

ОСОБЕННОСТИ ГЛУБОКОГО ОБУЧЕНИЯ НЕЙРОННЫХ СЕТЕЙ

© 2022 А. В. Фирсов, А. В. Линкина

Воронежский институт высоких технологий (Воронеж, Россия)

В статье рассматриваются особенности глубокого обучения нейронных сетей. Рассматриваются такие методологии как система подкрепления, обучение с учителем, метод локального минимума или максимума функции при обучении с учителем. Описывается модификация градиентного спуска - стохастический градиентный спуск, приводятся примеры решения задач на основе данного инструмента.

Ключевые слова: искусственный интеллект, сверточные сети, вес ошибки, глубокое обучение, машинное обучение, обучение с учителем, перцептрон, градиентный спуск.

В ходе стремительного развития нейросетевых технологий, машинного зрения тема глубокого обучения (Deep learning) находится на острие дискуссионных вопросов в связи с тем обстоятельством, что большинство исследователей и практиков пытаются найти самый простой, быстрый, эффективный и менее затратный по ресурсам способ обучения нейросетей.

Рассмотрим далее примеры существующих способов и алгоритмов глубокого обучения нейронных сетей, их преимущества и недостатки.

История исследований на эту тему начинается свой отсчет с кибернетической модели мозга Фрэнка Розенблатта в 1958 году в виде созданной им электронной машины Марк-1. Это была простая модель однослойного перцептрона, которую он изложил в своей книге «Принципы нейродинамики: Перцептроны и теория механизмов мозга» в 1962 году. Исследователь смог создать алгоритмы для обучения однослойного перцептрона: система подкрепления (т.е. некий набор правил, с помощью которых можно изменять состояние памяти перцептрона), обучение с учителем (метод коррекции ошибок, суть которого заключатся в том, что вес связи не меняется если перцептрон, по мнению учителя, выдает верный результат), обучение без учителя (система подкрепления, в которой веса всех активных связей изменяются на одинаковую величину, а веса неактивных связей не изме-

няются). Также был предложен метод обратного распространения ошибки (придуманый и описанный в 1974 году А. И. Галушкиным, а также независимо и одновременно с ним Полом Дж. Вербосом для многослойных перцептронов).

Первый рабочий алгоритм для глубоких многослойных перцептронов прямого распространения был опубликован в книге советских учёных Алексея Григорьевича Иващенко и Валентина Григорьевича Лапы «Кибернетические предсказывающие устройства» ещё в 1965 г. Хотя сам термин глубокого обучения появился в 1986 году в работе Рины Дехтер.

Революция в глубоком обучении произошла в 2011 году. Несмотря на то, что сверточные нейронные сети, использовавшие метод обратного распространения ошибки, существовали достаточно давно, внедрение их в графический процессор (GPU) позволило добиться впечатляющей производительности в визуальном распознавании образов.

Таким образом, на современном этапе под машинным обучением можно считать преобразование одних данных в другие, но в отличие от классического программирования, это преобразование зависит от параметров, определяющих, как эта модель работает, а подбор этих параметров как раз называется обучением. Чаще всего используются обучение с так называемым учителем, которое подразумевает собой пару объектов, в которые

Фирсов Александр Викторович – Воронежский институт высоких технологий, магистрант, e-mail: anna-linkina@rambler.ru.

Линкина Анна Вячеславовна – Воронежский институт высоких технологий, старший преподаватель, e-mail: anna-linkina@rambler.ru.

входят изначальные данные и правильный ответ к ним (рис. 1).

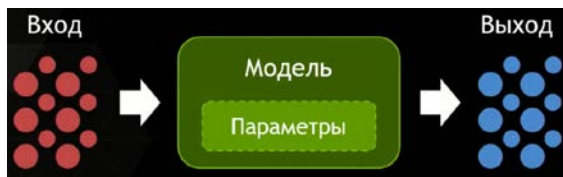


Рисунок 1. Модель представления обучения с учителем

Само по себе машинное обучение – это решение задачи оптимизации, т. е. минимизация ошибки. Это возможно сделать с помощью методов градиентного спуска, а точнее с помощью некоторых его модификаций. Суть этого метода заключается в нахождении локального минимума или максимума функции.

