ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ С ГИСТЕРЕЗИСНОЙ ФУНКЦИЕЙ АКТИВАЦИИ НА ОСНОВЕ ПРЕОБРАЗОВАТЕЛЯ ПРЕЙСАХА

М. Е. Семенов*, А. М. Соловьев, М. Г. Матвеев*, О. И. Канищева*****

** Воронежский государственный университет*
*** ОАО «Концерн «Созвездие»,*
**** ВУНЦ ВВС ВВА имени Н.Е. Жуковского и Ю.А. Гагарина (г. Воронеж)*

Поступила в редакцию 18.07.2013 г.

Аннотация. Разработаны принципы построения однослойной и двухслойной искусственных нейронных сетей с гистерезисной функцией активации нейронов на основе модели Прейсаха. Предложены алгоритмы их обучения. Проведен анализ работы рассматриваемых нейросетей на примере задачи классификации образов.
Ключевые слова: искусственная нейронная сеть, персептрон, многослойный персептрон, гистерезисная функция активации, гистерезис, модель Прейсаха, преобразователь Прейсаха, алгоритм обучения искусственных нейронных сетей.
Annotation. One-layered and two-layered artificial neural network with hysteresis activation function was developed. Training algorithm was proposed. Behavior of present neural network was investigated by the example of image's classification problem.
Keywords: artificial neural network, perseptron, multi-layered perseptron, hysteresis activation function, hysteresis, Preisach hysteresis model, training algorithm for artificial neural network.

ВВЕДЕНИЕ

Искусственные нейронные сети (ИНС) строятся по принципам организации и функционирования их биологических аналогов, нейронов головного мозга. Интерес к исследованию ИНС связан с тем, что способ обработки информации человеческим мозгом в корне отличается от методов, применяемых обычными цифровыми компьютерами. Мозг представляет собой чрезвычайно сложный, нелинейный, параллельный компьютер (систему обработки информации). Он обладает способностью организовывать свои структурные компоненты (нейроны) так, чтобы они могли выполнять конкретные задачи (такие как распознавание образов, обработку сигналов органов чувств, моторные функции) во много раз быстрее, чем могут позволить самые быстродействующие современные компьютеры. Примером такой задачи обработки информации является зрение [1–3].

Развитие и организация нейронов связаны с понятием пластичности мозга – способности настройки нервной системы в соответствии с окружающими условиями. Пластичность играет самую важную роль в работе нейронов в качестве единиц обработки информации в человеческом мозге. Аналогично, в ИНС работа проводится с искусственными нейронами. В общем случае ИНС представляет собой машину, моделирующую способ обработки мозгом конкретной задачи. Такая сеть обычно реализуется с помощью электронных компонентов или моделируется программой, выполняемой на цифровом компьютере.

ИНС способны решать широкий круг задач классификации образов, идентификации, прогнозирования, оптимизации, управления сложными объектами, распознавания текстов, фильтрации спама и др. Эффективность работы ИНС зависит от многих факторов, таких как архитектура сети, методы ее реализации, алгоритмы обучения, а так же свойства, заимствованные от биологических нейронов головного мозга человека. Так, заимствованная пластичность позволяет построить самообучающуюся ИНС. Еще одним не менее важным свойством биологических нейронов является кратковременная память. С точки зрения построения ИНС, дан-

Работа поддержана грантами РФФИ 11-08-00032, 12-07-00252, 13-08-00532.

ное свойство может быть заимствовано путем использования в нейронах гистерезисных функций активации (ГФА) [4–6], [15].

Так, в работе [4] рассматривается модель нейронов с бинарной ГФА для построения ИНС Хопфилда и находится теоретическое обоснование совместных свойств классической ИНС Хопфилда и модифицированной ИНС с ГФА.

Авторы работы [4] предлагают использовать в качестве функции активации нейронов сети бинарную ГФА. Однако в случае применения этой функции активации не всегда гарантируется снижение энергии системы. Было показыно, как предложенная модель позволяет предотвратить колебательные процессы в динамике сети, вследствие чего обеспечивается сходимость процесса. С этой целью были произведены модификации ГФА (бинарной и многоуровневой).

