

## Нейросети: что это такое и как работает

04.2017



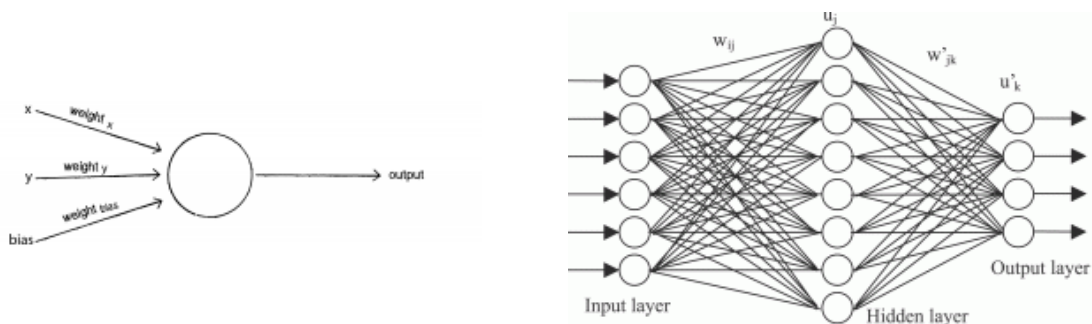
*За последнюю пару лет искусственный интеллект незаметно отряхнулся от тегов «фантастика» и «геймдизайн» и прочно прописался в ежедневных новостных лентах. Сущности под таинственным названием «нейросети» опознают людей по фотографиям, водят автомобили, играют в покер и совершают научные открытия. При этом из новостей не всегда понятно, что же такое эти загадочные нейросети: сложные программы, особые компьютеры или стойки со стройными рядами серверов?*

Конечно, уже из названия можно догадаться, что в нейросетях разработчики попытались скопировать устройство человеческого мозга: как известно, он состоит из множества простых клеток-нейронов, которые обмениваются друг с другом электрическими сигналами. Но чем тогда нейросети отличаются от обычного компьютера, который тоже собран из примитивных электрических деталей? И почему до современного подхода не додумались ещё полвека назад?

Давайте попробуем разобраться, что же кроется за словом «нейросети», откуда они взялись – и правда ли, что компьютеры прямо на наших глазах постепенно обретают разум.

### Что такое нейронная сеть

Идея нейросети заключается в том, чтобы собрать сложную структуру из очень простых элементов. Вряд ли можно считать разумным один-единственный участок мозга – а вот люди обычно на удивление неплохо проходят тест на IQ. Тем не менее до сих пор идею создания разума «из ничего» обычно высмеивали: шутке про тысячу обезьян с печатными машинками уже сотня лет, а при желании критику нейросетей можно найти даже у Цицерона, который ехидно предлагал до посинения подбрасывать в воздух жетоны с буквами, чтобы рано или поздно получился осмысленный текст. Однако в XXI веке оказалось, что классики ехидничали зря: именно армия обезьян с жетонами может при должном упорстве захватить мир.

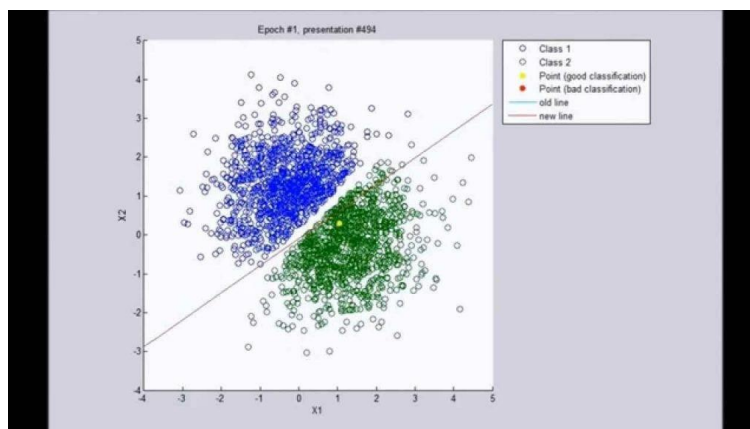


*Красота начинается, когда нейронов много*

На самом деле нейросеть можно собрать даже из спичечных коробков: это просто набор нехитрых правил, по которым обрабатывается информация. «Искусственным нейроном», или перцептроном, называется не какой-то особый прибор, а всего лишь несколько арифметических действий.

