
УДК 004.032.26:629.113

И. Ф. Дьяков

ИСПОЛЬЗОВАНИЕ НЕЙРОННОЙ ТЕХНОЛОГИИ ПРИ ВЫБОРЕ ТРАНСПОРТНОГО СРЕДСТВА

UDC 004.032.26:629.113

I. F. Dyakov

NEURAL TECHNOLOGY USAGE DURING VEHICLE SELECTION

Аннотация

Рассмотрен метод выбора транспортного средства с использованием нейронной технологии. Показан принцип работы нейронной сети с конечным числом слоев из однотипных элементов – аналогов нейронов с различными типами связей между слоями. Для сокращения времени решения задачи число нейронов в слоях выбирается исходя из необходимости обеспечения заданной точности решения задачи, а число слоев нейронов – как можно меньше. Приведены результаты применения нейронной технологии при выборе транспортного средства.

Ключевые слова:

нейронная технология, синапсы, архитектура нейронной сети, фаза обучения, функция ошибки, дендриты.

Abstract

This article describes some methods of selection of vehicles by using neural technology. It describes the principle of the neural technology with finite number of layers made of one-type elements, which are analogues with neurons with different types of connections. In order to reduce time for solving the problem, the number of neurons in layers is selected based on the need to ensure a specified accuracy of solving the problem and the number of layers of neurons shall be the least possible. The results of neural technology usage during vehicle selection are given.

Key words:

neural technology, synapses, architecture of neural network, training phase, error function, dendrites.

В период развития рыночной экономики актуален вопрос оснащения автотранспортных предприятий необходимым количеством (в зависимости от годового объема перевозимого груза) и качественным составом (в зависимости от разновидности грузов с учетом используемого вида перевозки) грузовых автомобилей. Решить данный вопрос можно путем моделирования конкуренции на рынке. Рынок состоит из большого числа независимых заводоизготовителей, автотранспортных предприятий, грузоотправителей и получа-

телей, самостоятельно определяющих тактику и стратегию своего поведения. В условиях равной осведомленности каждый из участников рыночного процесса стремится максимизировать свою выгоду. Найти оптимальную модель из множества параметров – задача математическая. Наиболее совершенным методом выбора транспортного средства является нейронная технология [1].

Появилось множество работ по новым математическим моделям и обучающим алгоритмам, называемым нейронными сетями [2, 3]. Искусствен-

ная нейронная сеть – это набор нейронов, соединенных между собой. Полносвязанной нейронной сетью называют многослойную структуру, в которой каждый нейрон произвольного слоя связан со всеми нейронами предыдущего слоя, а в случае первого слоя – со всеми входами нейронной сети.

При подаче любых чисел на входы сети получается набор чисел на ее выходах. Работа нейросети состоит в преобразовании входного вектора в выходной, причем это преобразование задается «весами» сети. Главная функция искусственного нейрона – формировать выходной сигнал в зависимости от сигналов, поступающих на его входы. Нейросети используют множество простых вычислительных элементов, называемых нейронами, каждый из которых имитирует поведение отдельной клетки человеческого мозга. Каждый нейрон в нейронной сети осуществляет преобразование входных сигналов в выходной сигнал и связан с другими нейронами. Входные нейроны формируют так называемый интерфейс нейронной сети.

Существует большое число алгоритмов обучения, ориентированных на решение разных задач. Среди них выделяется алгоритм обратного распространения ошибки, который является одним из наиболее успешных современных алгоритмов. Его основная идея заключается в том, что изменение «веса» синапсов происходит с учетом локального градиента функции ошибки. Разница

между реальными и правильными ответами нейронной сети, определяемыми на выходном слое, распространяется в обратном направлении – навстречу потоку сигналов. Для увеличения или уменьшения активности данного нейрона другими нейронами существуют так называемые *синапсы*. Они переносят величину активности от нейрона-отправителя к нейрону-получателю. Если синапс является возбуждающим, то величина активности нейрона-отправителя увеличивает активность нейрона-получателя. Если синапс тормозящий, то величина активности нейрона-отправителя уменьшает активность нейрона-получателя. Синапсы различаются не только по признаку торможения, но и по суммарному воздействию.