В нейронных сетях обычно стоит задача поиска минимума для многомерных функций, в примере на рисунке 2 для простоты восприятия представлена одномерная функция, функция ошибки E , в которой нам необходимо найти точку минимума W^* . Поиск минимума начинается с некоего случайного места, вычисляется вектор градиента. Полученный вектор изначально направлен в сторону возрастания функции, следовательно, если взять его со знаком минус, то он будет направлен в сторону убывания функции, что в свою очередь позволяет определить минимум этой функций. Однако ошибку локального минимума можно избежать с помощью некоторых модификаций этого метода, которые и используются в обучении нейросетей. В данном зацикленном процессе необходимо прибавить к текущему значению весов с некоторым параметром α (который еще называется скоростью обучения) полученный антиградиент.



Рисунок 2. Иллюстрация метода локального минимума или максимума функции при обучении с учителем.

Модификацией градиентного спуска, который непосредственно применяется в нейронных сетях, является стохастический градиентный спуск. Его суть заключается в том, что примеры из обучающей выборки для нейронной сети подаются по очереди или некоторыми пакетами. Для простоты рассмотрим подачу по очереди. На некоторой итерации имеется объект и правильный ответ к нему. Пусть заданный объект является входящим в нейронную сеть с уже заданными весами, которые при первой итерации задаются случайными числами. При получении выходного сигнала нейросети можно вычислить ошибку, так как нам известен правильный ответ. Далее с помощью полученной разности можно найти градиент, с помощью которого будут отредактированы веса для минимизации ошибки. В отличие от глобального минимума градиентного спуска ошибка вычисляется не на всех образцах, а локально, т. е. на подвыборке или даже на одном объекте.

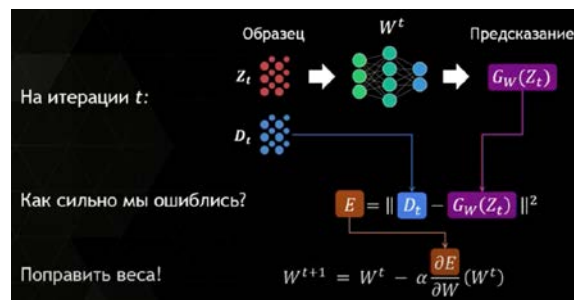


Рисунок 3. Пример обучения с использованием стохастического градиентного спуска

Несмотря на то, что сам по себе градиент достаточно сложен, для его расчета необходимо знать лишь связи между нейронами в нейронной сети. Если известно, как одни нейроны зависят от других нейронов, то можно применить правила дифференцирования сложной функции и всё сведётся к вычислению производной одной функции ошибки. При глубоком обучении этот метод называется методом обратного распространения ошибки. Его суть заключается в том, чтобы найти производную от конечного слоя к изначальному изменяя веса каждому нейрону.

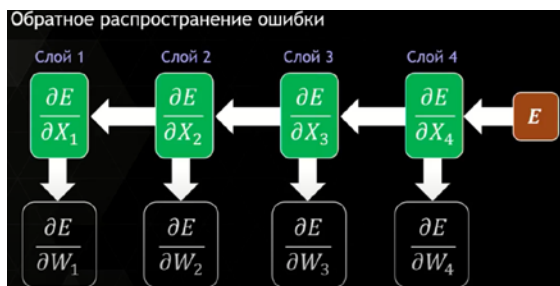


Рисунок 4. Метод обратного распространения ошибки

Разобранный выше многослойный перцептрон справляется задачами классификации и регрессии, но в некоторых частных случаях он может работать хуже, тогда для решения подобных задач целесообразно пользоваться иными методами.

Однако для задач машинного зрения по классификации изображений на основе отдельного признака нейронные сети имеют неоспоримое преимущество. Вместе с тем имеется риск появления такой проблемы как переобучение – нейронная сеть найдёт прямые связи между изображениями, запомнит их и будет выдавать некорректные результаты на других выборах.

Машинное зрение – это сопоставление картинок с неким семантическим описанием. Также выделяют детектирование объектов, т. е. определение разных объектов на изображении и их положение на нем. Еще более сложной задачей является семантическая сегментация – разметка каждого пикселя и определение его содержания.