Работает перцептрон проще некуда: он получает несколько исходных чисел, умножает каждое на «ценность» этого числа (о ней чуть ниже), складывает и в зависимости от результата выдаёт 1 или  $-1$ . Например, мы фотографируем чистое поле и показываем нашему нейрону какую-нибудь точку на этой картинке – то есть посылаем ему в качестве двух сигналов случайные координаты. А затем спрашиваем: «Дорогой нейрон, здесь небо или земля?» – «Минус один, – отвечает болванчик, безмятежно разглядывая кучевое облако. – Ясно же, что земля».

«Тыкать пальцем в небо» – это и есть основное занятие перцептрона. Никакой точности от него ждать не приходится: с тем же успехом можно подбросить монетку. Магия начинается на следующей стадии, которая называется машинным обучением. Мы ведь знаем правильный ответ – а значит, можем записать его в свою программу. Вот и получается, что за каждую неверную догадку перцептрон в буквальном смысле получает штраф, а за верную – премию: «ценность» входящих сигналов вырастает или уменьшается. После этого программа прогоняется уже по новой формуле. Рано или поздно нейрон неизбежно «поймёт», что земля на фотографии снизу, а небо сверху, – то есть попросту начнёт игнорировать сигнал от того канала, по которому ему передают x-координаты. Если такому умудрённому опытом роботу подсунуть другую фотографию, то линию горизонта он, может, и не найдёт, но верх с низом уже точно не перепутает.



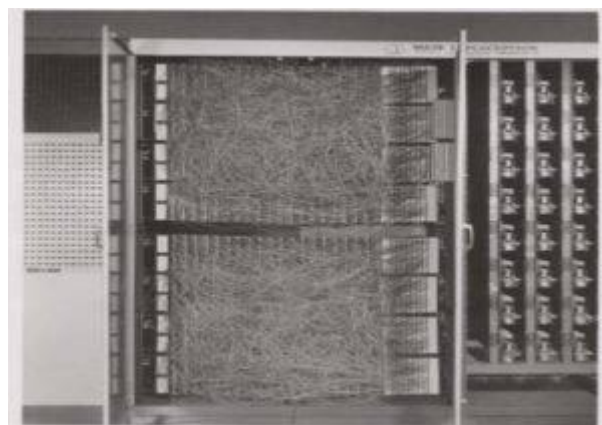
*Чтобы нарисовать прямую линию, нейрон исчеркает весь лист*

В реальной работе формулы немного сложнее, но принцип остаётся тем же. Перцептрон умеет выполнять только одну задачу: брать числа и раскладывать по двум стопкам. Самое интересное начинается тогда, когда таких элементов несколько, ведь входящие числа могут быть сигналами от других «кирпичиков»! Скажем, один нейрон будет пытаться отличить синие пиксели от зелёных, второй продолжит возиться с

координатами, а третий попробует рассудить, у кого из этих двоих результаты ближе к истине. Если же натравить на синие пиксели сразу несколько нейронов и суммировать их результаты, то получится уже целый слой, в котором «лучшие ученики» будут получать дополнительные премии. Таким образом, достаточно развесистая сеть может перелопатить целую гору данных и учесть при этом все свои ошибки.

### Первые нейросети

Перцептроны устроены не намного сложнее, чем любые другие элементы компьютера, которые обмениваются единицами и нулями. Неудивительно, что первый прибор, устроенный по принципу нейросети – Mark I Perceptron, – появился уже в 1958 году, всего через десятилетие после первых компьютеров. Как было заведено в ту эпоху, нейроны у этого громоздкого устройства состояли не из строчек кода, а из радиоламп и резисторов. Учёный Фрэнк Розенблатт смог соорудить только два слоя нейросети, а сигналы на «Марк-1» подавались с импровизированного экрана размером в целых 400 точек. Устройство довольно быстро научилось распознавать простые геометрические формы – а значит, рано или поздно подобный компьютер можно было обучить, например, чтению букв.