Математическую модель нейрона можно представить в виде

$$x_k^{(i+1)} = f \left(\sum_{j=1}^N w_j^{(k)} x_j^{(i)} \right),$$

где $x_j^{(i)}$ – выходной сигнал данного нейрона слоя $(i + 1)$; $w_j^{(k)}$ – коэффициент «веса» синаптической связи.

Нейронная сеть прямого распространения сигналов, показанная на рис. 1, имеет слой, принимающий входные сигналы, и слой, генерирующий выходные сигналы. Прямое распространение сигнала означает, что такая нейронная сеть не содержит петель.

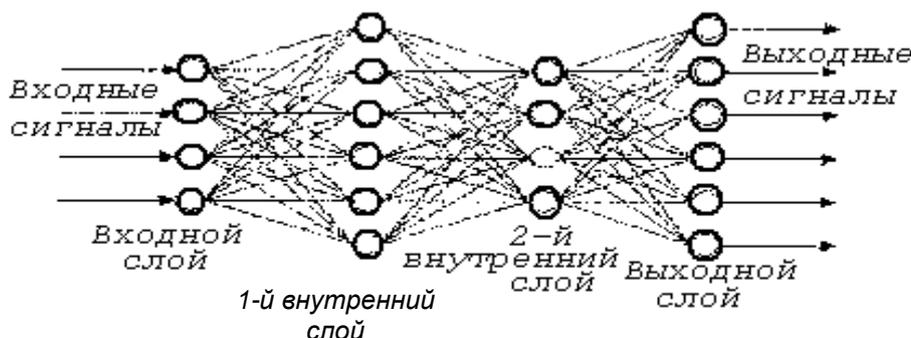


Рис. 1. Структура нейронной сети

Информация вводится в нейронную сеть через входной слой. Он бывает однослойный и многослойный. Все слои нейронной сети обрабатывают сигналы до тех пор, пока они не достигнут выходного слоя. В каждом слое нейроны между собой никак не связаны, но связаны с нейронами предыдущего и следующего слоев. Информация поступает с первого на второй слой, со второго – на третий и т. д. Задача нейронной сети – преобразование информации требуемым образом. Для этого сеть предварительно обучается. При обучении используются идеальные (эталонные) значения пар «входы–выходы» или «учитель», который оценивает поведение нейронной сети.

Понятие «учитель» применимо либо к математической функции, либо к лицу, которое оценивает качество поведения нейронной сети. Поскольку нейронные сети в основном используются в сложных расчетах, где нет хороших математических моделей, то обучение производится с помощью обучающей выборки (так называемого обучающего алгоритма), т. е. эталонных пар «входы-выходы».

Первый шаг в проектировании нейронной сети состоит в ее обучении. Это – фаза обучения, в результате которой нейронная сеть будет вычислять сигналы, близкие к эталонным данным при соответствующих входных сигналах.

Функция ошибки представляет собой разность между текущим выходом сети и идеальным, который необходимо получить. Для успешного обучения сети требуется приблизить выход сети к оптимальному выходу, т. е. последовательно уменьшить величину функции ошибки. Это достигается настройкой межнейронных связей. Каждый нейрон в сети имеет свои «веса», которые настраиваются, чтобы уменьшить величину функции ошибки.

Ненастроенная нейронная сеть не способна отображать достоверные результаты. Особый характер сложных задач не позволяет создавать универсальные сети, вынуждая использовать специализированные нейронные сети для каждой задачи. В настоящее время существует большое количество разнообразных типов сетей, которые отличаются структурой связей, правилами обучения и программирования. Возможность обучения – одно из главных преимуществ нейронных сетей перед традиционными алгоритмами. Технически обучение заключается в нахождении коэффициентов связей между нейронами. В процессе обучения нейронная сеть способна выявлять сложные зависимости между входными данными и выходными, а также выполнять обобщение.

Высокая эффективность нейронных технологий при решении задач адаптивного управления динамическими объектами может сделать их незаменимыми при создании новых сетей. Однако, если задача выходит за пределы простых вычислительных операций и затрагивает сферы анализа и обобщения информации, компьютерная техника демонстрирует свою несостоятельность по сравнению с человеком. Архитектура биологической нейронной системы совершенно не похожа на архитектуру машины фон Неймана.

