ИНФОРМАЦИОННОЕ МОДЕЛИРОВАНИЕ С ПРИМЕНЕНИЕМ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Г. Б. Жамалова
Преподаватель Каршинского инженерно-экономического института
(Узбекистан)

АННОТАЦИЯ

В статье представлены экспертные системы, искусственные нейронные сети, нечеткая логика, генетические алгоритмы и ряд других современных информационных технологий, позволяющих создавать интеллектуальные системы для принятие управленческих решений. Прежде всего, разные умные технологии в основе которых лежит концепции интеллекта, способность человека работать с формализованными знаниями (экспертные системы, нечеткая логика) или специфическими для человека способами обучения и мышления (искусственные нейронные сети и генетические алгоритмы).

Ключевые слова: *Интеллектуальные* системы, искусственные нейронные сети, экспертные системы, нечеткая логика, набор синапсов, сумматор, функция активации, нелинейная модель нейрона, кусочно-линейная функция, сигмоидальная функция.

INFORMATION MODELING USING ARTIFICIAL NEURAL NETWORKS

G. B. Jamalova
Lecturer of Karshi Engineering and Economical Institute, Uzbekistan

ABSTRACT

The article presents expert systems, artificial neural networks, fuzzy logic, genetic algorithms and a number of other modern information technologies that make it possible to create intelligent systems for making management decisions. First of all, various smart technologies based on the concept of intelligence, a person's ability to work with formalized knowledge (expert systems, fuzzy logic) or methods of learning and thinking specific to a person (artificial neural networks and genetic algorithms).

Keywords: Intelligent systems, artificial neural networks, expert systems, fuzzy logic, set of synapses, adder, activation function, nonlinear model of a neuron, piecewise linear function, sigmoidal function.

ВВЕДЕНИЕ

Новые способы представления информации, ее передачи и использования появляются ежегодно. Одними из самых свежих и самых динамичных направлений развития является информационное моделирование с использованием методик искусственных нейронных сетей. Совсем недавно известные только узкому кругу специалистов нейронные сети (НС), нечеткая логика, генетические алгоритмы и ряд других информационных технологий получили широкое применение за последнее десятилетие.Появился термин *интеллектуальные системы* – это системы,способные к «пониманию» и обучению в отношении возмущений, внешней среды и условий работы. Основное отличие интеллектуальных систем – наличие механизма системной обработки знаний.

В основе создания интеллектуальных систем лежат два принципа: ситуационное управление (управление на основе анализа внешних ситуаций или событий) и использование современных информационных технологий обработки знаний.

Существует несколько современных информационных технологий, позволяющих создавать данные системы: экспертные системы, искусственные нейронные сети, нечеткая логика, генетические алгоритмы и ряд других. Интеллектуальные технологии между собой различает прежде сего то, что именно положено в основу концепции интеллектуальности – либо умение работать с формализованными знаниями человека (экспертные системы, нечеткая логика), либо свойственные человеку приемы обучения и мышления (искусственные нейронные сети и генетические алгоритмы).

МЕТОДОЛОГИЯ

Под искусственными нейронными сетями (далее – просто нейронными сетями, НС) подразумевают вычислительные структуры, состоящие из большого количества однотипных элементов, каждый из которых выполняет относительно простые функции. Процессы в искусственных НС иногда ассоциируют с процессами, происходящими в нервной системе живых организмов. НС сходна с человеческим мозгом с двух точек зрения:

- Знания поступают в нейронную сеть из окружающей среды и используются в процессе обучения.
- Для накопления знаний используются связи между нейронами, называемые синаптическими весами.

Искусственные нейронные сети (ИНС) строятся по принципу организации и функционирования их биологических аналогов. Они способны решать широкий круг задач распознавания образов, идентификации, прогнозирования, оптимизации. Дальнейшее повышение производительности компьютеров связывают с ИНС, в частности с нейрокомпьютерами (НК), основу которых составляет искусственная нейронная сеть. Бурно развивающийся в последние годы аппарат НС предназначался в начале, в основном для решения задач классификации, кластеризации и распозна- вания образов, но дальнейшее развитие данного направления значительно расширило сферу применения нейросетевого подхода. Термин «нейронные сети» сформировался к середине 50-х годов XX века. Основные результаты в этой области связаны с именами У. Мак-Каллока, Д. Хебба, Ф. Розенблатта, М. Минского, Дж. Хопфилда. Приведем краткую историческую справку [1]. У. Мак-Каллок (W. McCulloch) и У. Питтс (W. Pitts) в 1943 году предложили модель нейрона и сформулировали основные положения теории функционирования головного мозга.