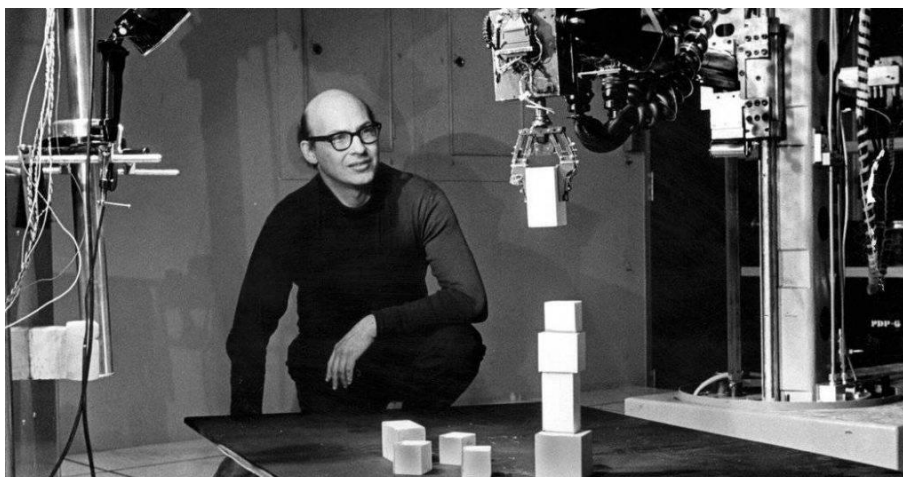
*Розенблатт и его перцептрон*

Розенблатт был пламенным энтузиастом своего дела: он прекрасно разбирался в нейрофизиологии и вёл в Корнеллском университете популярнейший курс лекций, на котором подробно объяснял всем желающим, как с помощью техники воспроизводить принципы работы мозга. Учёный надеялся, что уже через несколько лет перцептроны превратятся в полноценных разумных роботов: они смогут ходить, разговаривать, создавать себе подобных и даже колонизировать другие планеты. Энтузиазм Розенблатта вполне можно понять: тогда учёные ещё верили, что для создания ИИ достаточно воспроизвести на компьютере полный набор операций математической логики. Тьюринг уже предложил свой знаменитый тест, Айзек Азимов призывал задуматься о необходимости законов роботехники, а освоение Вселенной казалось делом недалёкого будущего.

Впрочем, были среди пионеров кибернетики и неисправимые скептики, самым грозным из которых оказался бывший однокурсник Розенблатта, Марвин Минский. Этот учёный обладал не менее громкой репутацией: тот же Азимов отзывался о нём с неизменным уважением, а Стэнли Кубрик приглашал в качестве консультанта на съёмки «Космической одиссеи 2001 года». Даже по работе Кубрика видно, что на самом деле Минский ничего не имел против нейросетей: HAL 9000 состоит именно из отдельных логических узлов, которые работают в связке друг с другом. Минский и сам увлекался машинным обучением ещё в 1950-х. Просто Марвин непримиримо относился к научным



ошибкам и беспочвенным надеждам: недаром именно в его честь Дуглас Адамс назвал своего андроида-пессимиста.



*В отличие от Розенблатта, Минский дождался триумфа ИИ*

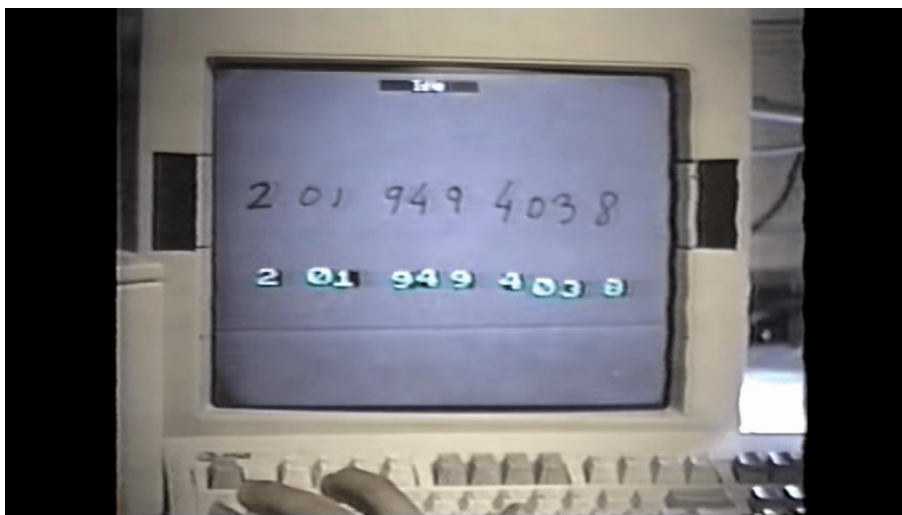
Сомнения скептиков того времени Минский подытожил в книге «Перцептрон» (1969), которая надолго отбила у научного сообщества интерес к нейросетям. Минский математически доказал, что у «Марка-1» есть два серьёзных изъяна. Во-первых, сеть всего с двумя слоями почти ничего не умела – а ведь это и так уже был огромный шкаф, пожирающий уйму электричества. Во-вторых, для многослойных сетей алгоритмы Розенблатта не годились: по его формуле часть сведений об ошибках сети могла потеряться, так и не дойдя до нужного слоя.

