

УДК 004.021

ОБУЧЕНИЕ НЕЙРОСЕТЕЙ

В.С. РАДЧЕНКО*(Представлено: Д.В. ПЯТКИН)*

В статье представлены методы для обучения нейронов в нейросетях и функции, которые помогут в понимании данного процесса.

1 Обучение нейронных сетей

Процесс обучения нейронной сети является необходимым, для ее способности выполнять поставленные задачи. Этот процесс может быть рассмотрен как настройка архитектуры сети и весов связей посредством моделирования среды, в которую эта сеть встроена. Свойство нейронной сети обучаться на примерах позволяет упростить задание условий для решения конкретной задачи, по сравнению с системами, которые следуют набору определенных правил, составленных экспертами. Разделяют алгоритмы обучение с учителем и без учителя [1].

Процесс обучения с учителем подразумевает предъявление сети выборки обучающих примеров. Каждый образец подается на вход нейронной сети, затем проходит процедуру обработки внутри ИНС. После вычисления выходного сигнала ИНС сравнивает полученный результат с соответствующим значением целевого вектора, представляющего собой требуемый выход сети. Вычислив ошибку, происходит изменение весовых коэффициентов связей внутри сети по выбранному алгоритму. Веса подстраиваются под каждый вектор до тех пор, пока ошибка по всему массиву входных данных не достигнет заданного уровня.

Обучение без учителя не требует знания правильного ответа на каждый пример обучающей выборки. В этом случае раскрывается внутренняя структура данных или корреляция в системе данных, что позволяет распределить образы по категориям.

В зависимости от решаемой задачи в обучающей выборке используются те или иные типы данных и различные размерности входных/выходных сигналов. Входные данные примеров обучающей выборки – изображения, таблицы чисел, распределения. Типы входных данных – бинарные (0 и 1), биполярные (-1 и 1) числа, целые или действительные числа из некоторого диапазона. Выходные сигналы сети – вектора целых или действительных чисел. Для решения практических задач часто требуются обучающие выборки большого объема. Из-за жестко ограниченного объема оперативной памяти компьютера разместить в ней большие обучающие выборки невозможно. Поэтому выборка делится на страницы – группы примеров. В каждый момент времени лишь одна страница примеров располагается в памяти компьютера, остальные – на жестком диске. Страницы последовательно загружаются в память компьютера. Обучение сети происходит по всей совокупности страниц примеров, по всей обучающей выборке.

В настоящее время отсутствует универсальная методика построения обучающих выборок. Набор обучающих примеров формируется по усмотрению пользователя программы моделирования нейронных сетей индивидуально для каждой конкретной решаемой задачи [2].

Если в необученную нейронную сеть ввести входной сигнал одного из примеров обучающей выборки, то выходной сигнал сети будет существенно отличаться от желаемого выходного сигнала, определенного в обучающей выборке. Функция ошибки численно определяет сходство всех текущих выходных сигналов сети и соответствующих желаемых выходных сигналов обучающей выборки. Наиболее распространенной функцией ошибки является среднеквадратичное отклонение. Однако предложены и другие функции ошибки.

Для обучения нейронных сетей могут быть использованы различные алгоритмы. Можно выделить две большие группы алгоритмов – градиентные и стохастические. Градиентные алгоритмы обучения сетей основаны на вычислении частных производных функции ошибки по параметрам сети. Среди градиентных различают алгоритмы первого и второго порядков. В стохастических алгоритмах поиск минимума функции ошибки ведется случайным образом [2].

При обучении сетей, как правило, используется один из двух следующих критериев останова: останов при достижении некоторого малого значения функции ошибки, останов в случае успешного решения всех примеров обучающей выборки.

Перед обучением выполняется инициализация нейронной сети, то есть присваивание параметрам сети некоторых начальных значений. Как правило, эти начальные значения – некоторые малые случайные числа.

Для формирования обучающих выборок, инициализации и обучения в программах моделирования нейронных сетей используются специальные процедуры. Возможность использования многостраничного обучения является очень важной при решении практических задач с помощью нейронных сетей, моделируемых на обычных компьютерах.

Обучение – это итерационная процедура, которая при реализации на обычных компьютерах, требует значительного времени. Алгоритмы обучения существенно различаются по скорости сходимости. Одной из самых важных характеристик программ для моделирования нейронных сетей является скорость сходимости алгоритма (или алгоритмов) обучения, которые реализованы в программе [2].

Теория обучения рассматривает три фундаментальных свойства, связанных с обучением по примерам: емкость, сложность образцов и вычислительная сложность. Под емкостью понимается, сколько образцов может запомнить сеть, и какие функции и границы принятия решений могут быть на ней сформированы. Сложность образцов определяет число обучающих примеров, необходимых для достижения способности сети к обобщению. Слишком малое число примеров может вызвать «переобученность» сети, когда она хорошо функционирует на примерах обучающей выборки, но плохо – на тестовых примерах, подчиненных тому же статистическому распределению. Известны 3 основных типа правил обучения: коррекция по ошибке, машина Больцмана и правило Хебба.

Правило коррекции по ошибке. При обучении с учителем для каждого входного примера задан желаемый выход d . Реальный выход сети y может не совпадать с желаемым. Принцип коррекции по ошибке при обучении состоит в использовании сигнала $(d-y)$ для модификации весов, обеспечивающей постепенное уменьшение ошибки. Обучение имеет место только в случае, когда перцептрон ошибается. Известны различные модификации этого алгоритма обучения [2].

