



ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: ПРОШЛОЕ, НАСТОЯЩЕЕ, БУДУЩЕЕ

Часть 1

Современная революция в области искусственного интеллекта

За последние десять лет наши оценки состояния и перспектив технологий в области искусственного интеллекта (ИИ) резко изменились. Еще относительно недавно мы говорили о «зиме искусственного интеллекта», подразумевая под этим то, что завышенные ожидания и надежды здесь оказались обманчивыми, а искусственные алгоритмы и программы почти во всех интеллектуальных задачах качественно отстают от возможностей человека и, возможно, никогда не смогут достичь человеческого уровня. Сегодня же мы можем в огромном количестве задач продемонстрировать не только человеческий, но даже «сверхчеловеческий» [superhuman] уровень результатов ИИ. В начале нынешнего века многие достаточно обоснованно предполагали, что «слабый» (или «функциональный») ИИ, необходимый для создания, например, автономных роботов, возникнет не раньше, чем вычислительные мощности искусственных процессоров сравнятся с мощностью человеческого мозга (и этот прогноз относили, согласно закону Мура, куда-то в район 2040 года). Сегодня, в 2020 году, мы можем уверенно констатировать, что функциональный ИИ уже практически создан, и это открывает перед нами фантастические перспективы дальнейшего технического и технологического развития.

Такие перемены сложно назвать иначе как *технологической революцией*. И революция эта была связана с появлением и развитием глубоких нейронных сетей (ГНС). Сегодня технологии на основе ГНС являются одними из наиболее многообещающих и прорывных средств изменения будущего.

Мне представляется, что внедрение ГНС и методов машинного обучения в ближайшие годы настолько же повлияет на все стороны нашей жизни, науки, производства и быта, как некогда внедрение цифровых ЭВМ и программирования.

Реальность XXI века во многом будет определяться уровнем достижений в области ИИ. И в этой изменившейся реальности России потребуются технологическая независимость и технологический паритет в сфере интеллектуальных технологий. Не секрет, что сегодня мировыми лидерами в развитии искусственного интеллекта являются США и Китай. В силу целого ряда финансовых, экономических и других причин Россия не сможет выйти на тот же уровень, просто повторяя пройденный ими путь. Нам нужен асимметричный ответ, иная стратегия, другие инструменты. Сегодня многие научные коллективы в нашей стране работают над решением этой задачи. В частности, мы надеемся, что технологической основой для широкого внедрения отечественных интеллектуальных решений может стать созданная нами унифицированная отечественная программная платформа нейросетевой разработки Plat.

Но чтобы рассказать о том, какие задачи стоят перед нами сегодня, нужно отступить на шаг назад и начать с того, что такое искусственный интеллект, как он развивался, и почему на острие той технологической революции, которую мы наблюдаем сегодня, оказались именно нейронные сети.

Моделирование рассуждений и машинное обучение

Термин «искусственный интеллект» часто трактуется как полный аналог «естественного» человеческого разума. То есть, речь идет о создании думающих и чувствующих машин, обладающих сознанием и самосознанием, внутренний мир которых подобен внутреннему миру людей. Такой ИИ принято называть «сильным». Мы к нему еще вернемся в конце нашего обсуждения, но та технологическая революция, о которой мы говорили выше, не потребовала создания «сильного» ИИ. Для решения практических задач на уровне человека и даже выше

достаточно оказалось «слабого» или «функционального» искусственного интеллекта, представляющего собой вычислительные устройства или компьютерные программы, способные автоматически выполнять функции, которые ранее могли быть выполнены только человеком.

В области ИИ всегда присутствовало два основных направления. Первое – искусственный интеллект, основанный на знаниях, логическом выводе и имитации человеческих рассуждений. Второе – машинное обучение, основанное на оптимизации целевых функций и обучении на примерах, а не на копировании человеческого мышления. Каждое из этих направлений сыграло свою роль.

Первый подход в истории науки можно проследить вплоть до Аристотеля и Евклида. В самом деле, как мы знаем, в основе геометрии Евклида находятся система постулатов и правила логического вывода, при помощи которых мы можем доказывать какие-то утверждения, то есть получать новые знания на основе существующих. Именно такой способ «машинного мышления», основанный на автоматизированных логических рассуждениях, и попытались реализовать в 1950-е годы, когда впервые заговорили о системах искусственного интеллекта. Предполагалось, что для любой области человеческой деятельности можно определить некоторую систему аксиом, после чего, используя универсальный автомат логического вывода, получать от ИИ ответы на любые возможные вопросы, касающиеся этой прикладной области. Благо, математическая основа для такого подхода в виде алгебры логики и исчисления предикатов была создана еще в докомпьютерную эру, а с появлением ЭВМ автоматизация логического вывода стала действительно делом техники. И успехи на этом пути действительно были достаточно впечатляющими.