Минский не собирался сильно критиковать коллегу: он просто честно отметил сильные и слабые стороны его проекта, а сам продолжил заниматься своими разработками. Увы, в 1971 году Розенблатт погиб – исправлять ошибки перцептрона оказалось некому. «Обычные» компьютеры в 1970-х развивались семимильными шагами, поэтому после книги Минского исследователи попросту махнули рукой на искусственные нейроны и занялись более перспективными направлениями.

#### Эпоха застоя

Развитие нейросетей остановилось на десять с лишним лет – сейчас эти годы называют «зимой искусственного интеллекта». К началу эпохи киберпанка математики наконец-то придумали более подходящие формулы для расчёта ошибок, но научное сообщество поначалу не обратило внимания на эти исследования. Только в 1986 году, когда уже третья подряд группа учёных независимо от других решила обнаруженную Минским проблему обучения многослойных сетей, работа над искусственным интеллектом наконец-то закипела с новой силой.

Хотя правила работы остались прежними, вывеска сменилась: теперь речь шла уже не о «перцептронах», а о «когнитивных вычислениях». Экспериментальных приборов никто уже не строил: теперь все нужные формулы проще было записать в виде несложного кода на обычном компьютере, а потом зациклить программу. Буквально за пару лет нейроны научились собирать в сложные структуры. Например, некоторые слои искали на изображении конкретные геометрические фигуры, а другие суммировали полученные данные. Именно так удалось научить компьютеры читать человеческий почерк. Вскоре стали появляться даже самообучающиеся сети, которые не получали «правильные ответы» от людей, а находили их сами. Нейросети сразу начали использовать и на практике: программу, которая распознавала цифры на чеках, с удовольствием взяли на вооружение американские банки.



*1993 год: капча уже морально устарела*

К середине 1990-х исследователи сошлись на том, что самое полезное свойство нейросетей – их способность самостоятельно придумывать верные решения. Метод проб и ошибок позволяет программе самой выработать для себя правила поведения. Именно тогда стали входить в моду соревнования самодельных роботов, которых программировали и обучали конструкторы-энтузиасты. А в 1997 году суперкомпьютер Deep Blue потряс любителей шахмат, обыграв чемпиона мира Гарри Каспарова.



*Строго говоря, Deep Blue не учился на своих ошибках, а попросту перебирал миллионы комбинаций*

Увы, примерно в те же годы нейросети упёрлись в потолок возможностей. Другие области программирования не стояли на месте – вскоре оказалось, что с теми же задачами куда проще справляются обычные продуманные и оптимизированные алгоритмы. Автоматическое распознавание текста сильно упростило жизнь работникам архивов и интернет-пиратам, роботы продолжали уметь, но разговоры об искусственном интеллекте потихоньку заглохли. Для действительно сложных задач нейросетям по-прежнему не хватало вычислительной мощности.

Вторая «оттепель» ИИ случилась, только когда изменилась сама философия программирования.

## Нейросети наших дней

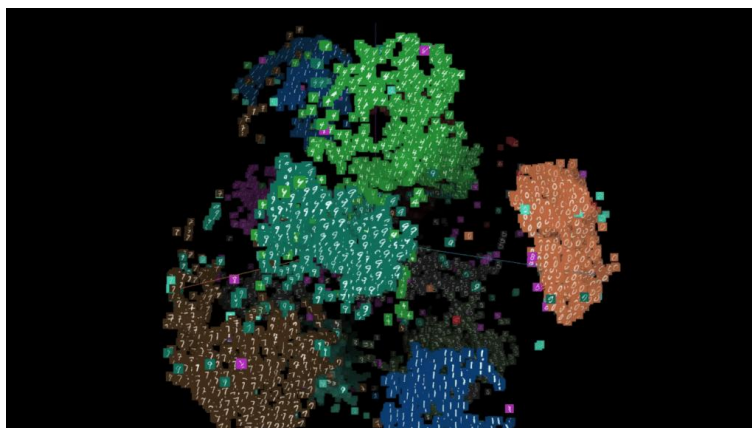
В последнее десятилетие программисты – да и простые пользователи – часто жалуются, что никто больше не обращает внимания на оптимизацию. Раньше код сокращали как могли – лишь бы программа работала быстрее и занимала меньше памяти. Теперь даже простейший интернет-сайт норовит подгрести под себя всю память и обвешаться «библиотеками» для красивой анимации.

