

А. М. СОЛОВЧУК

Беларусь, Брест, БрГУ имени А. С. Пушкина

ГЛУБОКИЕ НЕЙРОННЫЕ СЕТИ

Глубокая нейронная сеть (ГНС) представляет собой сложную программу, состоящую из большого количества внутренних (скрытых) слоев с настраиваемыми параметрами – весовыми коэффициентами искусственных нейронов, составляющих каждый слой сети. На первый, входной слой сеть принимает вектор признаков, описывающих объект, – данные в виде сигналов. На внутренних слоях происходит их обработка: входной вектор умножается на матрицу связей, и сформированный таким образом вектор новых признаков передается в следующий слой. Результат обработки сигнала отправляется на выходной слой сети.

Каждая нейросеть содержит множество параметров, которые невозможно настроить вручную, и поэтому мы обучаем нейросеть на массиве данных: в процессе обучения и переобучения весовые коэффициенты нейронов непрерывно меняются и настраиваются так, что результат вычислений и обработки сигналов становится осмысленным. Настройка весовых коэффициентов производится при помощи простого алгоритма оптимизации, основанного на методе градиентного спуска, используя который можно проследить, как меняется результат обработки сигналов нейросетью при небольшом изменении каждого ее весового коэффициента. Настройка весов происходит в каждом слое нейросети, именно поэтому ее и называют глубокой [1].

Сеть продвигается по слоям, рассчитывая вероятность каждого выхода. Например, ГНС, которая обучена распознавать породы собак, пройдет по заданному изображению и вычислит вероятность того, что собака на изображении относится к определенной породе. Пользователь может просмотреть результаты и выбрать вероятности, которые должна отображать сеть (выше определенного порога, например), и вернуть в сеть предположенную метку. Каждое математическое преобразование считается слоем, а сложные ГНС имеют много слоев, отсюда и название «глубинные» или «глубокие» сети.

Чем сложнее устроена нейросеть, т. е. чем больше слоев и нейронов она имеет и чем больше вычислительных операций она выполняет, тем, как правило, лучший результат она выдает, но и тем труднее нам понять, что происходит в ее внутренних слоях. Но в последнее время мы стали лучше разбираться в том, как выполняются операции в скрытых слоях нейронной сети, благодаря появлению новых алгоритмов, помогающих визуализировать все процессы внутри нее.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Глубокие нейронные сети: пути применения [Электронный ресурс]. – Режим доступа: <https://postnauka.ru/longreads/155983>. – Дата доступа: 27.09.2021.

А. М. СОЛОВЧУК

Беларусь, Брест, БрГУ имени А. С. Пушкина

МЕТОДЫ ОБУЧЕНИЯ НЕЙРОННЫХ СЕТЕЙ

Прежде чем начинать обучать нейросеть для решения любой задачи, необходимо собрать и подготовить обучающую выборку. Важно собрать большой массив данных, поскольку, даже имея нейросеть с простой архитектурой, вы получите лучший результат, чем если бы задачу решала нейросеть со сложной архитектурой, но обучаясь на меньшей выборке. Набирает популярность обучение на смешанных наборах данных, когда сеть обучается на двух наборах данных – большом и разнообразном, но либо имеющем ошибки в разметке данных, либо вообще неразмеченном, а другой набор данных правильно обработан и размечен, но маленький и недостаточно разнообразный. В таком случае необходимо придумать специальный алгоритм, с помощью которого можно будет обучать нейросеть по обоим таким выборкам одновременно.

Среди парадигм обучения искусственных нейронных сетей различают алгоритмы обучения с учителем и без учителя. Первый предполагает наличие обучающей выборки данных, которая состоит из пар входных и целевых векторов (т. е. известны условия задачи и само решение) и которую предъявляют нейронной сети. Каждый обучающий пример подается на вход сети, далее во внутренних слоях происходит его обработка, вычисляется выходной сигнал сети, и он сравнивается со значением целевого вектора – ожидаемым результатом. На основе этого вычисляется сигнал ошибки, который учитывается при дальнейшей перенастройке весовых коэффициентов нейронов сети. Процесс такой корректировки весов происходит, пока значение ошибки не станет минимальным на проверочном наборе данных. Благодаря этому сеть учится повторять разметку данных, сделанную учителем, и делать разумные предсказания для новых данных [1].

Обучение без учителя происходит тогда, когда известны только входные векторы, на основе которых нейронная сеть учится искать в этих данных какие-то закономерности и давать лучший результат – значение выходного сигнала. Какие именно векторы будут сформированы на выходе, зависит от конкретного типа обучения без учителя.