Исследований возможности применения нейронной технологии в организации управления в транспортной сфере крайне мало. Это делает актуальным построение модели конкуренции в условиях рыночной экономики с использованием базовых положений нейронной технологии и основ оптимизации. Для полного представления о нейронных сетях рассмотрим структуру типичного биологического нейрона (рис. 2) и будем применять ее фрагменты при решении поставленной задачи.

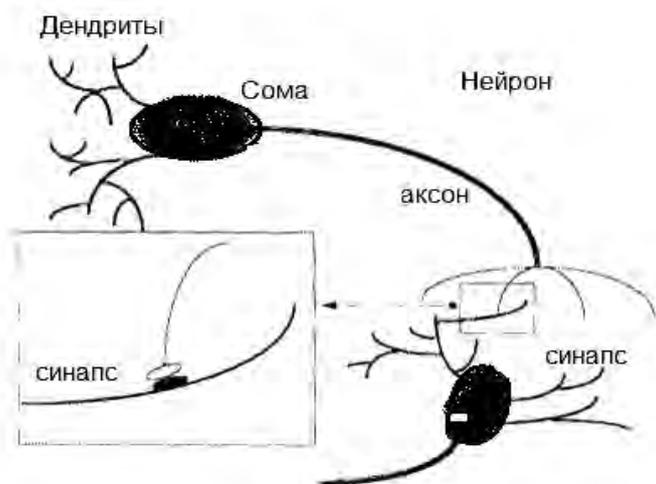


Рис. 2. Схема биологического нейрона

Выходной сигнал каждого нейрона передается по так называемому *аксону*. Входные сигналы, принятые синапсом, подводятся к телу нейрона – *сому*. Здесь они суммируются, причем одни входы стремятся возбудить нейрон, другие – воспрепятствовать его возбуждению. Когда суммарное возбуждение в теле нейрона превышает некоторый порог, нейрон возбуждается, посылая по аксону сигнал другим нейронам. Дендриты ответственны за сбор поступающих сигналов (параметров), сома – за основную обработку и суммирование сигналов, а аксон отвечает за передачу сигналов другим *дендритам*. Дендриты связаны с телом нервной клетки, где они принимают сигналы в точках соединения. Передача далее зависит от ряда факторов, но здесь можно смоделировать это поведение как передаточную функцию, которая принимает входные данные, обрабатывает их и готовит выходные данные, если выполняются свойства передаточной функции. Кроме того, в реальных нейронах вывод данных нелинеен, т. е. не цифровые, а аналоговые сигналы. В сущности, нейрон представляет собой тип памяти.

Нейроны имеют как кратковременную память (STM), так и долговре-

менную (LTM). STM – способность нейронной сети вспомнить что-то, что она узнала только что, а LTM – способность нейронной сети вспомнить что-то, что она узнала некоторое время назад в свете только что полученной информации. Одно из основных применений нейронных сетей – создание механизма воспоминаний, который сможет обрабатывать неполные либо нечеткие входные данные и возвращать результат. Результатом могут быть сами входные данные (ассоциация) или совершенно отличающийся от входных данных ответ (гетероассоциация). Нейронная сеть – это сеть с конечным числом слоев из однотипных элементов – аналогов нейронов с различными типами связей между слоями. При этом число нейронов в слоях выбирается исходя из необходимости обеспечения заданного качества решения задачи, а число слоев нейронов – как можно меньше для сокращения времени решения задачи.

Таким образом, используемый искусственный нейрон имитирует свойства биологического нейрона. Если на вход искусственного нейрона поступает некоторое множество сигналов, каждый из которых является выходом другого нейрона, то каждый вход умножается на

соответствующий ω_i «вес». Импульсы, поступившие к нейрону одновременно по нескольким дендритам, суммируются. Если суммарный импульс превышает некоторый порог, нейрон возбуждается, формирует собственный импульс и передает его далее по аксону. «Веса» синапсов могут изменяться со временем, значит, меняется и поведение соответствующего нейрона. Пусть к синапсам поступают импульсы силы $1, 2, \dots, h$. Тогда после прохождения синапсов и дендритов к нейрону также поступают импульсы. Нейрон преобразует полученный суммарный импульс согласно с некоторой передаточной функцией. Каждый «вес» соответствует «силе» одной биологической синаптической связи, определяя уровень активации нейрона.