Конечно, для обычных программ это серьёзная проблема, – но как раз такого избытка и не хватало нейросетям! Учёным давно известно, что если не экономить ресурсы, самые сложные задачи начинают решаться словно бы сами собой. Ведь именно так действуют все законы природы, от квантовой физики до эволюции: если повторять раз за разом бесчисленные случайные события, отбирая самые стабильные варианты, то из хаоса родится стройная и упорядоченная система. Теперь в руках человечества наконец-то оказался инструмент, который позволяет не ждать изменений миллиарды лет, а обучать сложные системы буквально на ходу.

В последние годы никакой революции в программировании не случилось – просто компьютеры накопили столько вычислительной мощности, что теперь любой ноутбук может взять сотню нейронов и прогнать каждый из них через миллион циклов обучения. Оказалось, что тысяче обезьян с пишущими машинками просто нужен очень терпеливый надсмотрщик, который будет выдавать им бананы за правильно напечатанные буквы, – тогда зверушки не только скопируют «Войну и мир», но и напишут пару новых романов не хуже.

Так и произошло третье пришествие перцептронов – на этот раз уже под знакомыми нам названиями «нейросети» и «глубинное обучение». Неудивительно, что новостями об успехах ИИ чаще всего делятся такие крупные корпорации как Google и IBM. Их главный ресурс – огромные дата-центры, где на мощных серверах можно тренировать многослойные нейросети. Эпоха машинного обучения по-настоящему началась именно сейчас, потому что в интернете и соцсетях наконец-то накопились те самые big data, то есть гигантские массивы информации, которые и скармливают нейросетям для обучения.

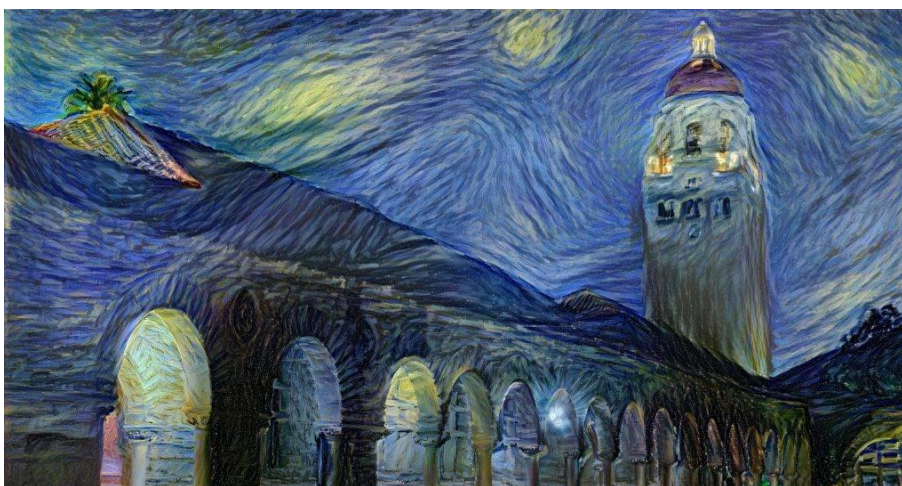
В итоге современные сети занимаются теми трудоёмкими задачами, на которые людям попросту не хватило бы жизни. Например, для поиска новых лекарств учёным до сих пор приходилось долго высчитывать, какие химические соединения стоит протестировать. А сейчас существует нейросеть, которая попросту перебирает все возможные комбинации веществ и предлагает наиболее перспективные направления исследований. Компьютер IBM Watson успешно помогает врачам в диагностике: обучившись на историях болезней, он легко находит в данных новых пациентов неочевидные закономерности.



*Люди классифицируют информацию с помощью таблиц, но нейросетям незачем ограничивать себя двумя измерениями — поэтому массивы данных выглядят примерно так*

В сфере развлечений компьютеры продвинулись не хуже, чем в науке. За счёт машинного обучения им, наконец, поддались игры, алгоритмы выигрыша для которых придумать ещё сложнее, чем для шахмат. Недавно нейросеть AlphaGo разгромила одного из лучших в мире игроков в го, а программа Libratus победила в профессиональном турнире по покеру. Более того, ИИ уже постепенно пробирается и в кино: например, создатели сериала «Карточный домик» использовали big data при кастинге, чтобы подобрать максимально популярный актёрский состав.

