

УДК 004.822

СОЗДАНИЕ МНОГОСЛОЙНОГО ПЕРСЕПТРОНА С ИСПОЛЬЗОВАНИЕМ СЕЛЕКЦИОННОГО АЛГОРИТМА АДАПТАЦИИ СТРУКТУРЫ СЕТИ

© Т.А. Калистратов

Ключевые слова: аппроксимация; нейронные сети; структура нейронной сети.

Предложен метод селекционной адаптации структуры нейронной сети, рассмотрены его положительные и отрицательные особенности, представлена блок-схема алгоритма адаптации структуры сети, а также результаты вычислительного эксперимента, подтверждающие работоспособность алгоритма.

Традиционные алгоритмы обучения нейронной сети (например, алгоритм обратного распространения ошибки) подразумевают нахождение наилучших коэффициентов этой сети при ее фиксированной структуре. Проблема выбора подходящей структуры нейронной сети при таком подходе лежит на плечах моделирующего и во многом предрешает успешность построения модели. Помимо этого, сама по себе проблема подбора структуры сети является крайне сложной.

Для упрощения построения нейронных сетей в настоящее время существует множество методов и алгоритмов. Все они могут быть разделены на две группы: методы и алгоритмы наращивания структуры сети и методы и алгоритмы упрощения структуры сети. Методы упрощения структуры сети являются более формализованными, однако они связаны с заведомо крупными затратами ресурсов на обучение начальных структур нейронных сетей. Алгоритмы наращивания структуры сети, как правило, опираются на эмпирические данные об улучшении степени обучения нейронной сети при наращивании ее структуры. Так, в [1] описан алгоритм наращивания структуры нейронной сети, основанный на последовательном заполнении скрытых слоев сети нейронами до состояния насыщения (в котором при добавлении нового нейрона в данный слой невязка обученной нейронной сети практически перестает уменьшаться), эмпирически ограничивая при этом число скрытых слоев сети.

В рассмотренном выше алгоритме отсутствует жесткая фиксация числа скрытых слоев, однако их количество можно ограничить двумя внутренними слоями ввиду того, что в [2] теоретически доказана достаточность одного внутреннего слоя, а также выводится эмпирическое утверждение об оптимальном количестве скрытых слоев, равном двум. Заметим, что использование нейронных сетей с большим количеством внутренних слоев может привести к получению удовлетворительных результатов обучения сети при меньших ее размерах в некоторых конкретных примерах, но не может существенно повлиять на уровень невязки обученной сети при отсутствии жестких ограничений на ее размер. На основе данного утверждения можно построить генетический алгоритм адаптации структуры нейронной сети, который будет обладать рядом преимуществ:

1) в отличие от алгоритмов последовательного заполнения скрытых слоев нейронной сети он не имеет шагов, на которых происходит потеря вычислительной мощности сети (такое может происходить при добавлении нового скрытого слоя, содержащего один нейрон – в этом случае такой нейрон становится «узким местом» всей сети и может приводить к большой невязке обученной сети;

2) данный алгоритм учитывает возможность изменения уровня насыщения скрытых слоев сети при наращивании структуры других скрытых слоев;

3) можно утверждать, что на каждом шаге наращивания структуры нейронной сети мы располагаем сетью с оптимальным распределением нейронов по скрытым слоям, что обеспечивает более широкий выбор критериев останова процесса адаптации структуры сети.

Сам алгоритм адаптации структуры нейронной сети основан на генах добавления нового нейрона в определенный скрытый слой и выглядит следующим образом:

- 1) задать начальную структуру нейронной сети N ;
 - 2) создать копию нейронной сети N' ;
 - 3) добавить для сети N новый нейрон в первый скрытый слой, для сети N' – во второй скрытый слой;
 - 4) произвести обучение сетей N и N' ;
 - 5) если невязка сети N' меньше невязки сети N , присвоить $N = N'$;
 - 6) проверить выполнение критерия останова.
- В случае его выполнения закончить адаптацию структуры сети, иначе вернуться к шагу 2.