Суммирующий блок, соответствующий сине, складывает взвешенные входы алгебраически, создавая выходной сигнал ψ : $\psi = \sum_{i=1}^l h_i \omega_k$, где h – множество входных сигналов; l – количество независимых нейронных сетей. Сигнал ψ далее преобразуется активационной функцией F и дает выходной нейронный сигнал $\psi_i = F(\psi)$. В итоге каждый нейрон способен определить вклад каждого своего «веса» в суммарную ошибку сети. «Весовые» коэффициенты меняются при передаче от слоя к слою по формуле

$$\Delta\omega = -\dot{v} \frac{\partial \delta}{\partial \omega_{ij}^{(k)}},$$

где $\omega_{ij}^{(k)}$ – коэффициент синаптической связи k -го нейрона от i -го слоя к j -му слою; \dot{v} – коэффициент скорости обучения; δ – функция суммарной ошибки сети.

Нейронная сеть может работать только с числовыми данными, из чего следует, что необходимо преобразовы-

вать символьные данные в числовые. Простейший способ заключается в создании таблицы соответствий между символьными данными и числовыми.

Применим вышеописанную концепцию к условиям транспортной конкуренции. Каждый грузовой автомобиль имеет множество конструктивных параметров (коэффициент использования собственной массы, надежность $P(t)$, путевой расход топлива $Q(l)$, время успокоения колебаний поддрессоренной массы $t_{\text{усп}}$, проходимость $P_{f(\sigma)}$, уровень звукового давления в кабине автомобиля $J(\ddot{v}(t))$, устойчивость движения Y_y и т. д.). Эти параметры правомочны для всех грузовых автомобилей. Правомочность суммирующего блока Σ нейронной сети подтверждается тем, что все учитываемые множества конструктивных параметров автомобилей N_i носят ценовой характер, т. е. они прямо влияют на эксплуатационные расходы. Обозначим их через h_p – множество рассматриваемых параметров. Пусть h_p выражает коэффициенты использования собственной массы грузовых автомобилей h_σ . Неценовые параметры автомобиля, такие как распределение нагрузки по осям q_i , максимальная скорость движения V_{max} , экологическая безопасность, виброустойчивость σ_z , не учитываются. В данном случае h_σ можно отразить одним нейроном (рис. 3). Таким образом, N_i -й грузовой автомобиль у разных производителей имеет набор этих параметров, но с различными значениями $(h_{\eta_\sigma}^l, h_{\eta_\sigma}^y, h_{\eta_\sigma}^w)$. Тогда каждый автомобиль будет описываться вектором характеристик, число которых будет равно числу изготовителей грузовых автомобилей.