Обучение Больцмана. Представляет собой стохастическое правило обучения, которое следует из информационных теоретических и термодинамических принципов. Целью обучения Больцмана является такая настройка весовых коэффициентов, при которой состояния видимых нейронов удовлетворяют желаемому распределению вероятностей. Обучение Больцмана может рассматриваться как специальный случай коррекции по ошибке, в котором под ошибкой понимается расхождение корреляций состояний в двух режимах.

Правило Хебба. Самым старым обучающим правилом является постулат обучения Хебба. Хебб опирался на следующие нейрофизиологические наблюдения: если нейроны с обеих сторон синапса активизируются одновременно и регулярно, то сила синаптической связи возрастает. Важной особенностью этого правила является то, что изменение синаптического веса зависит только от активности нейронов, которые связаны данным синапсом. Это существенно упрощает цепи обучения.

Для того что бы проверить навыки, приобретённые нейронной сетью в процессе обучения, используется имитация функционирования сети. В сеть вводится некоторый сигнал, который, как правило, не совпадает ни с одним из входных сигналов примеров обучающей выборки. Далее анализируется получившийся выходной сигнал сети. Тестирование обученной сети может проводиться на одиночных входных сигналах, либо на контрольной выборке, которая имеет структуру, аналогичную обучающей выборке [2].

1.1 Алгоритм обратного распространения ошибки

Алгоритм обратного распространения ошибки (back propagation) относится к методу обучения с коррекцией по ошибке и, как правило, применяется к многослойным перцептронам. Это метод обучения «с учителем», при котором «наставник» обучает сеть, также как ребенка обучают читать и писать. При обучении на входной слой многократно подаются образы сигналов, распознаванию которых нейронная сеть должна быть обучена, и корректируются веса нейронов для достижения желаемого выходного сигнала. Для улучшения качества распознавания образы, подаваемые на входной слой, могут быть слегка изменены (добавлен шум и т.п.). Детально процедура обучения выглядит так:

1) Выборка входных данных (множество образов, классифицированных учителем) разбивается на две: обучающую и контролирующую последовательности. Обычно обучающая последовательность содержит больше образов, чем проверочная.

2) Производится инициализация всех весов, включая пороговые, небольшими случайными величинами (обычно в диапазоне $[-1; +1]$). Это определяет начальную точку на поверхности ошибок для метода градиентов, позиция может оказаться решающей для сходимости сети.

3) Производится прямой проход сети для первого образа из обучающей выборки от входного слоя через скрытые слои к выходному слою: каждый нейрон суммирует произведения входов на веса и выдает результат функции активации, примененной к этой сумме, на нейроны следующего уровня.

4) Вычисляется разность между действительным и желаемым выходным значением каждого нейрона выходного слоя. При их несовпадении имеет место ошибка в распознавании (классификации) образов.

5) Производится процедура обратного распространения этих ошибок по связям от выходных нейронов к входным и определяются ошибки для каждого нейрона. Рассмотрим эту процедуру подробнее. Предположим, выходное значение нейронной сети было 0.5, а желаемое 0. Пусть ошибка определяется по формуле

$$E = y_i - d_i,$$

где E – ошибка сети,

y_i – желаемое значение на i -ом выходе нейронной сети,

d_i – выходное значение сети на i -ом выходе.

В данном случае, на выходе ошибка будет равна $E = 0.5 - 0 = 0.5$. Если рассматривать в качестве функции активации сигмоид, то расчет Δw , на который необходимо сдвинуть веса синаптических связей предпоследнего последнего слоя, будет производиться по формуле

$$\Delta w = E * \text{sigmoid}(x)dx$$

Здесь $\text{sigmoid}(x)dx$ – производная функции активации, равная $\text{sigmoid}(x)(1-\text{sigmoid}(x))$. Под x подразумевается значение, которое было получено путем суммирования всех выходных значений нейронов предпоследнего слоя, соответственно умноженных на w_i . Завершающий шаг коррекции весов синаптических связей между последним и предпоследним слоем будет производиться по следующей формуле

$$w_i = w_i - m_i * \Delta w * l,$$

где m_i – выходные значения i нейрона предпоследнего слоя,

l – параметр, отвечающий за скорость обучения.

После того как веса посчитаны, мы можем посчитать ошибку выходного значения m_i по формуле $E = w_i * \Delta w$. Зная ошибку каждого нейрона на предыдущем слое, применяя формулы, которые использовались выше, можем найти отклонения весов и значения нейронов на предыдущих слоях.

б) Вновь производится прямой проход сети уже для очередного образа обучающей выборки...

Шаги 3–6 повторяются до достижения некоторого критерия, например, достижения ошибкой установленного предела.

По завершению обучающей фазы сеть проверяется при помощи контролирующей последовательности, содержащей образы, не предъявленные ранее. (При этом не производится коррекция весов, а лишь вычисляется ошибка). В случае, если качество работы найдено удовлетворительным, сеть считается готовой к работе. В противном случае сеть подвергается повторному обучению, при котором возможно изменение некоторых параметров (начальные веса, количество нейронов в скрытых слоях, дополнительные обучающие образы и т.п.) [3].

Заключение

При обучении нейросетей используются можно использовать несколько способов основные из которых перечислены в данной статье.

ЛИТЕРАТУРА

1. Нейронные сети для обработки информации. С. Осовский. Пер. с польского И.Д. Рудинского. – М. : Финансы и статистика, 2012. – 344 с.
2. Neural Networks: A Systematic Introduction. R. Rojas. Springer Science & Business Media – 502 с.
3. Лекции по нейронным сетям [Электронный ресурс] – <https://www.intuit.ru/studies/courses/6/6/lecture/178> (15.09.2019).