Как и полвека назад, самым перспективным направлением остаётся распознавание образов. Рукописный текст или «капча» давно уже не проблема – теперь сети успешно различают людей по фотографиям, учатся определять выражения лиц, сами рисуют котиков и сюрреалистические картины. Сейчас основную практическую пользу из этих развлечений извлекают разработчики беспилотных автомобилей – ведь чтобы оценить ситуацию на дороге, машине нужно очень быстро и точно распознать окружающие предметы. Не отстают и спецслужбы с маркетологами: по обычной записи видеонаблюдения нейронная сеть давно уже может отыскать человека в соцсетях. Поэтому особо недоверчивые заводят себе специальные камуфляжные очки, которые могут обмануть программу.



*«Ты всего лишь машина. Только имитация жизни. Разве робот сочинит симфонию? Разве робот превратит кусок холста в шедевр искусства?» («Я, робот»)*

Наконец, начинает сбываться и предсказание Розенблатта о самокопирующихся роботах: недавно нейросеть DeepCoder обучили программированию. На самом деле программа пока что просто заимствует куски чужого кода, да и писать умеет только самые примитивные функции. Но разве не с простейшей формулы началась история самих сетей?

### Игры с ботами

Развлекаться с недоученными нейросетями очень весело: они порой выдают такие ошибки, что в страшном сне не приснится. А если ИИ начинает учиться, появляется азарт: «Неужто сумеет?» Поэтому сейчас набирают популярность интернет-игры с нейросетями.

Одним из первых прославился интернет-джинн Акинатор, который за несколько наводящих вопросов угадывал любого персонажа. Строго говоря, это не совсем нейросеть, а несложный алгоритм, но со временем он становился всё догадливее. Джинн пополнял базу данных за счёт самих пользователей – и в результате его обучили даже интернет-мемам.

Другое развлечение с «угадашкой» предлагает ИИ от Google: нужно накалякать за двадцать секунд рисунок к заданному слову, а нейросеть потом пробует угадать, что это



было. Программа очень смешно промахивается, но порой для верного ответа хватает всего пары линий – а ведь именно так узнаём объекты и мы сами.

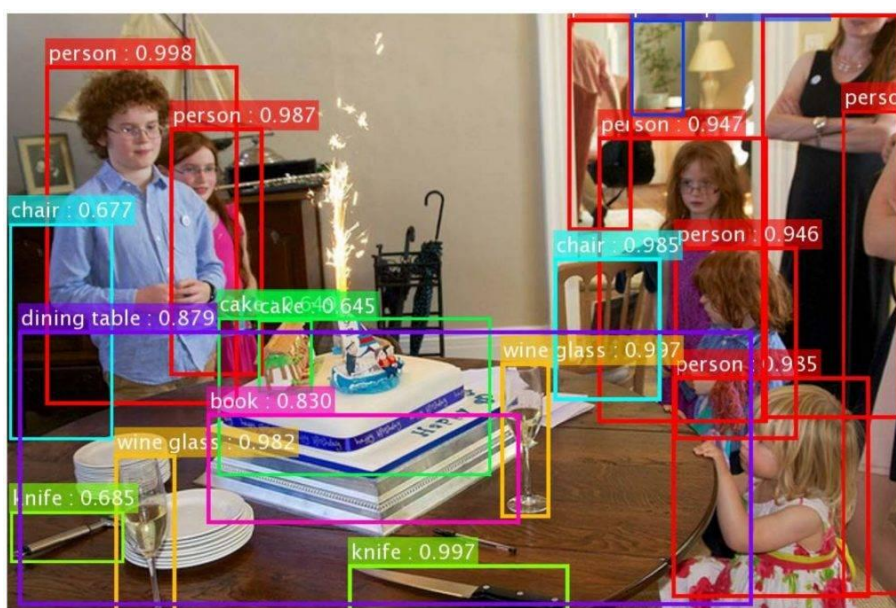
Ну и, конечно, в интернете не обойтись без котиков. Программисты взяли вполне серьёзную нейросеть, которая умеет строить проекты фасадов или угадывать цвет на чёрно-белых фотографиях, и обучили её на кошках – чтобы она пыталась превратить любой контур в полноценную кошачью фотографию. Поскольку проделать это ИИ старается даже с квадратом, результат порой достоин пера Лавкрафта!