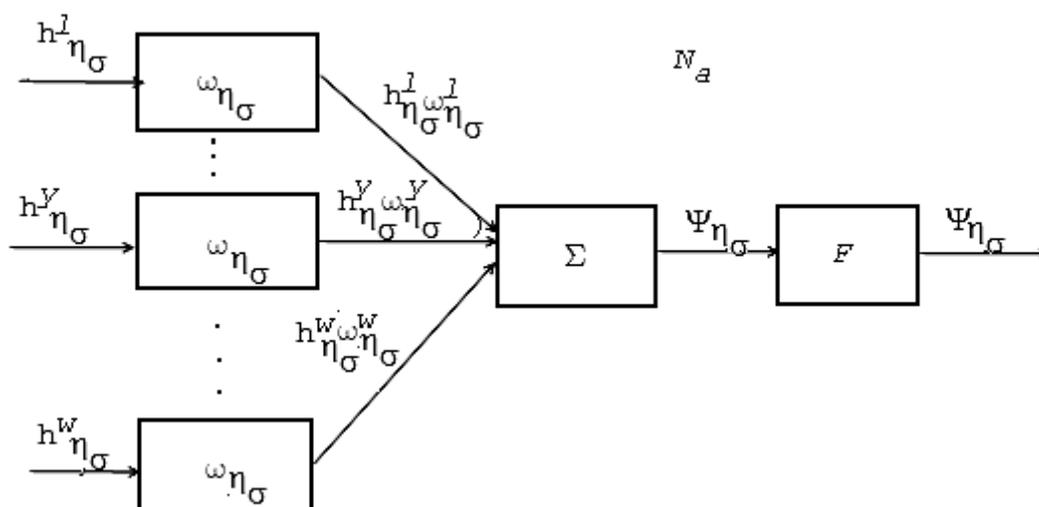


Рис. 3. Модель искусственного нейрона для автомобиля с заданными коэффициентами использования собственной массы η_{σ} разных заводов-изготовителей

Для получения характеристик необходимо построить нейронную сеть. На первом этапе следует выбрать, какие нейроны можно использовать (число входов, передаточные функции), каким образом соединить их между собой, что взять в качестве входов и выходов сети. Здесь важно определить наилучшую структуру сети, т. е. оптимальное число скрытых слоев и число нейронов в них. Непрерывная функция может быть аппроксимирована с высокой точностью нейросетью с единственным слоем, но неограниченным числом нейронов. При ограничении числа нейронов возрастает ошибка сети. Это может потребовать ее периодического переобучения. Нейросеть со скрытым слоем обладает лучшей способностью к обобщению при ограниченном числе нейронов, но процесс ее обучения становится сложнее. Нейроны внутренних слоев, которые связаны с выходами, имеющими большую погрешность, должны изменять свои «веса» значительно сильнее, чем нейроны, соединенные с корректными выходами. Следовательно, «веса» нейрона должны изменяться прямо пропорционально ошибке тех нейронов, с которыми

данней нейрон связан.

В силу этого обратное распространение ошибок через сеть позволяет корректно настраивать «веса» связей между всеми слоями. В таком случае величина функции ошибки уменьшается и сеть обучается. Рассмотрим основные процедуры расчета распространения ошибки на одном параметре коэффициента использования собственной массы h_{σ} для N_i -го грузового автомобиля при следующих условиях: $\delta_{h_{\sigma}}$ – величина функции ошибки для грузового автомобиля; $[h_{\sigma j N_i}]$ – допускаемый показатель выхода j -го нейрона; $\bar{h}_{\sigma j N_i}$ – действительный показатель выхода j -го нейрона; ω_{ij} – «вес» связи между i -м и j -м нейронами.

Пусть функция ошибки прямо пропорциональна квадрату разности между допускаемым и действительным выходами для всей обучающей выборки:

$$\delta_{h_{\sigma}} = \frac{1}{2} \sum_j ([h_{\sigma j N_i}] - \bar{h}_{\sigma j N_i})^2. \quad (1)$$

Множитель $1/2$ вводится для упрощения операции дифференцирования. Активация каждого j -го нейрона для N_i -го грузового автомобиля записывается в виде взвешенной суммы

$$\text{net}_{N_{ij}} = \sum_i \omega_{ij} \bar{h}_{\sigma_j N_{ij}}. \quad (2)$$

Выход каждого j -го нейрона является значением пороговой функции F_j , которая активизируется взвешенной суммой. Часто используют активационные функции: жесткую пороговую функцию, линейный порог, сигмоидную функцию. В многослойной сети применяют обычно сигмоидную функцию $F(\text{net})$, хотя может использоваться любая непрерывно дифференцируемая монотонная функция

$$\bar{h}_{\sigma_j N_i} = F_j(\text{net}_{N_{ij}}). \quad (3)$$

Сигмоидная функция определяется из выражения [3]

