

ОБУЧЕНИЕ НЕЙРОННЫХ СЕТЕЙ И ЕГО ЗНАЧЕНИЕ ДЛЯ РАЗВИТИЯ ПРОГРАММНОЙ ИНЖЕНЕРИИ

Козак Евгений

Старший разработчик, Memery Crystal LLP, Лондон,
Англия
eugeniu.cozac@gmail.com

NEURAL NETWORK TRAINING AND ITS SIGNIFICANCE FOR THE DEVELOPMENT OF SOFTWARE ENGINEERING

E. Cozac

Summary. Currently, there is an increased interest in neural networks in various spheres of public life. However, the success of neural networks when used is not unconditional, since the main problem associated with the use of neural networks is their training. At the same time, the process of training neural networks is quite time-consuming, since it is necessary to choose both the task itself that the trained neural network should solve, and to prepare data sets for solving such a task—and the more complex the task that is assigned to the neural network, the larger and more disparate the data sets that need to be prepared.

At the same time, when training neural networks, it is quite standard to have a situation in which the neural network copes with the task of training quite easily and quickly, and the learning process itself can be very slow. This is due both to the nature of neural networks, and to the fact that the solution of the problem of training neural networks and the choice of an algorithm from training depends on the direction of using neural networks.

In the context of this article, the author considers certain aspects related to the training of neural networks and concludes that the training of neural networks is of great practical importance for the purposes of software engineering, since it allows solving more and more new tasks every year, and the importance of training neural networks is due to the ability of trained neural networks to accumulate large volumes of disparate data, the possibility of adaptation and improvement of neural networks with an increase in the performance of electronic computing equipment, as well as the ability of neural networks to adapt to constantly improving learning algorithms.

Keywords: machine learning, neural network training, deep learning, artificial intelligence, neural networks, neural network training algorithms.

Аннотация. В настоящее время наблюдается повышенный интерес к нейронным сетям в различных сферах общественной жизни. Однако успех нейронных сетей при их использовании не безусловен, поскольку главной проблемой является их обучение. При этом сам процесс обучения нейронных сетей довольно трудоемкий, поскольку необходимо выбрать как саму задачу, которую должна решать обученная нейронная сеть, так и подготовить наборы данных для решения такой задачи. При этом чем сложнее задача, которая поставлена перед нейронной сетью, тем больше и разрозненнее наборы данных, которые необходимо подготовить.

При этом при обучении нейронных сетей вполне стандартной является ситуация, при которой нейронная сеть достаточно легко и быстро справляется с поставленной задачей обучения, а сам процесс обучения может происходить очень медленно. Обусловлено это как непосредственно самой природой нейронных сетей, так и тем, что решение задачи обучения нейронных сетей и выбор алгоритма их обучения зависит от направления использования нейронных сетей.

В контексте настоящей статьи автором рассматриваются отдельные аспекты, связанные с обучением нейронных сетей, и делается вывод о том, что обучение нейронных сетей имеет большое практическое значение для целей программной инженерии, поскольку позволяет с каждым годом решать все новые и новые задачи, а значение обучения нейронных сетей обусловлено возможностью уже обученных нейронных сетей аккумулировать большие объемы разрозненных данных, возможностью адаптации и совершенствования нейронных сетей, связанной с ростом производительности электронно-вычислительной техники, а также возможностью нейронных сетей приспосабливаться к постоянно совершенствующимся алгоритмам обучения.

Ключевые слова: машинное обучение, обучение нейронных сетей, глубокое обучение, искусственный интеллект, нейронные сети, алгоритмы обучения нейронных сетей.

Нейронные сети в настоящий момент являются основой большинства технологий. Особенности этих сетей обусловлены тем, что они состоят из набора алгоритмов, который моделируется на человеческом мозге. Эти алгоритмы могут интерпретировать сенсорные данные с помощью машинного восприятия и маркировать или группировать необработанные

данные. Они предназначены для распознавания числовых паттернов, содержащихся в векторах, в пределах которых должны быть переведены все реальные данные (изображения, звук, текст, временные ряды и т.д.).

Отсюда можно сделать вывод о том, что основная задача нейронных сетей состоит в группировании и клас-

сификации необработанных данных — они группируют немеченные данные на основе сходств, найденных во входных данных, а затем классифицируют данные на основе помеченного обучающего набора данных. Нейронные сети могут автоматически адаптироваться к изменяющимся входным данным. Таким образом, человеку не нужно перестраивать выходные критерии каждый раз, когда входные данные меняются, чтобы получить наилучший возможный результат.