Трудности машинного зрения заключаются в очень больших объемах неструктурированной информации. Дело в том, что изображение представляет собой матрицу из некоторых чисел.

Еще до создания машинного зрения и глубокого обучения нейросетей люди умели извлекать признаки, но эти признаки извлекались вручную. Признак – это некая локальная особенность для конкретной картинке. После выделения признаков и преобразования их в некоторые числа можно отправить на обработку в нейронную сеть.

Свёртка – это некий алгоритм, позволяющий автоматизировать процесс извлечения локальных признаков. На рисунке 5 представлен алгоритм свертки. Зеленым окошком

представлено ядро свертки (небольшая матрица весов), к нему применяется искомым шаблон W и с помощью формулы рассчитывается корреляция. В результате получается некоторое число, которое записывается в память.



Рисунок 5а. Алгоритм свертки

В результате получается карта признаков, которая описывает изображение по заданному шаблону.



Рисунок 5б. Алгоритм свертки

Возникает вопрос – как определить признак (шаблон)? Ранее данные признаки определялись вручную (на приведенном выше примере, в частности, очевидно, что процесс был трудоемок и занимал длительное время), в связи с чем был создан лучший подход. Свёртка вычисляется с помощью линейной комбинации взвешенной суммы входных пикселей.

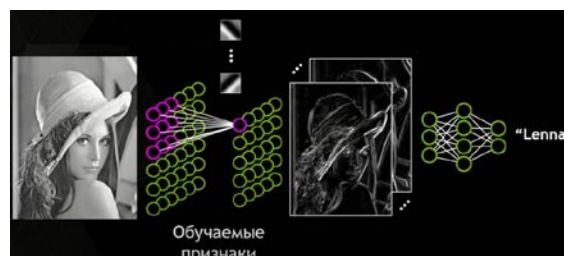


Рисунок 6. Модернизированный алгоритм свертки

Определим свертку как специальный слой нейронной сети. Тогда каждый входной нейрон не будет связан с каждым выходным нейроном. В данном случае связь представлена как некая группа нейронов 1 слоя (выделенная розовым), влияющая только на один нейрон следующего слоя. Веса будут определяться из ядра свертки, за счет того, что это линейная комбинация. Тогда получим однородную карту признаков. Признаки получены автоматическим способом во время обучения нейронной сети и отражают силу связи между нейронами. Далее имеем карту признаков, а точнее некоторый набор этих карт для разных ядер свертки. Имеется возможность использовать полученные результаты в полносвязной нейронной сети или извлекать иерархию признаков (признаки в пространстве признаков предыдущего уровня).

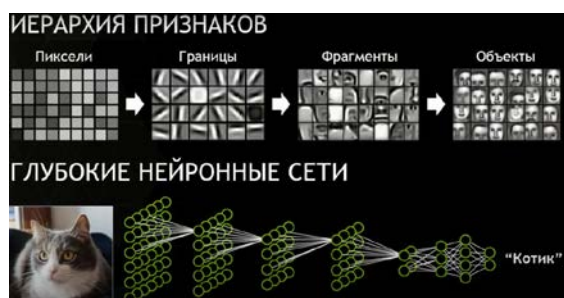


Рисунок 7. Иерархия признаков

На рисунке 7 представлен процесс извлечения иерархии признаков. Из 1-го уровня, где есть пикселем, выделяются некоторые границы, из которых отбираются более высокоуровневые семантические признаки, далее из этих фрагментов на следующем уровне выделяются еще более сложные признаки и т. д. На выходе имеем настолько высокоуровневую карту признаков, что каждый нейрон соответствует высокоуровневому фрагменту изображения и идентификацию изображения можно завершить полносвязной нейронной сетью (поскольку можно провести т. н. разделяющую поверхность между результирующими объектами).

Таким образом, глубокое обучение нейронных сетей (Deep learning) – это группа методов машинного обучения, основанная на обучении иерархии признаков представлений, а не специализированных на алгорит-

мах, созданных под конкретные задачи. Особенностью использования нейронных сетей с глубоким обучением является использование значительных массивов исходных данных, в том числе и зашумленных. Кроме того, при решении сложных задач распознавания текстов, изображений и т.п. чем больше таких массивов, тем задача решается успешнее.