$$F(\text{net}) = \frac{1}{1 + e^{-k\text{net}}}, \quad (4)$$

где k – положительное число, влияющее на растяжение функции: увеличение k сжимает функцию и имеет диапазон $0 < F(\text{net}) < 1$, при $k \rightarrow \infty$ функция $F(\text{net})$ приближается к функции Хевисайда [2]. Данный коэффициент может использоваться в качестве параметра усиления, т. к. для слабых входных сигналов угол наклона сигмоидной функции будет крутым и функция будет изменяться быстро, производя значительное усиление сигнала. Для больших входных сигналов угол наклона функции будет немного меньше. Это означает, что сеть может принимать большие сигналы и при этом оставаться чувствительной к слабым изменениям сигнала. Если выход нейрона $h_{\sigma N_{ij}}$ задается в виде

$$h_{\sigma N_{ij}} = F(\text{net}) = \frac{1}{1 + e^{-k\text{net}}},$$

то производная по отношению к данному нейрону $F'(\text{net})$ вычисляется как

$$\begin{aligned} F'(\text{net}) &= \frac{ke^{-k\text{net}}}{(1 + e^{-k\text{net}})^2} = \frac{kF(\text{net})}{1 - F(\text{net})} = \\ &= k\bar{h}_{\sigma N_{ij}}(1 - \bar{h}_{\sigma N_{ij}}), \end{aligned}$$

т. е. является простой функцией от выходов нейронов.

Функцию ошибки (1) представим через частные производные по правилу цепочки в виде

$$\frac{\partial h_{\sigma N_i}}{\partial \omega_{ij}} = \frac{\partial \bar{h}_{\sigma N_i}}{\partial \text{net}_{N_{ij}}} \frac{\partial \text{net}_{N_{ij}}}{\partial \omega_{ij}}. \quad (5)$$

Для второго сомножителя в (5) правой части, используя (2), получим

$$\begin{aligned} \frac{\partial \text{net}_{N_{ij}}}{\partial \omega_{ij}} &= \frac{\partial}{\partial \omega_{ij}} \sum_k \omega_{kj} \bar{h}_{\sigma N_{ik}} = \\ &= \sum_k \frac{\partial \omega_{jk}}{\partial \omega_{ij}} \bar{h}_{\sigma N_{ik}} = \bar{h}_{\sigma N_{ij}}, \end{aligned} \quad (6)$$

поскольку $\frac{\partial \omega_{kj}}{\partial \omega_{ij}} = 0$, за исключением

$k = i$, когда эта производная равна единице. Изменение ошибки как функция изменения входов нейрона имеет вид:

$$-\frac{\partial \delta_{h_{\sigma}}}{\partial \text{net}_{N_{ij}}} = \Delta_{h_{\sigma j}}. \quad (7)$$

Тогда выражение (5) преобразуется к виду

$$\frac{\partial h_{\sigma N_i}}{\partial \omega_{ij}} = \Delta_{h_{\sigma}} \bar{h}_{\sigma N_{ij}}. \quad (8)$$

Следовательно, уменьшение величины $\delta_{h_{\sigma}}$ означает изменение «веса» пропорционально $k_n \Delta_{h_{\sigma}} \bar{h}_{\sigma N_{ij}} = \Delta_{h_{\sigma j}} \omega_{ij}$,

где k_n – коэффициент пропорциональности, влияющий на скорость обучения. При обучении сети «веса» внутри блоков не меняются. Настройке подлежат только свободные «веса» на выходах фрагментов, которые могут быть получены при помощи алгоритмов минимизации целевой функции. Теперь необходимо найти значение для каждого нейрона. Используя выражение (7) и правило цепочки, можно записать следующее:

$$\Delta h_{\sigma N_{ij}} = -\frac{\partial \delta_{h_{\sigma}}}{\partial \text{net}_{N_{ij}}} = \frac{\partial \delta_{h_{\sigma N_i}}}{\partial \bar{h}_{\sigma N_{ij}}} \frac{\partial \bar{h}_{\sigma N_{ij}}}{\partial \text{net}_{N_{ij}}}. \quad (9)$$

Исходя из выражения (2) записываем второй сомножитель в (8):

$$\frac{\partial \bar{h}_{\sigma N_{ij}}}{\partial \text{net}_{N_{ij}}} = F'_j(\text{net}_{N_{ij}}). \quad (10)$$