Существующие нейросетевые технологии начали развиваться в середине 2000-х годов, когда накопилось не только достаточное количество знаний в такой предметной области, но и далеко вперед шагнули наука и техника. Так, например, если до 2010 года не существовало базы данных, позволяющей обучить нейронные сети решать задачи по распознаванию изображений и их классификации и по указанной причине нейронные сети совершали ошибки, путая изображения, то с появлением первой базы данных, в которой содержалось более 15 миллионов изображений, классифицированных по 22 различным категориям, обучение нейронных сетей стало более эффективным.

Примерно к этому же времени появились новые достижения ученых в области искусственного интеллекта: предобучение сети (при помощи машины Больцмана обучался каждый слой многослойной сети отдельно), сверточные нейронные сети для распознавания изображений, каскадный автокодировщик для задействования всех слоев в глубокой нейронной сети и пр.

Основными современными трендами в области обучения нейронных сетей являются: глубокое обучение, гибридное обучение и автоматическое обучение [2, 8, 10]. При этом первое — глубокое обучение, развивается быстрее всех остальных. Обусловлено это спецификой глубокого обучения нейронных сетей. Как отмечается в специализированной литературе, в основе глубокого обучения нейронных сетей лежит сочетание сразу нескольких алгоритмов обучения [9]. Среди таких алгоритмов — алгоритмы обучения нейронных сетей с учителем, алгоритм обучения сетей без учителя, а также алгоритм обучения сетей с подкреплением полученных результатов.

Глубокое обучение определяется в специализированной литературе как одно из направлений машинного обучения, которое использует серию нелинейных процессоров, состоящих из нескольких слоев, для преобразования и извлечения объектов. В рамках глубокого обучения несколько слоев искусственных нейронных сетей выполняют процесс машинного обучения. Первый слой нейронной сети обрабатывает исходные данные и передает информацию на второй слой. Затем

второй уровень обрабатывает эту информацию дальше, добавляя дополнительную информацию (например, IP-адрес пользователя) и передает ее следующему уровню. Этот процесс продолжается на всех уровнях сети глубокого обучения до тех пор, пока не будет достигнут желаемый результат.

С технологической точки зрения глубокое обучение позволяет нейронной сети сначала тренироваться на большом количестве обобщенных данных, а уже потом адаптироваться под решение конкретных задач [1, 2, 4, 6]. Традиционное машинное обучение можно охарактеризовать как обучение делать прогнозы на основе прошлых наблюдений, подходы же глубокого обучения работают, позволяя нейронным сетям не только предсказывать, но и правильно представлять данные для того, чтобы они были пригодны для прогнозирования. Учитывая большой набор желаемых отображений, подходы глубокого обучения работают путем подачи данных в сеть, которая производит последовательные преобразования входных данных до тех пор, пока окончательное преобразование не предскажет выходные данные.

Даже гибридное обучение в сравнении с глубоким обучением видится уже не настолько совершенным, поскольку в большей степени ориентировано на вероятностное моделирование и вычисления причинно-следственных отношений в биоинформатике (генетические сети, структура белков), медицине, классификации документов, обработке изображений и системах поддержки принятия решений [4].

Нейронные сети, обучаемые с использованием технологий глубокого обучения, выполняют автоматическое извлечение признаков необходимого объекта без вмешательства человека, чего сделать, например, не позволяет большинство традиционных алгоритмов машинного обучения.

Кроме того, нейронная сеть глубокого обучения, обученная на структурированных данных, может быть затем применена к неструктурированным данным, предоставляя ей доступ к гораздо большему количеству входных данных, чем сети машинного обучения. Таким образом, чем совершеннее нейронная сеть и чем больше данных сеть может обучать, тем более точным будет результат.

Так, например, сверточные нейронные сети при глубоком обучении позволяют решать задачи классификации, в которых ожидаемо нахождение локальных подсказок относительно принадлежности объекта к тому или иному классу, но эти подсказки могут появляться в разных местах во входных данных. Например,

в задаче классификации документов при гибридном обучении одна ключевая фраза может помочь в определении нейронной сетью темы документа при нахождении ключевой фразы в определенном месте документа. При глубоком обучении сверточной нейронной сети не имеет значения, в каком месте текста документа будет встречаться ключевая фраза. Сверточные и пуловые слои позволяют модели научиться находить любые заданные локальные индикаторы вне зависимости от их положения.

Принципы работы глубокого обучения наиболее показательны при их рассмотрении на примере логистической регрессии (механизм, который используется для классификации с целью преобразования непрерывных сигналов в двоичные) [7]. Механизм логистической регрессии вычисляет вероятность того, что набор входных данных соответствует заданной метке. Сказанное можно наглядно продемонстрировать с помощью следующего выражения:

$$F(x) = \frac{1}{1 + n^{-x}}$$

Для того чтобы непрерывные входные данные были выражены в виде вероятностей, они должны иметь положительные результаты, поскольку отрицательной вероятности не существует. Теперь рассмотрим отношение показателя n к дроби $1/1$. Единица — это потолок вероятности, за пределы которого результаты не могут выйти.