Д. Хебб (D. Hebb) в 1949 году высказал идеи о характере соединений нейронов мозга и их взаимодействии (клеточные ансамбли, синаптическая пластичность). Впервые в 1957 г. Ф. Розенблатт (F. Rosenblatt) предложил правила обучения нейронной сети и разработал принципы организации и функционирования персептронов, предложил вариант технической реализации первого в мире нейрокомпьютера Mark. Д. Хьюбел (D. Hubel) и Т. Визель (T. Wiesel) в 1959 году показали распределенный и параллельный характер хранения и обработки информации в биологических нейронных сетях.

Активные исследования в области искусственных нейронных сетей велись в 1960–1968 гг., например, Адалина и Мадалина, В. Уидроу (1960–1962 гг.), ассоциативные матрицы К. Штайнбухг (1961 г.). 1969 г. Публикация книги М. Минского (M. Minsky) и С.Пейперта (S.Papert) «Персептроны», в которой доказывается принципиальная ограниченность возможностей персептронов. Угасание интереса к искусственным нейронным сетям. Активные разработки в области персептронов (основные заказчики – военные ведомства) проводились в 1970–1976 гг.

Возобновление интереса к искусственным нейронным сетям как следствие накопления новых знаний о деятельности мозга, а также значительного прогресса в области микроэлектроники и компьютерной техники велись в конце 1970-х гг.

Дж.Хопфилд (J.Hopfield) в 1982–1985 гг. предложил семейство оптимизирующих нейронных сетей, моделирующих ассоциативную память.

Появление в 1985 году первых коммерческих нейрокомпьютеров, например, Mark III фирмы TRW (США). Широкомасштабное финансирование разра-боток в области ИНС и НК в США, Японии и Западной Европе (японская программа «Human Frontiers» и европейская программа «Basic Research in Adaptive Intelligence and Neurocomputing») было начато в 1987 году.

РЕЗУЛЬТАТЫ

Разработки и исследования в области ИНС и НК ведутся с 1989 года практически всеми крупными электротехническими фирмами. Нейрокомпьютеры становятся одним из самых динамичных секторов рынка (за два года объем продаж вырос в пять раз). Агентством DARPA (Defence Advanced Research Projects Agency) министерства обороны США начато финансирование программы по созданию сверхбыстро-действующих образцов НК для разнообразных применений. Активизация советских исследовательских организаций в области ИНС и НК (Институт кибернетики им. Глушкова в Киеве, Институт многопроцессорных вычислительных систем в Таганроге, Институт нейрокибернетики в Ростове-на-Дону) произошли в 1990 г. Общее число фирм, специализирующихся в области ИНС и НК, достигло трехсот. Годовой объем продаж на рынке ИНС и НК в 1991 году приблизился к 140 млн долларам. Создались центры нейрокомпьютеров в Москве, Киеве, Минске, Новосибирске, С.Петербурге. Работы в области ИНС в 1992 году находился на стадии интенсивного развития. Ежегодно проводились десятки международных конференций и форумов по нейронным сетям, число специализированных периодических научных изданий по указанной тематике достигло двух десятков наименований. Число международных конференций по ИНС и НК в 1996 году достигло до ста.

Годовой объем продаж на рынке ИНС и НК в 1997 году превысил 2 млрд долларов, а ежегодный прирост составил 50 %. Переход в 2000 году на субмикронные и нанотехнологии, а также успехи молекулярной и биомолекулярной технологии приводили к принципиально новым архитектурным и технологическим решениям по созданию нейрокомпью-теров.Глубокое изучение ИНС требовало знания нейрофизиологии, науки о познании, психологии, физики (статистической механики), теории управления, теории вычислений, проблем искусственного интеллекта, статистики / математики, распознавания образов, компьютерного зрения, параллельных вычислений и аппаратных средств (цифровых и аналоговых). С другой стороны, ИНС также стимулируют эти дисциплины, обеспечивая

их новыми инструментами и представлениями. Этот симбиоз жизненно необходим для исследования нейронных сетей. Представим некоторые проблемы, решаемые искусственными нейронными сетями. Задачи, решаемые информационными системами**,** в большинстве случаев можно свести к ряду типовых [2], среди которых можно выделить следующие:

1. Классификация образов – определение принадлежности образа к одному или нескольким предварительно определенным классам.