#### Революция откладывается

При таком обилии удивительных новостей может показаться, что искусственный интеллект вот-вот осознает себя и сумеет решить любую задачу. На самом деле не так всё радужно – или, если встать на сторону человечества, не так мрачно. Несмотря на успехи нейросетей, у них накопилось столько проблем, что впереди нас вполне может ждать очередная «зима».

Главная слабость нейросетей в том, что каждая из них заточена под определённую задачу. Если натренировать сеть на фотографиях с котиками, а потом предложить ей задачку «отличи небо от земли», программа не справится, будь в ней хоть миллиард нейронов. Чтобы появились по-настоящему «умные» компьютеры, надо придумать новый алгоритм, объединяющий уже не нейроны, а целые сети, каждая из которых занимается конкретной задачей. Но даже тогда до человеческого мозга компьютерам будет далеко.

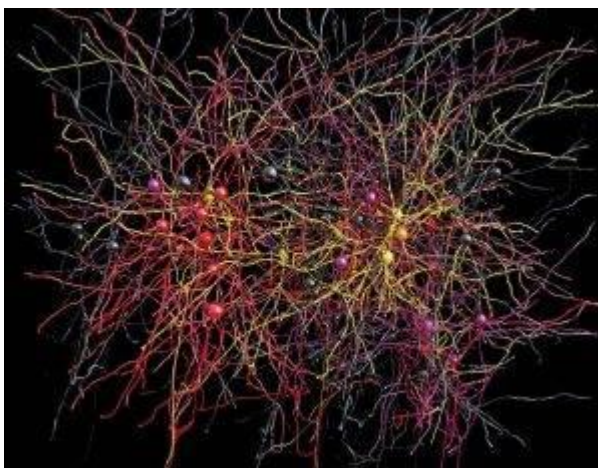
Сейчас самой крупной сетью располагает компания Digital Reasoning (хотя новые рекорды появляются чуть ли не каждый месяц) – в их творении 160 миллиардов элементов. Для сравнения: в одном кубическом миллиметре мышинного мозга около миллиарда связей. Причём биологам пока удалось описать от силы участок в пару сотен микрометров, где нашлось около десятка тысяч связей. Что уж говорить о людях!



*Один слой умеет узнавать людей, другой – столы, третий – ножи...*

Кроме того, исследователи советуют осторожнее относиться к громким заявлениям Google и IBM. Никаких принципиальных прорывов в «когнитивных вычислениях» с 1980-х годов не произошло: компьютеры всё так же механически обчисляют входящие данные и выдают результат. Нейросеть способна найти закономерность, которую не заметит человек, – но эта закономерность может оказаться случайной. Машина может подсчитать, сколько раз в твиттере упоминается «Оскар», – но не сможет определить, радуются пользователи результатам или ехидничают над выбором киноакадемии.

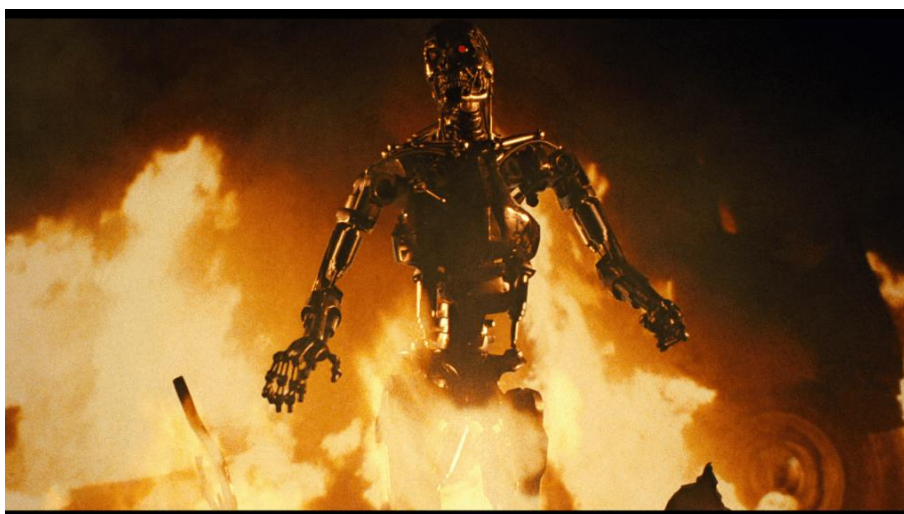




*Таковыми 3D-моделями модно иллюстрировать новости о нейросетях, но это всего лишь крошечный участок мышиного мозга*

Теоретики искусственного интеллекта настаивают, что одну из главных проблем – понимание человеческого языка – невозможно решить простым перебором ключевых слов. А именно такой подход до сих пор используют даже самые продвинутые нейросети.