По мере того, как входной сигнал x , который запускает метку, растет, выражение n для x сжимается до нуля, оставляя долю $1/1$, или равную 100%, что означает, что осуществляется приближение к абсолютной уверенности в том, что метка применима. Вход, который отрицательно коррелирует с выходом, будет иметь свое значение, перевернутое отрицательным знаком на экспоненте n , и по мере того, как этот отрицательный сигнал растет, величина n к значению x становится больше, подталкивая всю дробь все ближе к нулю.

То есть сигналы, которые нейронная сеть получает в качестве входных данных, будут охватывать диапазон

значений и включать любое количество метрик в зависимости от проблемы, которую нейронная сеть должна будет решить.

Подводя итог, необходимо отметить, что обучение нейронных сетей в настоящее время является достаточно перспективной и развивающейся областью исследования, что обусловлено рядом причин. Во-первых, потому, что нейронные сети позволяют обрабатывать большие объемы данных. Так, в специализированной литературе отмечается, что огромное количество доступных данных, собранных за последнее десятилетие, в значительной степени способствовало популярности глубокого обучения, поскольку позволило нейронным сетям действительно показать свой потенциал, так как нейронные сети становятся совершеннее, чем больше данных в них вводится [3, 5]. Для сравнения, традиционные алгоритмы машинного обучения, например, не становятся производительнее с ростом объема обрабатываемых данных.

Во-вторых, рост глубокого обучения нейронных сетей обусловлен ростом производительности электронной вычислительной техники. По словам Рэя Курцвейла, ведущего специалиста в области искусственного интеллекта, вычислительная мощность умножается на постоянный коэффициент для каждой единицы времени (например, удваивается каждый год), а не просто прибавляется постепенно. Это означает, что вычислительная мощность растет экспоненциально [7]. На момент зарождения концепции глубокого обучения исследователи не имели доступа к достаточному количеству данных или вычислительной мощности для создания и обучения значимых моделей глубокого обучения.

В-третьих, еще одним фактором, который увеличил популярность глубокого обучения нейронных сетей, стал прогресс, достигнутый в алгоритмах обучения. Новые алгоритмы обучения позволяют обучать сети намного быстрее, чем даже пять лет назад, что позволяет использовать все больше и больше данных. Основное преимущество нейронных сетей заключается в их способности превосходить почти любой другой алгоритм машинного обучения.

ЛИТЕРАТУРА

1. Барский А.Б. Введение в нейронные сети / Барский А.Б. — Электрон. текстовые данные. — М.: Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. — 358 с.
2. Бова В.В., Дуккардт А.Н. Применение искусственных нейронных сетей для коллективного решения интеллектуальных задач. Известия ЮФУ. Технические науки. Хайкин С. Нейронные сети: Пер. с англ. — М.: Издательский дом «Вильямс», 2006.
3. Брагин А.Д., Спицын В.Г. Распознавание моторных образов на электроэнцефалограммах с применением сверточных нейронных сетей // КО. — 2020. — № 3. — С. 482–489.
4. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. — 2-е изд., стереотип. — М.: Горячая линия-Телеком, 2002. — 382 с.

5. Редько В.Г. Эволюция, нейронные сети, интеллект: Модели и концепции эволюционной кибернетики / В.Г. Редько. — М.: Ленанд, 2019. — 224 с.
6. Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы / Рутковская Д., Пилиньский М., Рутковский Л. — Электрон. текстовые данные. — М.: Горячая линия — Телеком, 2013. — 384 с.
7. Созыкин Андрей Владимирович Обзор методов обучения глубоких нейронных сетей // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. — 2017. — № 3. — С. 28–59.
8. Сысоев Д.В. Введение в теорию искусственного интеллекта: учебное пособие / Сысоев Д.В., Курипта О.В., Проскурин Д.К. — Электрон. текстовые данные. — Воронеж: Воронежский государственный архитектурно-строительный университет, ЭБС АСВ, 2014. — 171 с.
9. Тихонов А.А. Большие данные и глубокое машинное обучение в искусственных нейронных сетях // Наука и образование сегодня. — 2018. — № 6 (29). — С. 35–38.
10. LeCun Y., Bengio Y., Hinton G. Deep Learning // Nature. — 2015. — Vol. 521. — P. 436–444.

© Козак Евгений (eugeniu.cozac@gmail.com).

Журнал «Современная наука: актуальные проблемы теории и практики»



г. Лондон