2. Кластеризация/категоризация – разделение образов на заранее не определенные классы по каким-либо признакам.

3. Аппроксимация функций – оценку неизвестной зависимости по экспериментальным данным.

Предположим, что имеется обучающаяся выборка ((x_1, y_1**),** **(**x_2, y_2**), …,** **(**x_n, y_n**)**), которая генерируется неизвестной функцией, искаженной шумом. Задача аппроксимации состоит в нахождении оценки этой функции. 4. Прогноз (предсказание) – определение будущего процесса по его прошлому и настоящему.

Пусть заданы n дискретных отсчетов {y**(**t_1**),** y**(**t_2**), …,** y**(**t_n**)}** в последовательные моменты времени t_1, t_2, …, t_n. Задача состоит в предсказании значения y**(**t_n+1**)** в момент t_n+1 Прогноз имеет значительное влияние на принятие решений в бизнесе, науке и технике.

5. Оптимизация – нахождение решений, которые максимизируют или минимизируют определенный критерий качества при заданных ограничениях.

Многочисленные проблемы в математике, статистике, технике, науке, медицине и экономике могут рассматриваться как проблемы оптимизации. Задачей оптимизации является нахождение решения, которое удовлетворяет системе ограничений и максимизирует или минимизирует целевую функцию.

6. Память**,** адресуемая по содержанию (ассоциативную память), а также доступная по указанному содержанию.

В модели вычислений фон Неймана обращение к памяти доступно только посредством адреса, который не зависит от содержания памяти. Более того, если допущена ошибка в вычислении адреса, то может быть найдена совершенно иная информация. Память, адресуемая по содержанию, или ассоциативная память, доступна по указанию заданного содержания. Содержимое памяти может быть вызвано даже по частичному или искаженному содержанию. Ассоциативная память чрезвычайно желательна при создании перспективных информационно-вычислительных систем.

7. Управление – перевод и поддержание системы в требуемом состоянии. Рассмотрим динамическую систему, заданную совокупностью **{u(t), y(t)},** где **u(t)** является входным управляющим воздействием, а **y(t)** – выходом системы в момент времени t. В системах управления с эталонной моделью целью управления является расчет такого входного воздействия **u(t),** при котором система следует по желаемой траектории, диктуемой эталонной моделью.

Нейрон представляет собой единицу обработки информации в нейронной сети. На блок-схеме рис. 1 показана модель (model) нейрона, лежащего в основе искусственных нейронных сетей. В этой модели можно выделить три основных элемента.

1. Набор *синапсов* (synapse) или *связей* (connecting link), каждый из которых характеризуется своим весом (weight) или силой (strength). В частности, сигнал *xj* на входе синапса *j*, связанного с нейроном *k*, умножается на вес *wkj*. Важно обратить внимание на то, в каком порядке указаны индексы синаптического веса *wkj*. Первый индекс относится к рассматриваемому нейрону, а второй – к входному окончанию синапса, с которым связан данный вес. В отличие от синапсов мозга синаптический вес искусственного нейрона может иметь как положительные, так и отрицательные значения.

2. *Сумматор* (adder) складывает входные сигналы, взвешенные относительно соответствующих синапсов нейрона. Эту операцию можно описать как *линейную комбинацию*.

3. *Функция активации* (activation function) ограничивает амплитуду выходного сигнала нейрона. Эта функция также называется функцией сжатия (squashing function). Обычно нормализованный диапазон амплитуд выхода нейрона лежит в интервале [0,1] или [–1,1].

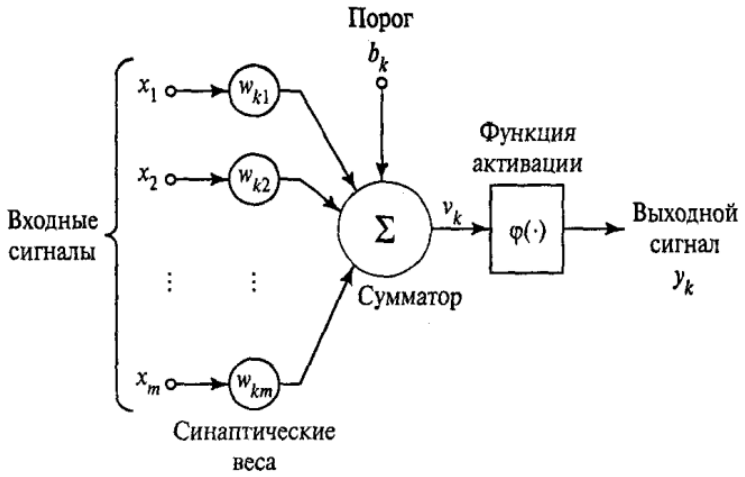


Рис. 1 Нелинейная модель нейрона [3]