Например, компьютеры, основываясь на логических правилах и машинном расчете возможных будущих позиций, научились играть в шахматы и шашки. В 1960-1970 годы шахматные программы уже вполне могли соперничать с шахматистами-мастерами, а в 1997 году компьютер IBM Deep Blue впервые обыграл чемпиона мира по шахматам Гарри Каспарова.

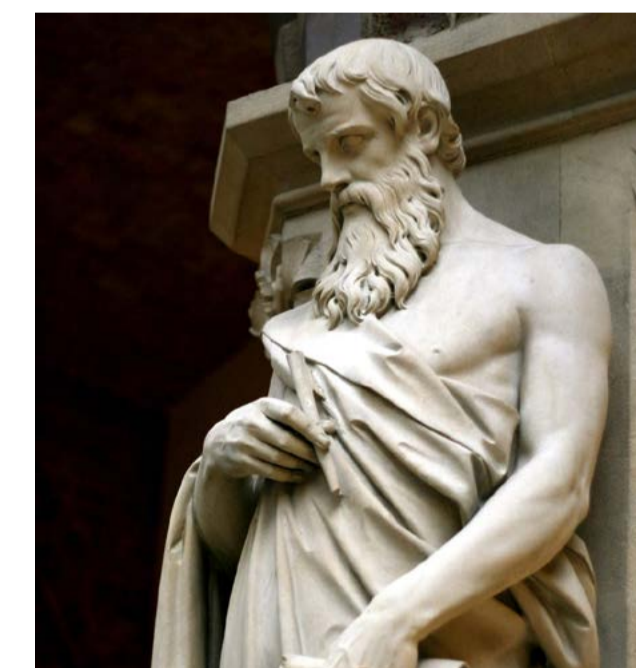
Однако вскоре оказалось, что эйфория от успехов логиче-



Юрий ВИЗИЛЬТЕР,
начальник подразделения
интеллектуального анализа данных и
технического зрения ФГУП «ГосНИИАС»,
доктор физико-математических наук,
профессор РАН



Аристотель (IV век до н. э.)



Евклид (IV-III век до н. э.)

ского ИИ была преждевременной. За пределами математики и условных игр по правилам добыть достаточно полные и непротиворечивые базовые знания о мире было крайне непросто: проблему подбора удачной системы аксиом для реальных задач мы до сих пор как следует решать не умеем. Кроме того, логический вывод, например, автоматическое доказательство новых теорем, сам по себе является задачей крайне трудоемкой. Для компьютерной реализации таких интеллектуальных систем было придумано логическое программирование. Наиболее известный язык логического программирования называется Пролог (название происходит от французского *Programmation en Logique*). Пролог оперирует фактами, правилами логического вывода и запросами, что позволяет описывать базы знаний, процедуры формальных рассуждений и принятия решений. Автоматический логический вывод в Прологе и других подобных системах основан на алгоритме поиска с возвратом. Идею поиска с возвратом легко представить себе на примере классической задачи о восьми ферзях. Как расставить на 64-клеточной шахматной доске 8 ферзей так, чтобы ни один из них не «бил» другого? Ферзь, как известно, бьет все поля по вертикали, горизонтали и диагонали относительно той клетки, на которой стоит. Вертикалей и горизонталей восемь, поэтому, очевидно, что больше 8 ферзей так поставить нельзя. Попробуем это сделать. Поставим сначала на доску одного ферзя (в любую точку – все поля еще свободны). Затем следующего (на такое поле, которое первый не бьет), затем следующего (на такое поле, которое не бьет предыдущие), и так далее – до тех пор, пока имеется хотя бы одна «небитая» клетка. Если таких клеток не осталось, а ферзей на поле меньше 8, придется вернуться на шаг назад и попытаться изменить положение предыдущего установленного ферзя, а потом вновь попытаться поставить еще одного ферзя на появившееся небитое поле. Если этого оказалось недостаточно, нужно вернуться еще на шаг назад – к положению предыдущего ферзя... Легко увидеть, что перед нами на самом деле способ полного перебора всех возможных позиций, только он организован так, чтобы можно было написать алгоритм упорядоченного поиска вариантов (например, программу для ЭВМ).

Именно таким способом и происходит автоматический подбор цепочки логических переходов от аксиом к теореме. Если сложность задачи небольшая, перебор вариантов занимает не очень много времени. Но если сложность задачи растет, то и сложность решения методом поиска с возвратом растет, причем экспоненциально – все равно, что от доски с 64 клетками мы переходим к доске с миллионами клеток. Ни один современный и даже мыслимый в будущем компьютер не в состоянии решать таким способом серьезные логические задачи за разумное время. Таким образом, следуя по этому пути, исследователи в области ИИ оказались в серьезном тупике, связанном с принципиальными вычислительными ограничениями. Нужно было искать обходной путь к искусственному интеллекту.

