



Глава 1

От биологии к информатике, *или We need to go deeper*

TL;DR

Первая глава — вводная. В ней мы:

- узнаем, что в машинном обучении недавно произошла революция и революция эта все еще длится;
 - вспомним историю искусственного интеллекта как науки;
 - познакомимся с разными задачами машинного обучения, узнаем, как они связаны и чем отличаются друг от друга;
 - выясним, почему человеческий мозг может служить образцом для подражания при создании искусственного интеллекта;
 - придем к выводу, что нейробиология пока не слишком точная наука;
 - закончим кратким обзором самых ярких примеров применения современных нейронных сетей и областей, где им еще есть чему поучиться.
-

1.1. Революция обучения глубоких сетей

— А как повернулось дело? — задумчиво пробормотал он. — Трах-та-бабах! Революция!

Ему стало весело. Он глотал пахнувший росой и яблоками воздух и думал: «Столб, хлеб, дом, рожь, больница, базар — слова все знакомые, а то вдруг — Революция! Бейте, барабаны!»

А. Гайдар. Бумбараши

Прежде всего, спросим себя, положив руку на сердце:

— Да есть ли у нас сейчас революция?..

Разве та гниль, глупость, дрянь, копоть и мрак, что происходит сейчас, — разве это революция? Революция — сверкающая прекрасная молния, революция — божественно красивое лицо, озаренное гневом Рока, революция — ослепительно яркая ракета, взлетающая радугой среди сырого мрака!.. Похоже на эти сверкающие образы то, что сейчас происходит?..

А. Аверченко. Дюжина ножей в спину революции

Десять лет назад, в середине 2000-х годов, в машинном обучении началась революция. В 2005–2006 годах группы исследователей под руководством Джеффри Хинтона (Geoffrey Hinton) в университете Торонто и Йошуа Бенджи (Yoshua Bengio) в университете Монреаля научились обучать *глубокие нейронные сети*. И это перевернуло весь мир машинного обучения! Теперь в самых разнообразных предметных областях лучшие результаты получаются с помощью глубоких нейронных сетей. Одним из первых громких промышленных успехов стало распознавание речи: разработанные группой Хинтона глубокие сети очень быстро радикально улучшили результаты распознавания по сравнению с оттачивавшимися десятилетиями классическими подходами, и сегодня любой распознаватель, включая голосовые помощники вроде Apple Siri и Google Now, работает исключительно на глубоких нейронных сетях. А сейчас, к 2017 году, люди научились обучать самые разные архитектуры глубоких нейронных сетей, и те решают абсолютно разные задачи: от распознавания лиц до вождения автомобилей и игры в го.

Но идеи большинства таких моделей появились еще в 80–90-х годах XX века, а то и раньше. Искусственные нейронные сети стали предметом исследований очень давно; скоро мы увидим, что они были одной из первых хорошо оформленных идей искусственного интеллекта, когда и слов-то таких — «искусственный интеллект» — еще никто не слышал. Но с начала 90-х годов XX века до середины нулевых этого¹ нейронные сети были, мягко говоря, не в моде. Известный

¹ Мы искренне надеемся, что написали хорошую книгу, но все-таки не будем здесь делать специальных оговорок в расчете на читателей XXII века...

исследователь нейронных сетей Джон Денкер (John Denker) в 1994 году сказал: «Нейронные сети — это второй лучший способ сделать практически что угодно». И действительно, на тот момент уже было известно, что нейронная сеть теоретически может приблизить любую функцию и обучиться решать любую задачу, а глубокая нейронная сеть способна еще и более эффективно решить гораздо больше разных задач... Но обучать глубокие сети никому не удавалось, другие методы на конкретных практических примерах работали лучше. Из общих методов машинного обучения это были сначала «ядерные методы» (kernel methods), в частности метод опорных векторов, а затем байесовские сети и в целом графические вероятностные модели. Но вообще Денкер вел речь о том, что более простая вероятностная модель, разработанная специально с учетом специфики конкретной задачи, обычно работала лучше, чем нейронная сеть «общего назначения», даже если выбрать ей подходящую для задачи архитектуру. А глубокие нейронные сети обучаться никак не хотели; о том, почему именно, мы подробно поговорим в разделе 3.5.