СПИСОК ИСТОЧНИКОВ

1. Голенков В. В. Инструментальные средства проектирования интеллектуальных обучающих систем / В. В. Голенков, Н. А. Гулякина, О. Е. Елисеева. – Минск: Белорусский государственный университет информатики и радиоэлектроники, 1999. – 101 с. – URL: <https://www.bsuir.by/m/48815.pdf> (дата обращения: 06.10.2022).

2. Горбаченко В. И. Интеллектуальные системы: нечеткие системы и сети: учебное пособие для вузов / В. И. Горбаченко, Б. С. Ахметов, О. Ю. Кузнецова. – 2-е изд., испр. и доп. – Москва: Издательство Юрайт, 2023. – 105 с. – (Высшее образование). – ISBN 978-5-534-08359-0. – Текст: электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/514580> (дата обращения: 19.10.2022).

3. Инструменты для разработки искусственного интеллекта и машинного обучения [Электронный ресурс]. Режим доступа: <http://digitrode.ru/articles/1644-instrumenty-dlya-razrabotki-iskusstvennogo-intellekta-i-mashinnogo-obucheniya.html> (дата обращения: 17.10.2022).

4. Линкина А. В. Информационное обеспечение цифровых технологий в агропромышленном комплексе / А. В. Линкина, И. Ю. Богданчиков // Вестник Воронежского института высоких технологий. – 2021. – № 2 (37). – С. 25-27.

5. Львович А. И. Алгоритмизация процесса визуально-экспертного моделирования при оптимизации управления развитием организационных систем с использованием мониторинговой информации / А. И. Львович, А. П. Преображенский. // Моделирование, оптимизация и информационные технологии. – 2022. – Т. 10. – №1 (36). – [Электронный ресурс]: <https://moitvvt.ru/ru/journal/article?id=1154> (дата обращения 10.09.2022).

6. Мельникова Т. В. Моделирование обработки больших массивов данных в распределенных информационно-телекоммуникационных системах / Т. В. Мельникова, М. В. Питолин, Преображенский Ю. П. // Моделирование, оптимизация и информационные технологии. – 2022. – Т. 10. – № 1 (36). – [Электронный ресурс]: <https://moitvvt.ru/journal/article?id=1117> (дата обращения: 10.09.2022)

7. Платонов А. В. Машинное обучение: учебное пособие для вузов / А. В. Платонов. – Москва: Издательство Юрайт, 2023. – 85 с. – (Высшее образование). – ISBN 978-5-534-15561-7. – Текст: электронный // Образовательная платформа Юрайт [сайт]. –

URL: <https://urait.ru/bcode/520544> (дата обращения: 15.10.2022)

8. Системы искусственного интеллекта [Электронный ресурс]. Режим доступа: [http://www.machinelearning.ru/wiki/index.php?title=Системы_искусственного_интеллекта_\(курс_лекций,_Д.В.Михайлов\)](http://www.machinelearning.ru/wiki/index.php?title=Системы_искусственного_интеллекта_(курс_лекций,_Д.В.Михайлов)) (дата обращения: 15.10.2022).

9. Чупринская Ю. Л. Краткий обзор современных технологических трендов в контексте цифровой трансформации / Ю. Л. Чупринская, А. В. Линкина // Вестник Воронежского института высоких технологий. – 2022. – № 1 (40). – С. 107-109.

FEATURES OF DEEP LEARNING OF NEURAL NETWORKS

© 2022 A. V. Firsov, A. V. Linkina

Voronezh Institute of High Technologies (Voronezh, Russia)

The article discusses the features of deep learning of neural networks. Methodologies such as the reinforcement system, supervised learning, the method of local minimum or maximum function in supervised learning are considered. A modification of gradient descent – stochastic gradient descent – is described, examples of solving problems based on this tool are given.

Keywords: artificial intelligence, convolutional networks, error weight, deep learning, machine learning, supervised learning, perceptron, gradient descent.