Описанный алгоритм имеет некоторые недостатки в сравнении с алгоритмом последовательного наращивания сети. Первым его недостатком можно считать жесткое задание количества скрытых слоев. Однако заметим, что данное количество может быть отличным от двух. Алгоритм позволяет наращивать сеть с n скрытыми слоями, однако в этом случае придется работать с n конкурирующими генами и n параллельно обучающимися нейронными сетями. Отсюда вытекает второй недостаток – описанный алгоритм адаптации структуры нейронных сетей работает в n раз медленнее, чем алгоритмы последовательного наращивания.

Таким образом, определив заранее число скрытых слоев нейронной сети, можно построить генетический алгоритм адаптации структуры нейронной сети, кото-

рый на каждой итерации будет предоставлять нейронную сеть с оптимальным распределением нейронов по скрытым слоям.

Данный алгоритм можно рассматривать как генетический в широком смысле определения генетического алгоритма, т. к. этап копирования нейронной сети представляет собой механизм наследования, этап различной модификации каждой нейронной сети – механизм мутации, а последующий выбор наилучшей нейронной сети – механизм отбора. Однако в отличие от традиционных генетических алгоритмов нет необходимости в накладывании дополнительных ограничений на размер нейронной сети, кроме тех, что входят в критерий останова обучения.

Реализация алгоритма обучения нейронной сети с адаптацией структуры с использованием селекционного метода

Блок-схема описанного ранее алгоритма изображена на рис. 1.

Выполнение блоков 1, 2, 6 происходит, в точности следуя алгоритму обратного распространения ошибки. Введенная первичная структура нейронной сети определяет дальнейшие возможности развития и адаптации структуры сети в рамках данного алгоритма: наиболее важными характеристиками являются количество внутренних слоев сети и наличие или отсутствие связей между нейронами. Алгоритм может работать с любой начальной структурой многослойного персептрона, однако на данный момент остановимся на следующей структуре как наиболее универсальной: сеть содержит два скрытых слоя, каждый из которых содержит по одному нейрону. Сеть является полносвязной, т. е. каждый нейрон i -го слоя связан со всеми нейронами $(i+1)$ -го и $(i-1)$ -го слоев.

Критерий окончания обучения в блоке 3 может быть как абсолютным (преодоление уровнем ошибки сети некоторого порогового значения), так и относительным (уменьшение ошибки сети при наращивании элементов прекратилось или замедлилось). Помимо этого, нельзя забывать об ограничениях на размер нейронной сети, не связанных с уровнем ее ошибки – эти ограничения определяются имеющимися вычислительными мощностями и размером обучающей выборки.

Блок 4 определяет метод копирования нейронной сети. Являясь тривиальным в концептуальном смысле, он представляет достаточно высокую сложность при реализации алгоритма. Для реализации копирования необходимо задать линейное упорядочение для всех нейронов сети, т. е. пронумеровать их. Первыми копируются нейроны без связей таким образом, чтобы нумерация нейронов осталась неизменной. Перенос каждой связи осуществляется с использованием не самих сущностей нейронов, а их номеров. Веса связей, очевидно, копируются. Количество копий нейронных сетей должно равняться количеству скрытых слоев сети.

Блок 5 осуществляет модификацию структуры каждой нейронной сети. В каждую нейронную сеть добавляется один нейрон. Слой, в который происходит добавление нейрона, определяется порядковым номером копии сети. То есть в каждой сети наращивается один из слоев, остальные при этом остаются без изменений. Если модифицируемая сеть является полносвязной, место добавления нейрона не имеет значения.



Рис. 1. Алгоритм адаптации структуры нейронной сети

Если же в сети некоторые связи отсутствуют, необходимо произвести поиск «наиболее узкого» места. Оно определяется следующим образом: тот нейрон в модифицируемом слое, который сохранил наибольшую ошибку, копируется. Связи для нового нейрона оказываются идентичны связям нейрона с наибольшей ошибкой, однако их коэффициенты заменяются случайными величинами.