Модели нейрона. В модель нейрона, показанную на рис. 1, включен пороговый элемент (bias), который обозначен символом *bk*. Эта величина отражает увеличение или уменьшение входного сигнала, подаваемого на функцию активации. В математическом представлении функционирование нейрона *k* можно описать следующей парой уравнений:

$$u_k = \sum_{j=1}^{m} w_{kj} x_j ,$$

$$y_k = \phi(u_k + b_k), \quad (1)$$

где *x1, x2,…, xm* – входные сигналы; *w1, w2,…, wkm* – синаптические веса нейрона *k; uk* – линейная комбинация входных воздействий (linear combiner output); *bk* – порог; φ(·) – функция активации (activation function); *yk* – выходной сигнал нейрона. Использование порога *bk* обеспечивает эффект аффинного преобразования (affine transformation) выхода линейного сумматора *uk*. В модели, показанной на рис. 2, постсинаптический потенциал вычисляется следующим образом:

$$\vartheta_k = u_k + b_k. \quad (2)$$

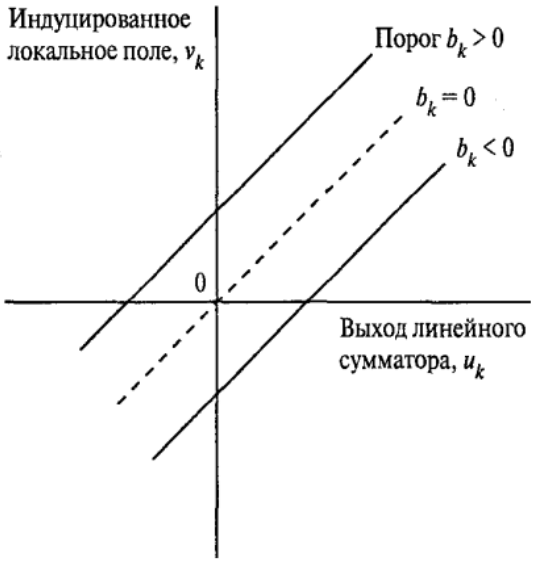


Рис. 2. Аффинное преобразование, вызванное наличием порога

Порог *bk* является внешним параметром искусственного нейрона *k*. Его присутствие мы видим в выражении (1), которое преобразуем с учётом (4)

$$\vartheta_k = \sum_{j=1}^{m} w_{kj} x_j ,$$

$$y_k = \phi(\vartheta_k). \quad (3)$$

В выражении (3) добавился новый синапс. Его входной сигнал равен: а его вес:

$$x_0=+1, \quad (4)$$

$$w_{k0}=b_k \quad (5)$$

Это позволило трансформировать модель нейрона к виду, показанному на рис. 3. На этом рисунке видно, что в результате введения порога добавляется новый входной сигнал фиксированной величины +1, а также появляется новый синаптический вес, равный пороговому значению *bk*. Хотя модели, показанные на рис.2 и рис.3, внешне совершенно не схожи, математически они эквивалентны [3].