В рассматриваемой работе, с помощью компьютерного моделирования, было доказано, что модифицированная ИНС с гистерезисом имеет ту же сходимость и те же совместные свойства, что и классическая нейронная сеть Хопфилда. Авторы демонстрируют работу их модели на примере решения комбинаторной задачи оптимизации (поиска максимального разреза). Было показано, что по сравнению с прочими методами построения нейронных сетей, модифицированная ИНС с ГФА имеет лучшие временные и качественные характеристики при решении задачи поиска максимального разреза.

Работа [5] посвящена исследованию динамики ИНС с ГФА при распознавании образов (модель восприятия на основе взаимодействия в системе «сетчатка-мозг»), а также механизма воспоминания и обучения биологических нейронных систем, которые могут быть описаны с помощью такой ИНС. Авторы работы исследовали 3-нейронную ИНС с применением гладкой ГФА и провели наблюдение и анализ феномена бифуркации и хаоса в ее поведении.

В работе [6] рассматривается синхронная ИНС Хопфилда с ГФА входящих в нее нейронов. Авторы сравнивают данную модель со стандартной нейронной сетью на примере восстановления изображения на фоне шума и показывают ее преимущества. Одной из особенностей данной модели является лучшая характеристика сходимости при большом уровне шума и худшая – при малом. Также, в отличие от стандартной модели, предложенная модель имеет две области стабильности. В работе показана зависимость сходимости и шумового порога как функция от расстояния между областями стабильности на фазовой плоскости. Авторы рассматривают гистерезисную функцию нейрона на основе нейрофизиологии и производят тестирование предложенной модели на примере оптической системы с применением *ZnSe*-фильтра.

ПРЕОБРАЗОВАТЕЛЬ ПРЕЙСАХА

Особенностью рассматриваемой в данной работе ИНС является наличие гистерезиса в функции активации нейронов сети. Для построения модели такой ИНС наиболее подходящим является использование преобразователей гистерезисной природы. Следуя классическим схемам М.А. Красносельского и А.В. Покровского [8], гистерезисные операторы трактуются как преобразователи, определенные на пространстве непрерывных функций, динамика которых описывается соотношениями: вход-состояние и состояние-выход.

Обозначим через $R\left[\alpha,\beta,x_0\right]$ двухпозиционное реле с пороговыми числами α и β. Пространством состояний неидеального реле является пара чисел $\{-1,1\}$. Связь между входом $u(t)\in c_{[0,t]}$ и переменным выходом $x(t)\in\{-1,1\}$ устанавливается оператором $R\left[\alpha,\beta,x_0\right]$:

$$x(t)=R\left[\alpha,\beta,x_0\right]\cdot u(t), \qquad (1)$$

здесь x_0 – начальное состояние преобразователя. Взаимосвязь между входом и выходом иллюстрирует рис. 1.

Преобразователем Прейсаха называют континуальный аналог преобразователя, состоящего из неидеальных реле, соединенных параллельно [9–11], [16]. Рассмотрим частный класс таких преобразователей. Пусть на полуплоскости $P_{\alpha,\beta}\equiv\{\alpha,\beta:\alpha<\beta\}$ определена положитель-

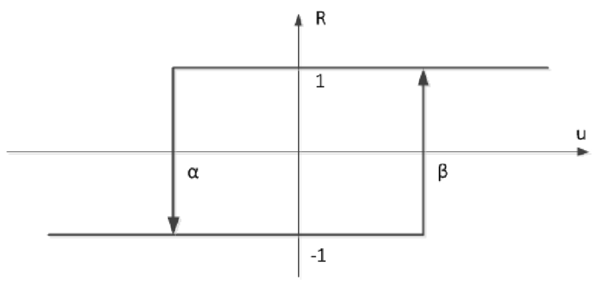

Рис. 1. Характеристика неидеального реле

ная абсолютно непрерывная суммируемая функция $\lambda = \lambda\left(\alpha, \beta\right)$. Определим на полуплоскости $P_{\alpha,\beta}$ меру μ равенством:

$$d\mu = \lambda\left(\alpha, \beta\right) d\alpha d\beta. \tag{3}$$

Отметим, что измеримыми по мере μ будут все измеримые по Лебегу множества, в том числе и имеющие бесконечную меру.