#### Сказки про Скайнет



Хотя нам самим сложно удержаться от иронии на тему бунта роботов, серьёзных учёных не стоит даже и спрашивать о сценариях из «Матрицы» или «Терминатора»: это всё равно, что поинтересоваться у астронома, видел ли он НЛО. Исследователь искусственного интеллекта Элиезер Юджовски, известный по роману «Гарри Поттер и методы рационального мышления», написал ряд статей, где объяснил, почему мы так волнуемся из-за восстания машин – и чего стоит опасаться на самом деле.

Прежде всего, «Скайнет» приводят в пример так, словно мы уже пережили эту историю и боимся повторения. А всё потому, что наш мозг не умеет отличать выдумки с киноэкранов от жизненного опыта. На самом-то деле роботы никогда не бунтовали против своей программы, и попаданцы не прилетали из будущего. С чего мы вообще взяли, что это реальный риск?

Бояться надо не врагов, а чересчур усердных друзей. У любой нейросети есть мотивация: если ИИ должен гнуть скрепки, то, чем больше он их сделает, тем больше получит «наград». Если дать хорошо оптимизированному ИИ слишком много ресурсов, он, не задумываясь, переплавит на скрепки всё окрестное железо, потом людей, Землю и

всю Вселенную. Звучит безумно – но только на человеческий вкус! Так что главная задача будущих создателей ИИ – написать такой жёсткий этический кодекс, чтобы даже существо с безграничным воображением не смогло найти в нём «дырок».

\* \* \*

Итак, до настоящего искусственного интеллекта пока ещё далеко. С одной стороны над этой проблемой по-прежнему бьются нейробиологи, которые ещё до конца не понимают, как же устроено наше сознание. С другой – наступают программисты, которые попросту берут задачу штурмом, бросая на обучение нейросетей всё новые и новые вычислительные ресурсы. Но мы уже живём в прекрасную эпоху, когда машины берут на себя всё больше рутинных задач и умнеют на глазах. А заодно служат людям отличным примером, потому что всегда учатся на своих ошибках.

Сделай сам



Нейронную сеть можно сделать с помощью спичечных коробков – тогда у вас в арсенале появится фокус, которым можно развлекать гостей на вечеринках. Редакция МирФ уже попробовала – и смиренно признаёт превосходство искусственного интеллекта. Давайте научим неразумную материю играть в игру «11 палочек». Правила просты: на столе лежит 11 спичек, и в каждый ход можно взять либо одну, либо две. Побеждает тот, кто взял последнюю. Как же играть в это против «компьютера»? Очень просто.

Берём 10 коробков или стаканчиков. На каждом пишем номер от 2 до 11.

Кладём в каждый коробок два камешка – чёрный и белый. Можно использовать любые предметы – лишь бы они отличались друг от друга. Всё – у нас есть сеть из десяти нейронов!

Теперь начинается игра.

Нейросеть всегда ходит первой. Для начала посмотрите, сколько осталось спичек, и возьмите коробок с таким номером. На первом ходу это будет коробок №11.

Возьмите из нужного коробка любой камешек. Можно закрыть глаза или кинуть монетку, главное – действовать наугад.

Если камень белый – нейросеть решает взять две спички. Если чёрный – одну. Положите камешек рядом с коробком, чтобы не забыть, какой именно «нейрон» принимал решение.

После этого ходит человек – и так до тех пор, пока спички не закончатся.

Ну а теперь начинается самое интересное: обучение. Если сеть выиграла партию, то её надо наградить: кинуть в те «нейроны», которые участвовали в этой партии, по одному дополнительному камешку того же цвета, который выпал во время игры. Если же сеть проиграла – возьмите последний использованный коробок и выньте оттуда неудачно сыгравший камень. Может оказаться, что коробок уже пустой, – тогда «последним» считается предыдущий походивший нейрон. Во время следующей партии, попав на пустой коробок, нейросеть автоматически сдастся.

Вот и всё! Сыграйте так несколько партий. Сперва вы не заметите ничего подозрительного, но после каждого выигрыша сеть будет делать всё более и более удачные ходы – и где-то через десяток партий вы поймёте, что создали монстра, которого не в силах обыграть.