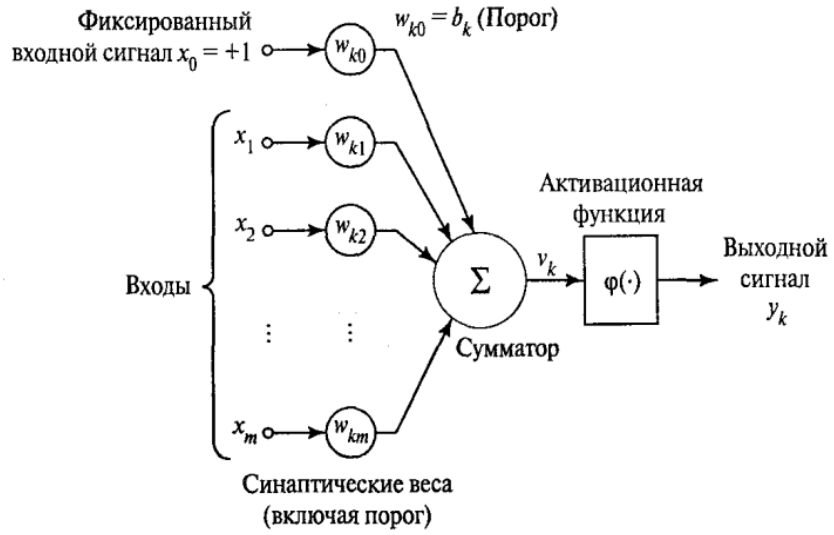
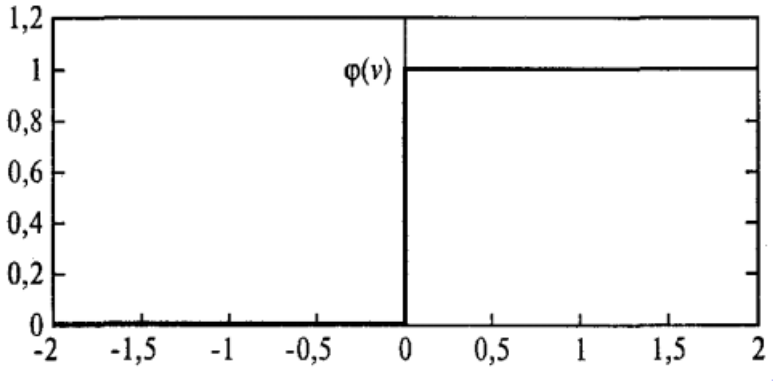


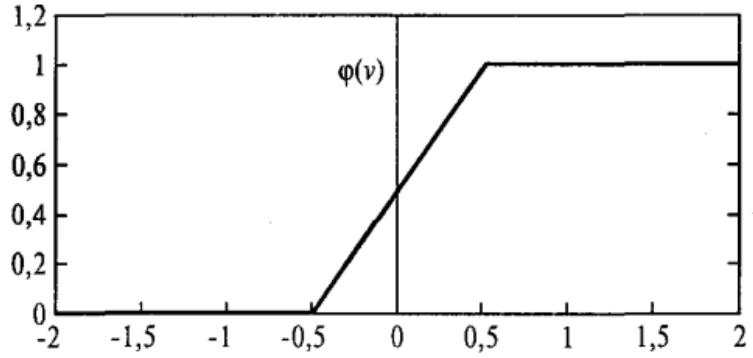
Рис. 3. Нелинейная модель нейрона

Функции активации, представленные в формулах как $\varphi(\vartheta)$, определяют выходной сигнал нейрона в зависимости от индуцированного локального поля ϑ. Можно выделить три основных типа функций активации.1. *Функция единичного скачка*, или пороговая функция (threshold function). Этот тип функции показан на рис.4, *а* и описывается следующим образом:

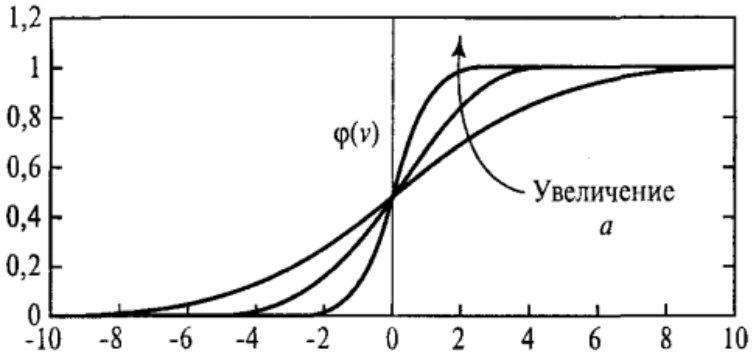
$$\varphi(\vartheta)=\begin{cases}1, \text{ если } \vartheta \geq 0\\ 0, \text{ если } \vartheta \prec 0\end{cases} \quad (6)$$



а)



б)



в)

Рис.4. Виды активационных функций [3]:
функция единичного скачка (а); кусочно-линейная функция (б) и сигмоидальная функция для различных значений параметра а (в)В технической литературе эта форма функции единичного скачка
обычно называется функцией Хэвисапда (Heaviside function). Соответственно выходной сигнал нейрона к такой функции можно представить как

$$y_k = \begin{cases} 1, \text{если } \vartheta_k \geq 0 \\ 0, \text{если } \vartheta_k \prec 0 \end{cases} \quad (7)$$

