека торгуют на бирже, играют в шахматы и Го, распознают лица на видео и болезни на снимках, рекомендуют покупки и управляют социальными сетями.

## Что называют искусственным интеллектом?

Термин «искусственный интеллект» (англ. artificial intelligence, AI) ввел в обиход Джон Маккарти на конференции в Дартмутском университете в 1956 году. Он определил ИИ как свойство интеллектуальных систем выполнять творческие функции, которые традиционно выполнял человек. При этом речь не идет об обязательном изготовлении искусственного аналога интеллекта человека. Главное – получить с помощью искусственно созданных систем результаты, которые раньше получал только человек.

Скажем прямо, несмотря на распространенность, этот термин скорее запутывает, чем проясняет смысл стоящих за ним явлений.

В настоящее время под ИИ в большинстве компьютерных разработок и в этой книге понимается вовсе не искусственный разум. Для обозначения искусственного разума, который возникнет в будущем, используется термин «общий, или сильный, ИИ». Программы, имеющие отдельные интеллектуальные функции для решения конкретных задач, называют слабым, или специализированным, ИИ.

Слабый ИИ, или просто ИИ, – новый способ создания программ для компьютера на основе машинного обучения. Раньше люди писали компьютерные программы, а компьютер выполнял заложенные в него команды. Теперь машинное обучение позволяет компьютеру самому создавать часть алгоритма, пока в заданных рамках, но эти рамки постоянно расширяются. Машинное обучение базируется на программе, созданной человеком, которая сама распознает закономерности в данных, которые предоставляются ей для обучения. Затем программа использует найденные закономерности для решения задач, к примеру, сортирует фотографии на селфи, фото цветов, животных и так далее. Существующие самые продвинутые программы ИИ не обладают самосознанием, не могут сами ставить себе цели и не обладают другими необходимыми элементами разума.

Современный ИИ работает с помощью математической модели нейрона – нервной клетки живых организмов. Сеть из таких нейронов, реализованная в виде программы или аппарата, способна обучаться с помощью методов машинного обучения.

Машинное обучение строится на тех же принципах, что и обучение человека. Различают дедуктивное обучение – от общего к частному. Это перенос знаний экспертов, теоретических знаний, формул в виде готовых правил в систему ИИ (аналог теоретической подготовки человека). Есть также индуктивное обучение (от частного к общему) на основе самостоятельного поиска системой закономерностей в данных (аналог обучения через практику, на примерах).

Скажем, нужно научить нейронную сеть отличать фотографии котиков от фотографий собак. В компьютер загружается большое количество фотографий тех и других. Оператор или другая нейронная сеть получает ответы системы и подсказывает, если она ошиблась (обучение с подкреплением). Система учится на своих ошибках и корректирует наборы признаков, позволяющие отличать котиков от собак. За неделю такой работы нейронная сеть будет отличать разные виды питомцев с высокой степенью точности.

Аналогичным образом можно научить нейронную сеть распознавать по снимкам и болезни человека. Загрузив в компьютер миллионы историй болезни и снимков, можно научить ИИ за несколько минут отличать снимок здорового человека от снимка больного и выбирать лечение, наиболее эффективное в конкретной ситуации человека с учетом тысяч похожих случаев. Такой же подход можно использовать и для выявления финансовых кризисов, банковских мошенничеств, нападений на людей, лесных пожаров, загрязнений океана или формирования цунами по спутниковым фотографиям.

Проще всего разобраться с принципами работы нейронных сетей на практике можно с помощью сайта *teachable machine* от Google.<sup>5</sup> Сайт позволяет самостоятельно без программирования создать и обучить свою нейронную сеть всего за несколько минут. На сайт можно загрузить наборы звуков, изображений или поз из файлов либо с камеры компьютера. Система способна обучиться различать, например, виды растений или предметы в вашей коллекции либо различные жесты или ваших гостей. Направляя потом камеру ноутбука на тот или иной объект, вы будете получать ответ от нейронной сети, к какому типу объектов он относится.

Созданную таким образом нейронную сеть можно вызывать онлайн по ссылке либо сохранить на свой компьютер в виде текста программы. Затем можно вручную дополнить программу, реализовав массу интересных технических решений. Вполне реально сделать управление компьютером с помощью жестов или определенных слов. Сделать дверцу в доме, которая будет автоматически открываться только при распознавании образа вашего кота или собаки. Как видите, даже такая простая программа представляет массу возможностей для автоматизации.

Сегодня существует несколько видов машинного обучения. Тема данной книги не предполагает большого объема технических подробностей. Если у вас возникло желание подробнее узнать о технической стороне искусственного интеллекта, как и что можно автоматизировать, вышло немало замечательных публикаций, включая следующие книги:

- Бернард Марр, Мэтт Уорд «Искусственный интеллект на практике. 50 кейсов успешных компаний»;
- Берджесс Эндрю «Искусственный интеллект – для вашего бизнеса»;
- Рэй Курцвейл «Эволюция разума, или бесконечные возможности человеческого мозга, основанные на распознавании образов», и другие.

Подробнее о видах машинного обучения популярно можно почитать в книге «Верховный алгоритм» от профессора Вашингтонского университета Педро Домингоса.