Решение, предложенное группой Хинтона в середине 2000-х годов, пришло в виде *предобучения без учителя*, когда сеть сначала обучается на большом наборе данных без разметки, а потом уже дообучается на размеченных данных, используя это приближение. Например, если мы хотим распознавать человеческие лица, то давайте сначала пообучаем нейронную сеть на фотографиях с людьми вообще, без разметки (таких фотографий можно легко набрать сколь угодно много), а уже потом, когда сеть «насмотрится» на неразмеченные фотографии, дообучим ее на имеющемся размеченном наборе данных. Оказалось, что при этом конечный результат становится намного лучше, а хорошие и интересные признаки сеть начинает выделять еще на этапе предобучения без учителя. Именно таким образом произошел, в частности, прорыв в распознавании речи. Конечно, здесь пока совершенно непонятно, что сеть собственно должна делать с этими неразмеченными фотографиями, и в этом и заключался прорыв середины нулевых годов. Мы немного поговорим об этих методах в разделе 4.2, однако без подробностей, потому что сегодня эти методы уже практически не используются.

Почему? Потому что все неплохо работает и без них! Оказалось, что сейчас мы можем успешно обучать нейронные сети, в том числе глубокие, фактически теми же методами, которыми раньше это сделать никак не удавалось. Методы предобучения без учителя оказались лишь «спусковым крючком» для революции глубокого обучения. Второй важнейшей причиной стал, собственно, прогресс в вычислительной технике и в размерах доступных для обучения наборов данных. С развитием Интернета данных становилось все больше: например, классический набор данных MNIST, о котором мы начнем подробный разговор в разделе 3.6 и на котором добрый десяток лет тестировались модели компьютерного зрения, — это 70 тысяч изображений рукописных цифр размером 28×28 пикселей, суммарно около 10 Мбайт данных; а современный стандартный датасет для моделей компьютерного зрения ImageNet содержит уже около 1,2 Тбайт изображений. В части

вычислений помогают и закон Мура¹, и важные новшества, которые позволяют обучать нейронные сети быстрее. Вся глава 4 будет посвящена таким новшествам. В основном мы, конечно, будем говорить о продвижениях в самих алгоритмах обучения, но будут и фактически чисто «железные» продвижения: обучение современных нейронных сетей обычно происходит на графических процессорах (GPU, то есть по сути на видеокартах), что часто позволяет ускорить процесс в десятки раз... Эта идея появилась около 2006 года и стала важной частью революции глубокого обучения.

Но, конечно, содержание революции глубоких сетей не ограничивается тем, что компьютеры стали быстрее, а данных стало больше. Если бы это было так, можно было бы никаких книг не писать, а закончить разговор прямо сейчас. Техническое развитие позволило не только вернуться к идеям 80–90-х годов XX века, но и придумать много новых идей.

Развивались и общие методы обучения нейронных сетей (о них мы поговорим в главе 4 и вообще в первой части книги), и классические архитектуры нейронных сетей, которым в основном посвящена вторая часть, — сверточные сети и автокодировщики (глава 5), рекуррентные сети (глава 6). Но появлялись и совершенно новые архитектуры: порождающие состязательные сети (глава 8) сумели превратить нейронные сети в порождающие модели, нейронные сети в обучении с подкреплением (глава 9) привели к невиданным ранее прорывам, нейробайесовские методы (глава 10) соединили нейронные сети и классический вероятностный вывод с помощью вариационных приближений, а нужды конкретных приложений привели в разработке таких новых архитектур, как сети с вниманием (раздел 8.1) и сети с памятью, которые уже находят и другие применения.

Прежде чем двигаться дальше, скажем пару слов об основных источниках. Книг по обучению глубоких сетей пока не так уж много. Некоторые считают, что эта область еще не вполне устоялась и систематизировать ее пока не обязательно. Как вы наверняка догадались, мы с этим не вполне согласны. Фактически сейчас есть только две известные книги об обучении глубоких сетей: в 2014 году вышла книга сотрудников Microsoft Research Ли Денга (Li Deng) и Донга Ю (Yu Dong) *Deep Learning: Methods and Applications* [117], в которой основной упор делается на приложения для обработки сигналов, то есть для анализа изображений и речи. А совсем недавно, в конце 2016 года, вышла книга Иэна Гудфеллоу (Ian Goodfellow), Йошуа Бенджи и Аарона Курвилля (Aaron Courville) *Deep Learning* [184], в которой дается подробное и идейное введение в тему. Хотя эта книга, формально говоря, наш прямой конкурент, мы ее всячески рекомендуем; кстати, русский перевод