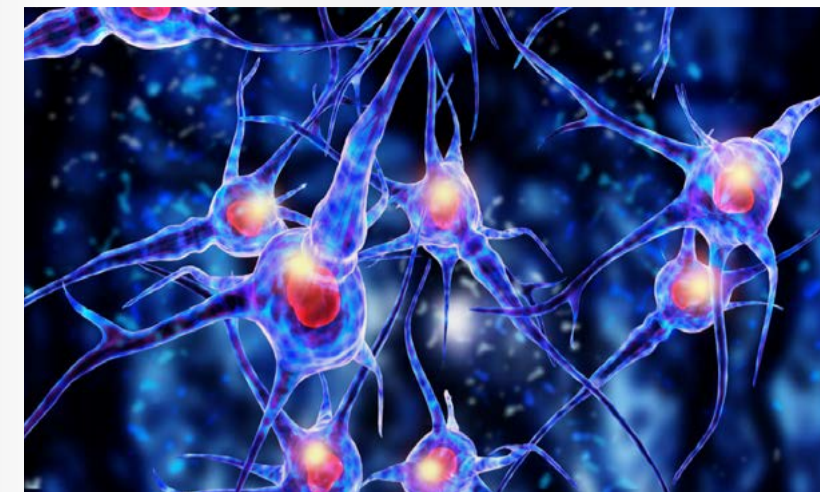
Между тем, параллельно с первым направлением ИИ, основанным на воспроизведении человеческих рассуждений, развивалось и второе направление, связанное с машинным обучением. Задача машинного обучения состоит в том, чтобы из экспериментально наблюдаемых фактов попытаться автоматически извлекать некое обобщенное знание, которое компьютеры смогут в дальнейшем использовать для решения новых задач за пределами тех данных, на которых происходило обучение. И если

выше мы говорили о Евклиде с его системой аксиом и логических построений, то в связи с машинным обучением необходимо, видимо, упомянуть Декарта и его идею системы координат. Используя координаты, мы можем описывать геометрические объекты не только словами или рисунками, но и в численном виде. А что если мы могли бы описать наборами чисел не только точки пространства, но вообще любые объекты, с которыми мы имеем дело на практике? Тогда мы могли бы вывести из опыта наблюдений математические «уравнения» и «неравенства» не только для прямых и окружностей, но и для таких семантических классов, как «человек» или «самолет», что позволило бы их автоматически распознавать и вообще работать с высокими интеллектуальными материями как с обычными инженерными расчетами. Эта идея «координатизации мира», описания его числами – не менее общая, чем идея описания мира базами знаний, состоящими из логических или словесных описаний – оказалась также очень удобной для реализации на ЭВМ и легла в основу систем машинного обучения.

Машинное обучение всегда нацелено на решение конкретной прикладной задачи. Например, с точки зрения постановки медицинского диагноза человека можно охарактеризовать некоторым набором признаков – значениями измеренных давления, температуры, содержания в крови определенных веществ, результатами других анализов. Так мы получаем для каждого человека свой вектор чисел (признаков), и, тем самым, описываем его некоторой точкой в соответствующем декартовом пространстве (пространстве признаков, в данном случае – медицинских). Как теперь поставить диагноз в этом условном воображаемом пространстве? Для этого нужно построить такую разделяющую поверхность (границу) в пространстве признаков, чтобы по одну сторону от нее находились примеры «больных», а по другую – примеры «здоровых». Такое уравнение (точнее, неравенство), разделяющее интересующие нас классы в пространстве признаков, называется классификатором. При этом алгоритмы машинного обучения вычисляют классификаторы, не пытаясь угадать, как решал бы такую задачу классификации человек. Они всего лишь пытаются найти в заданном классе решений такую границу между классами, которая обеспечивает минимальное число ошибок классификации на обучающей выборке примеров. Так, без обращения к экспертному опыту людей, базам знаний и логическим рассуждениям, могут решаться многие практические задачи, имитирующие различные функции человека. Таков путь машинного обучения на примерах, и именно на этом пути, в качестве одного из способов построения обучаемых классификаторов, были придуманы искусственные нейронные сети.