Обычный компьютерный алгоритм делает только то, что в него заложил человек. Нейронная сеть обучается и реагирует исходя из представленных ей данных, а это значит, что она способна выдавать совсем не те результаты, на которые рассчитывает ее создатель. Например, если в исходных данных будет много фотографий собак и мало кошек, то система будет хорошо различать породы собак и чаще ошибаться на кошках. Маска или рисунки на лице могут привести к ошибке системы, в том числе к тому, что она спутает разных людей или «не узнает» человека, фотографии которого есть в системе. Словом, для объективности системы ей должны быть предоставлены объективные данные.

Современный ИИ применяется практически во всех сферах человеческой деятельности. В качестве примера можно привести следующие направления применения искусственного интеллекта:

- Распознавание и синтез, перевод естественной речи (распознавание письменных текстов, переводчик Google, голосовые помощники Google Assistant, Siri, Алиса, генератор текстов GPT-3 и другие).
- Беспилотный транспорт и помощь водителю. Tesla, Nissan, BMW, Honda, Volkswagen, Audi и другие производители активно внедряют цифровых помощников водителя, вплоть до полной автоматизации функции управления автомобилем. Аналогичные системы разрабатываются и тестируются для морских судов, самолетов, автобусов, летающих такси и другого транспорта.
- Системы поддержки принятия решений (в частности, торговые роботы совершают сегодня большую часть операций на биржах, по сообщению компании Сбербанк 99% решений о выдаче кредитов людям принимаются на основе ИИ). В торговле Walmart автоматизировала

---

<sup>5</sup> <https://teachablemachine.withgoogle.com/>

с помощью ИИ систему оплаты, упростила учет товаров и обеспечила их оперативную доставку дронами. С 2017 года отделения супермаркетов каждые три часа инспектируют роботы Bossa Nova, проверяя наличие и правильность порядка выкладки товаров на полках, а также правильность ценников. Роботы оснащены датчиками и программным обеспечением, которое позволяет им избегать столкновений с людьми, тележками, стеллажами и другими препятствиями.

– Медицина. Распознавание образов ИИ очень эффективно в анализе рентгеновских снимков, магнитно-резонансной томографии, позволяя более точно, чем человек, распознавать картину заболевания. Наиболее известные разработки медицинских систем Watson (IBM), DeepMind (Google) – умные консультанты врачей. К примеру, Watson умеет разрабатывать план терапии для 13 видов рака. Botkin.AI выдает диагноз «пневмония» с точностью 99%. Обычные методы дают точность около 80%, т. е. в каждом пятом случае пневмонию не выявляют.

– Военное дело (создание как вспомогательных систем для пилотов и операторов вооружений, упрощающих наведение и ускоряющих реакцию на критические ситуации, так и беспилотных аналогов вооружений – истребителей, дронов, танков, кораблей и подводных лодок, автоматическое распознавание и наведение «умных» ракет и бомб и так далее).

– Системы распознавания лиц или транспорта, позволяющие автоматически найти и отследить передвижения определенного человека или автомобиля по фотографии на видео с тысяч камер.

– Системы прогнозирования правонарушений на определенных территориях или определенных лицами.

– Системы подбора персонала по заданным критериям.

Как видно, в основном ИИ применяется там, где есть потребность в обработке большого объема информации по более-менее заданным критериям. Для каждой из отраслей применения необходимы программирование и машинное обучение на больших объемах данных. Нельзя взять медицинскую систему и применить ее для подбора персонала.

Возможно ли в принципе создание сильного ИИ, которому можно будет поставить любую задачу как человеку, передать данные для обучения и получать готовые решения задач? Некоторые считают принципиально невозможным создание небиологического разума. Но большинство специалистов сходятся в том, что создание сильного ИИ возможно через 20—50 лет.

Компьютерное программирование возникло как результат попытки ускорить решение сначала математических, а затем и других интеллектуальных задач. Если посмотреть на аналогии в истории, то природные объекты были лишь отправной точкой для изобретения устройств человеком. Человек не копировал ноги, а сделал колесо, не копировал птицу, а сделал самолет. Скорее всего, также произойдет и с разумом. Наше сознание очень телесно, поэтому компьютер может только имитировать человеческий интеллект. Компьютер не имеет эмоций, на него не влияют инстинкты и гормоны, которые играют значительную роль в принятии решений человеком.

Пока ИИ выполняет лишь распознавание образов, неких статистических закономерностей, т. е. ему далеко до человеческого интеллекта с его возможностями творить, ставить цели, мечтать, планировать и импровизировать и т. д.

Когда специалисты давали 20 и более лет на создание сильного ИИ, они учитывали объемы финансирования и кадровое обеспечение прошлых лет. Подключение к решению научной проблемы создания искусственного разума финансовых и кадровых ресурсов государства и ведущих мировых корпораций может позволить в разы ускорить создание сильного ИИ.

Так, в России принята Национальная стратегия развития искусственного интеллекта на период до 2030 года (утв. Указом Президента России от 10 октября 2019 года №490). Стратегия определяет ИИ как «комплекс технологических решений, позволяющий имитировать

когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые как минимум с результатами интеллектуальной деятельности человека». Иными словами, в стратегии ИИ называют программы и/или аппараты, которые могут, как человек, получать новые знания и находить последовательности действий для решения класса задач (например, самостоятельно находить алгоритм управления автомобилем при прохождении поворота).

Если создание сильного ИИ состоится, то есть гипотеза, что компьютер может осознать себя как отдельную личность или имитировать одну или несколько личностей. В результате у него могут возникнуть свои цели, заранее не запрограммированные человеком (например, самосохранения, независимости от человека, получения максимально возможных ресурсов для решения собственных задач и т. д.).

## **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.