¹ Гордон Мур (Gordon Moore) сформулировал свой закон, по которому число транзисторов в интегральной схеме удваивается примерно каждые два года, а общая производительность систем — примерно каждые 18 месяцев, еще в 1965 году (точнее, тогда он предсказывал удвоение каждый год, а в 1975-м стал более пессимистичным). Конечно, это не закон природы, а просто эмпирическая закономерность, но поразительно, насколько долго эта экспоненциальная зависимость сохраняет силу: закон Мура в наше время (2016–2017 годы) лишь немного замедлился, с двух лет до двух с половиной. Правда, сейчас речь идет уже не только о числе транзисторов в одной схеме, но и о числе ядер в процессоре.

тоже уже существует. Кроме того, есть целый ряд важных больших обзоров, которые до статуса книг не дотягивают, но дают очень широкое представление о глубоком обучении: ранний обзор Бенджи [38] показывает состояние дел на 2009 год, его же более поздний обзор рассматривает глубокое обучение как обучение *представлений* (representation learning) [39], обзор Ли Денга сделан с уклоном в обработку сигналов [115], а обзор Юргена Шмидхубера (Jurgen Schmidhuber) [475] дает очень подробный анализ истории основных идей глубокого обучения.

Есть, конечно, и более специальные обзоры для конкретных областей: обработки изображений [107], обработки естественных языков [182] и многих других. Из статей, в которых речь идет об обучении глубоких сетей «вообще», можно выделить программную статью в Nature трех главных представителей глубокого обучения: Джеффри Хинтона, Йошуа Бенджи и Яна ЛеКуна (Yann LeCun) [314]. На более специализированные источники мы будем ссылаться на протяжении всей этой книги: область молодая, развивается она очень быстро, и в списке литературы, в который мы старались помещать только действительно важные статьи, источников уже накопилось немало. Кстати, следить за последними новостями глубокого обучения лучше всего на *arXiv*¹: новые работы по обучению глубоких сетей появляются там постоянно, и если вы не заходили на arXiv два-три месяца, может оказаться, что в вашей области уже произошла какая-нибудь очередная мини-революция, а вы и не заметили...

В этой книге мы поговорим о разных архитектурах нейронных сетей, разберем, как именно обучать нейронные сети, какие задачи можно ими решать. Более того, мы дойдем и до практики: почти каждый наш рассказ будет сопровождаться вполне рабочими примерами кода с использованием современных библиотек для обучения нейронных сетей, главным образом TensorFlow и Keras. Подобные библиотеки сделали обучение нейронных сетей общедоступным и стали важной компонентой взрыва интереса и быстрого развития нашей области. Нас ждет очень много интересного. А в этой, вводной, главе мы начнем с того, что поговорим о том, что вообще такое искусственный интеллект и машинное обучение, как они появились, как развивались и как искусственные нейронные сети, которые мы рассматриваем в этой книге, связаны с естественными, с помощью которых вы эту книгу читаете.

¹ Хранилище научных препринтов arXiv, расположенное по адресу <http://arxiv.org>, содержит статьи по многим научным дисциплинам. Но в некоторых областях arXiv стал стандартом де-факто, куда выкладываются практически все серьезные статьи, обычно еще до официальной публикации. К таким областям относится и машинное обучение, в частности обучение глубоких сетей. Правда, стоит отметить, что на arXiv нет никакого рецензирования, так что совсем свежие статьи там могут оказаться и с ошибками, иногда важными. Например, в 2010 году немало шума наделало опубликованное на arXiv решение проблемы $P \neq NP$, предложенное Винаем Деодаликаром. К сожалению, тревога оказалась ложной, и в решении нашлись ошибки. А в 2017-м появилось решение Норберта Блюма, и, когда мы пишем эти слова, решение еще не опровергнуто — впрочем, шансов мало: Александр Разборов уже указал соображения, по которым это доказательство не должно работать... А недавняя инициатива <https://openreview.net/> делает вещи, для «классической науки» совершенно немыслимые: там публикуются статьи, еще только поданные на конференции, и рецензии на них.