В блоке 7 для выбора наилучшей нейронной сети сравниваются их интегральные ошибки. Сеть с наименьшей ошибкой оставляется, остальные уничтожаются. Иными словами, модификация, которая привела к наибольшему улучшению функционирования нейронной сети, принимается в ее «базовую» структуру, остальные модификации отбрасываются.

ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ

Для проверки работоспособности алгоритма с его помощью была построена и обучена нейронная сеть для решения задачи умножения двух чисел.

В качестве обучающей выборки используются множители и произведения чисел от 0,1 до 10,0 с шагом 0,1. Всего обучающая выборка содержит 10000 примеров.

Обучающая выборка разбивается на две части – 90 % выборки случайным образом перемешивается перед каждой эпохой обучения и используется для обучения и корректировки весов нейронной сети. 10 % выборки используется для проверки результатов обучения, по этим элементам обучающей выборки веса сети не корректируются. Разбиение выборки на обучающее и проверочное множества происходит случайным образом.

В качестве критерия обученности нейронной сети используется невязка сети, находящаяся как среднее арифметическое ошибок всех выходных нейронов по всем примерам проверочного множества. Структура нейронной сети считается обученной, если по окончании очередной эпохи невязка сети не уменьшилась. Аналогичный критерий использовался для прекращения наращивания структуры нейронной сети.

Обучение нейронной сети происходит согласно формулам (4.1–4.25), описанным в [2, с. 225-232]. Параметр скорости обучения $n = 0,03$.

Начальная структура нейронной сети представляет собой два входных нейрона. Один выходной нейрон и два скрытых слоя по одному нейрону на каждом. Функция активации выходного нейрона линейная, $y = x$. Функция активации всех скрытых нейронов сигмоидальная, $y = \frac{1}{1+e^{-x}}$. Сеть является слоисто-полносвязной, начальные значения коэффициентов связей нейронов задаются случайным образом.

По результатам моделирования получены следующие данные адаптации структуры нейронной сети: адаптация нейронной сети для решения поставленной задачи происходила в 10 этапов (итераций). На первых 7 этапах адаптации нейрон добавлялся в первый скрытый слой, на следующих двух этапах – во второй скрытый слой. На последнем этапе нейрон был добавлен в первый скрытый слой. Невязка полученной в результате нейронной сети была равна 0,00113.

РЕЗУЛЬТАТЫ

Полученные в ходе численного эксперимента данные показывают работоспособность алгоритма и практическую обоснованность подхода к непоследовательному наращиванию структуры нейронной сети.

ЛИТЕРАТУРА

1. Арзамасцев А.А., Крючин О.В., Азарова П.А., Зенкова Н.А. Универсальный программный комплекс для компьютерного моделирования на основе искусственной нейронной сети с самоорганизацией структуры // Вестник Тамбовского университета. Серия Естественные и технические науки. Тамбов, 2006. Т. 11. Вып. 4. С. 564-570.
2. Хайкин С. Нейронные сети: полный курс, 2-е изд.: пер. с англ. М.: Издат. дом «Вильямс», 2006.

Поступила в редакцию 21 мая 2015 г.

Kalistratov T.A. CREATION OF MULTI-LAYER PERCEPTRON WITH THE USE OF SELECTION ALGORITHM OF ADAPTATION OF NETWORK STRUCTURE

A method of selection adaptation of neural network's structure is provided, its positive and negative features are considered, a diagram of a network structure of the adaptation algorithm is submitted and also the results of computational experiment confirming the performance of the algorithm are proposed.

Key words: approximation; neural networks; structure of neural network.

Калистратов Тимофей Александрович, Тамбовский государственный университет им. Г.Р. Державина, г. Тамбов, Российская Федерация, аспирант, кафедра компьютерного и математического моделирования, e-mail: stubbytim@mail.ru

Kalistratov Timofey Aleksandrovich, Tambov State University named after G.R. Derzhavin, Tambov, Russian Federation, Post-graduate Student, Computing and Mathematical Modeling Department, e-mail: stubbytim@mail.ru