где ϑk – это индуцированное локальное поле нейрона, т. е.

$$\vartheta_k = \sum_{j=1}^{m} w_{kj} x_j + b_k \quad (8)$$

Эту модель в литературе называют моделью МакКаллока-Питца (McCalloch-Pitts model), отдавая дань работе [4]. В этой модели выходной сигнал нейрона принимает значение 1, если индуцированное локальное поле этого нейрона не отрицательно, и 0 – в противном случае. Это выражение описывает свойство «все или ничего» модели Мак-Каллока – Питца.

2. *Кусочно-линейная функция* (piecewise-linear function), показанная на рис. 4, б, описывается следующим выражением:

$$\varphi(\vartheta) = \begin{cases} 1, & \vartheta \geq +\frac{1}{2}; \\ |\vartheta|, & +\frac{1}{2} \succ \vartheta \succ -\frac{1}{2}; \\ 0, & \vartheta \leq -\frac{1}{2}, \end{cases} \quad (9)$$

где коэффициент усиления в линейной области оператора предполагается равным единице. Эту функцию активации можно рассматривать как аппроксимацию (approximation) нелинейного усилителя. Следующие два варианта можно считать особой формой кусочно-линейной функции.

ОБСУЖДЕНИЕ

Выходное значение нейрона лежит в диапазоне [0, 1]. Ценные свойства сигмоидальной функции – дифференцируемость на всей оси абсцисс и простое выражение для ее производной, что используется в некоторых алгоритмах обучения. Кроме того, она обладает свойством усиливать слабые сигналы лучше, чем большие, что предотвращает насыщение от больших сигналов, так как они соответствуют областям аргументов, где сигмоид имеет пологий наклон. Известно также большое количество других разновидностей активационных функций.

- Если линейная область оператора не достигает порога насыщения, он превращается в линейный сумматор (linear combiner).
- Если коэффициент усиления линейной области принять бесконечно большим, то кусочно-линейная функция вырождается в пороговую (threshold function).

3. *Сигмоидальная функция* (sigmoid function). Сигмоидальная функция, график которой напоминает букву S, является, пожалуй, самой распространенной функцией, используемой для создания искусственных нейронных сетей. Это быстро возрастающая функция, которая поддерживает баланс между линейным и нелинейным поведением. Примером сигмоидаль-ной функции может служить логистическая функция (logistic function), задаваемая следующим выражением:

$$\varphi(\vartheta) = \frac{1}{1 + exp(-a \cdot \vartheta)} \qquad (10)$$

где a – параметр наклона (slope parameter) сигмоидальной функции. Изменяя этот параметр, можно построить функции с различной крутизной (см. рис. 4, ***в***). Первый график соответствует величине параметра, равной ***а***/4. В пределе, когда параметр наклона достигает бесконечности, сигмоидальная функция вырождается в пороговую. Если пороговая функция может принимать только значения 0 и 1, то сигмоидальная функция принимает бесконечное множество значений в диапазоне от 0 до 1. При этом следует заметить, что сигмоидальная функция является дифференцируемой (дифференцируемость активационной функции играет важную роль в теории нейронных сетей), в то время как пороговая – нет.

Область значений функций активации, определенных формулами (6), (9) и (10), представляет собой отрезок от 0 до +1. Однако иногда требуется функция активации, имеющая область значений от –1 до +1. В этом случае функция активации должна быть симметричной относительно начала координат. Это значит, что функция активации является нечетной функцией индуцированного локального поля. В частности, пороговую функцию в данном случае можно определить следующим образом:

$$\varphi(\vartheta) = \begin{cases} 1, \text{ если } \vartheta \geq +\frac{1}{2}; \\ 0, \text{ если } \vartheta = 0; \\ -1, \quad \vartheta \prec -0, \end{cases} \qquad (11)$$

Эта функция обычно называется сигнум. В данном случае сигмоидальная функция будет иметь форму гиперболического тангенса:

$$\varphi(\vartheta) = \tan h(\vartheta) \qquad (12)$$

Такой вид сигмоидальной функции обеспечивает ряд преимуществ. Ценные свойства сигмоидальной функции – дифференцируемость на всей

оси абсцисс и простое выражение для ее производной, что используется в некоторых алгоритмах обучения. Кроме того, она обладает свойством усиливать слабые сигналы лучше, чем большие, что предотвращает насыщение от больших сигналов, так как они соответствуют областям аргументов, где сигмоид имеет пологий наклон. Известно также большое количество других разновидностей активационных функций.