Теперь рассмотрим первый сомножитель в (9). Согласно (1) можно получить

$$\frac{\partial h_{\sigma N_i}}{\partial \bar{h}_{\sigma N_i}} = -\left([h_{\sigma N_{ij}}] - \bar{h}_{\sigma N_{ij}} \right). \quad (11)$$

Тогда скорость изменения ошибки для j -го нейрона

$$\Delta h_{\sigma N_{ij}} = F'_j(\text{net}_{N_{ij}}) \left([h_{\sigma N_{ij}}] - \bar{h}_{\sigma N_{ij}} \right). \quad (12)$$

Последнее соотношение является полезным для выходных нейронов, поскольку для них известны целевые и действительные значения выходов. Однако для внутренних слоев целевые значения выходов неизвестны. Если j нейрон невыходной, то повторно используем правила цепочки и соотношения (2) и (7), которые представим в следующем виде:

$$\begin{aligned} \frac{\partial \delta_{h_{\sigma N_i}}}{\partial \bar{h}_{\sigma N_{ij}}} &= \sum_k \frac{\partial \delta_{h_{\sigma N_i}}}{\partial \text{net}_{N_{ik}}} \frac{\partial \text{net}_{N_{ik}}}{\partial \bar{h}_{\sigma N_{ij}}} = \\ &= \sum_k \frac{\partial \delta_{h_{\sigma N_i}}}{\partial \text{net}_{N_{ik}}} \frac{\partial}{\partial \bar{h}_{\sigma N_{ij}}} \sum_i \omega_{jk} \bar{h}_{\sigma N_{ij}}; \end{aligned} \quad (13)$$

$$\begin{aligned} \sum_k \frac{\partial \delta_{h_{\sigma N_i}}}{\partial \text{net}_{N_{ik}}} \frac{\partial}{\partial \bar{h}_{\sigma N_{ij}}} \sum_i \omega_{jk} \bar{h}_{\sigma N_{ij}} &= \\ &= -\sum_k \Delta h_{\sigma N_{ij}} \omega_{jk}, \end{aligned} \quad (14)$$

где k – порядковый номер используемого слоя.

В полученных выражениях сумма по i исчезает, т. к. частная производная не равна нулю лишь в одном случае, так же, как и в (6). Подставим (14) в (10). В результате окончательное выражение изменения ошибки по данному показателю N_i -го грузового автомобиля будет иметь вид:

$$\Delta h_{\sigma N_{ij}} = F'_j(\text{net}_{N_{ij}}) \sum_k \Delta h_{\sigma k} \omega_{jk}. \quad (15)$$

Таким образом, уравнения (12) и (15) составляют базу метода обучения многослойной нейронной сети, который называют обобщенным дельта-правилом. Нейронная сеть может работать только с числовыми данными, из чего следует, что необходимо преобразовывать символьные данные в числовые. Простейший способ заключается в создании таблицы соответствий между символьными данными и числовыми, а более сложный подход – в принятии хэш-функций для создания уникальных числовых данных, соответствующих i -й строке. Одно из главных преимуществ нейронных сетей состоит в том, что они теоретически могут аппроксимировать любую непрерывную функцию, что позволяет не принимать заранее каких-либо предположений относительно рассматриваемой модели.

СПИСОК ЛИТЕРАТУРЫ

1. **Дьяков, И. Ф.** Основы оптимизации в автомобилестроении / И. Ф. Дьяков. – М. : Машиностроение, 2012. – 386 с.
2. **Кофман, А.** Введение теории нечетких множеств в управление предприятиями : пер. с исп. / А. Кофман, Х. Хил Алуха. – Минск : Выш. шк., 1992. – 224 с.
3. **Каллан, Р.** Основные концепции нейронных сетей : пер. с англ. / Р. Каллан. – М. : Изд. дом «Вильямс», 2001. – 283 с.

Статья сдана в редакцию 26 января 2016 года

Иван Федорович Дьяков, д-р техн. наук, проф., Ульяновский государственный технический университет. E-mail: i.dyakov@ulstu.ru.

Ivan Fedorovich Dyakov, DSc (Engineering), Prof., Ulyanovsk State Technical University. E-mail: i.dyakov@ulstu.ru.