Искусственные нейронные сети и начало революции ИИ

Искусственные нейронные сети появились как простейшая математическая модель естественных нейронных сетей, которые отвечают за нервную деятельность, а значит, и за мышление у человека. Нейрофизиологи и врачи справедливо скажут, что настоящий человеческий нейрон – сложнейший объект, изучение которого и сегодня еще далеко от завершения. Но все же некоторые базовые



принципы его работы известны достаточно неплохо. Описывая нейроны как биологические устройства обработки информации, мы можем сказать, что каждый нейрон имеет множество входов (по которым он связывается с другими нейронами) и только один выход. При этом зависимость значения выходного сигнала от входных сигналов является нелинейной – выход нейрона может быть либо в «возбужденном» (включенном), либо в «заторможенном» (выключенном) состоянии. Исходя из этого, в середине прошлого века была предложена модель простейшего вычислительного элемента, названного «искусственным нейроном». На вход его поступает некоторый вектор чисел. Все элементы входного вектора умножаются на заданные коэффициенты, после чего складываются и подвергаются нелинейному преобразованию, что позволяет имитировать возбуждение и торможение естественных нейронов. Конечно, сегодня мы знаем, что наши биологические нейроны работают совсем не так элементарно. Зато с точки зрения математики, каждый «искусственный нейрон» представляет собой простейший линейный классификатор в пространстве признаков. Набор же таких нейронов, соединенных между собой по входам и выходам (нейронная сеть), может реализовать классификатор практически любой необходимой сложности, а это нам и нужно в реальных задачах машинного обучения.

Почему от отдельных нейронов перешли к нейронным сетям? Во-первых, в природе мы наблюдаем как раз нейронные сети. А во-вторых, задачи классификации образов редко удается решить «в один шаг» – одним линейным классификатором. Другое дело, когда признаки из признакового пространства анализируются с помощью нескольких слоев нейронов. Первый слой получает на вход исходные признаки. Затем второй слой нейронов принимает на вход выходные сигналы нейронов первого слоя, и так далее, слой за слоем. Такую структуру нейронной сети типа торта «Наполеон» назвали многослойным персептроном (*perceptron* – «распознающая машина»). Персептроны продемонстрировали, что искусственные системы могут обучаться на примерах так же, как живые. Это тут же породило огромный всплеск ожиданий прорыва в области искусственного интеллекта, связанного именно с искусственными нейронными сетями. Но в тот момент этого прорыва все-таки не произошло.

Сейчас мы понимаем, что персептронам XX века не хватило сложности и данных. Многослойные персептроны прошлого содержали 3, максимум 5 слоев. Этой глубины

было явно недостаточно для того, чтобы нейронная сеть достигла необходимой обобщающей способности для решения действительно интеллектуальных задач. И даже при такой небольшой глубине обучение персептронов превращалось в сложную вычислительную проблему. Количество весов в каждом нейроне персептрона, которые нужно было обучать («тренировать»), равно размерности входного вектора для всего слоя. И у каждого нейрона эти веса свои. С добавлением каждого нового слоя обучение нейронов происходит все сложнее, поскольку общее число параметров сети, которые нужно настраивать, драматически растет, и все они нелинейно связаны друг с другом. Казалось, на этом пути мы снова попали в ту же ловушку вычислительной сложности, о которой уже говорили выше, применительно к логическому ИИ. Кроме того, если бы кто-то и смог создать тогда действительно глубокий персептрон с тысячами слоев и миллионами параметров, все равно до появления Интернета в мире не существовало необходимых гигантских баз цифровых данных, на которых можно было бы обучить такой огромный персептрон. Ведь чем больше параметров заложено в решающее правило, тем большее количество примеров нужно, чтобы правильно оценить эти параметры.

И, тем не менее, обходной путь для преодоления проблемы вычислительной сложности все-таки нашелся. Идея, позволившая резко уменьшить число обучаемых параметров нейронной сети, пришла из области компьютерного зрения. Там давно уже пользовалась концепция «сверточного фильтра», который обрабатывает изображение не все сразу и целиком, а по фрагментам, как бы в бегущем по изображению «окне». При этом каждое такое «окно» должно обрабатываться одинаково, поскольку при сдвиге изображения результат обработки не должен изменяться. Именно так работает «сверточный нейрон» – базовый элемент сверточных нейронных сетей, которые стали первыми по-настоящему глубокими нейронными сетями. Каждый сверточный нейрон связан не со всеми нейронами предыдущего слоя, а лишь с нейронами в небольшой окрестности собственного положения. И у всех сверточных нейронов одного слоя одинаковые коэффициенты – будто один размноженный нейрон скользит по выходам предыдущего слоя, подобно тому, как оконный сверточный фильтр скользит по изображению. Так устроены все сверточные слои глубокой сети, и это резко – на много порядков – снижает сложность ее обучения.

Сверточные сети для обработки изображений придумали в начале 1990-х годов. Но долгожданный прорыв в качестве их работы произошел лишь тогда, когда с развитием интернета и поисковых сервисов исследователи, занимавшиеся машинным обучением, получили доступ к коллекциям данных, содержащих миллиарды изображений. В 2011 году сверточная сеть AlexNet впервые статистически доказала возможность решения задач распознавания образов на уровне человека. Так появились современные глубокие нейронные сети, и именно с этого момента принято отсчитывать начало современной технологической революции в искусственном интеллекте.

Продолжение следует