Модель нейрона, показанная на рис. 3, является *детерминистской***.** Это значит, что преобразование входного сигнала в выходной точно определено для всех значений входного сигнала. Однако в некоторых приложениях лучше использовать *стохастические нейросетевые модели*, в которых функция активации имеет вероятностную интерпретацию. В подобных моделях нейрон может находиться в одном из двух состояний: +1 или –1. Решение о переключении состояния нейрона принимается с учетом вероятности этого события. Обозначим состояние нейрона символом x, а *вероятность активации нейрона* (probability of firing) – функцией $P(\vartheta)$, где ϑ – индуцированное локальное поле нейрона. Тогда

$$x = \begin{cases} +1, \text{ с вероятностью } P(\vartheta); \\ -1, \text{ с вероятностью } 1 - P(\vartheta). \end{cases} \quad (13)$$

Вероятность описывается сигмоидальной функцией следующего вида:

$$P(\vartheta) = \frac{1}{1 + \exp\left(-\frac{\vartheta}{T}\right)}, \quad (14)$$

где T – это аналог температуры (temperature), используемый для управления уровнем шума, и, таким образом, степенью неопределенности переключения. При этом важно заметить, что T не описывает физическую температуру нейронной сети, будь то биологической или искусственной. Параметр T управляет термальными флуктуациями, представляющими эффект синаптического шума. Заметим, что если параметр T стремится к нулю, то стохастический нейрон, описанный выражением (14), принимает детерминированную форму (без включения шума) нейрона Мак-Каллока – Питца.

ЗАКЛЮЧЕНИЕ

Таким образом, нейронные сети идеально приспособлены для обнаружения сложных зависимостей в отсутствие априорных знаний об исследуемой системе или процессе. Нейронные сети также можно использовать

везде, где обычно применялись линейные методы и алгоритмы, и производилось оценивание при помощи статистических методов анализа, таких как регрессионный, кластерный, дискриминантный анализ, временные ряды.

Технологии нейронных сетей, нечеткой логики и генетических алгоритмов применимы практически в любой области. В технической физике, в процессах разделения возможности использования могут быть направлены на решение комплекса задач управления динамическими объектами, например каскадом по разделению изотопов и его оптимизацией, построения адаптивных и интеллектуальных систем управления.

Главный недостаток существующих разработок – это их малая серийность из-за узкой специализации и, в связи с этим, высокая их стоимость при сравнительно небольшой сложности нейронной сети.

REFERENCES

1. Fodor J.A. and Z.W. Pylyshyn. Connectionism and cognitive architecture: a critical analysis // Cognition, 1988 – Vol. 28. – P. 3–72.
2. Hornik K., Stinchcombe M., Wite H. Multilayer Feedforward. Networks are Universal Approximators // Neural Networks, 1989. – Vol. 2. – P. 359–366.
3. Хайкин Саймон. Нейронные сети. – М.: Вильямс, 2006. – 1103 с.
4. McCulloch W.S. and W. Pitts. A logical calculus of the ideas immanent in nervous activity // Bulletin of Mathematical Biophysics, 1943. – Vol. 5. – P. 115–133.
5. Генетические алгоритмы, искусственные нейронные сети и проблемы виртуальной реальности/ Г.К. Вороновский, К.В. Махотило, С.Н. Петрашев, С.А. Сергеев. – Харьков: Основа, 1997. – 112 с.
6. Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. – Новосибирск: Наука, 1996. – 276 с.
7. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. – М.: Горячая линия-Телеком, 2001. – 382 с.
8. Круглов В.В., Дли М.И., Голунов Р.Ю. Нечеткая логика и искусственные нейронные сети. – М.: Физматлит, 2001. – 224 с.
9. Усков А.А. Интеллектуальные технологии управления. Искусственные нейронные сети и нечётная логика. – М.: Горячая линия Телеком, 2004. – 143 с.
10. Жамалова Г.Б. «Использование компьютеров в сфере обслуживании» “Интеграция мировых научных процессов как основа общественного прогресса”. Международный сборник научных трудов. Россия город Казань за июнь 2016 г. Ст. 85-88. В.№ 38.