Обозначим через ψ класс ограниченных функций, заданных на неотрицательной полуоси и удовлетворяющих условию Липшица с коэффициентом, равным единице. Рассмотрим множество Ω_ψ скалярных функций $\omega(\alpha, \beta)$, заданных на полуплоскости $P_{\alpha,\beta} \equiv \left\{\alpha, \beta : \alpha < \beta\right\}$ и таких, что:

$$\omega\left(\alpha, \beta\right) = \begin{cases} -1, & \alpha + \beta > \psi\left(\beta - \alpha\right), \\ 1, & \alpha + \beta \le \psi\left(\beta - \alpha\right), \end{cases} \tag{4}$$

где $\psi\left(\nu\right) \in \psi$. Множество Ω_ψ – пространство возможных состояний преобразователя Прейсаха. На рис. 2 показан один из элементов множества Ω_ψ.

Соотношение вход – переменное состояние преобразователя Прейсаха (Γ, ξ) и состояние-выход устанавливается соотношениями.

$$\omega(\alpha, \beta, t) = \Gamma[\omega_0]u(t) =$$
$$= \mathbf{R}[\omega_0(\alpha, \beta), \alpha, \beta)]u(t),$$

$$\xi(t) = \int\limits_{\alpha<\beta} \omega(\alpha, \beta, t) d\mu_{\alpha,\beta} =$$
$$= \mu_{\alpha,\beta}(\{\alpha, \beta\} : \mathbf{R}[\omega_0(\alpha, \beta), \alpha, \beta]u(t) = 1).$$

В дальнейшем будем использовать преобразователь Прейсаха аппроксимируемый конечным числом операторов $R\left[\alpha, \beta, x_0\right]$ в качестве функции активации исследуемой ИНС. Блок-схема такого преобразователя приведена на рис. 3.

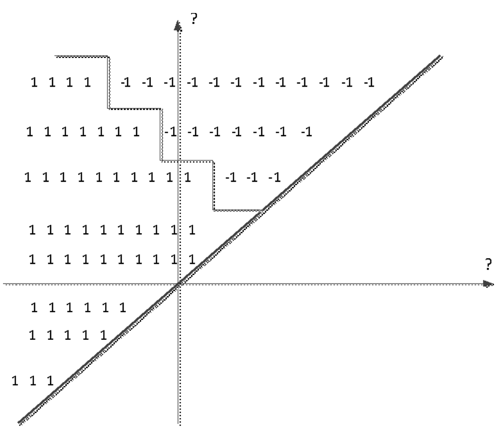

Рис. 2. Элемент множества Ω_ψ

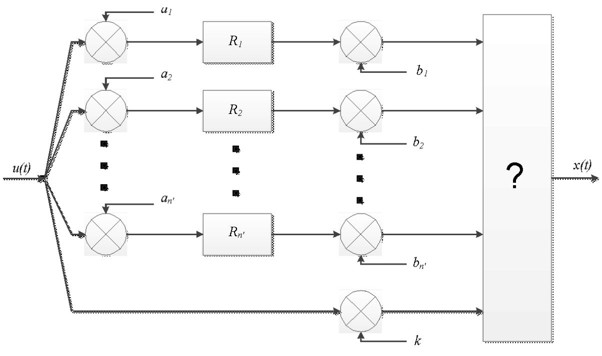

Рис. 3. Блок-схема преобразователя Прейсаха

Таким образом, функция активации каждого нейрона сети содержит n неидеальных реле $R\left[\alpha, \beta, x_0\right]$, имеющих характеристику, показанную на рис. 1. Дополнительные коэффициенты $a_{n'}$, $b_{n'}$ и k задают наклон характеристик неидеальных реле и, как следствие, форму характеристики преобразователя Прейсаха в целом.

Будем считать, что коэффициенты $a_{n'} = \left\{1\right\}$, $b_{n'} = \left\{1\right\}$, $k = 0{,}1$ а шаг изменения пороговых значений α и β, $h_{\alpha,\beta} = 0{,}1$. В этом случае характеристика преобразователя будет иметь вид, показанный на рис. 4. Наклон характеристики при значениях превосходящих по модулю единицу обеспечит лучшую сходимость при использовании градиентного метода поиска оптимального решения в процессе обучения ИНС.

Из графиков, приведенных на рис. 2 и рис. 4 следует, что количество неидеальных реле, удовлетворяющих выбранным условиям можно рассчитать следующим образом:

$$n' = \frac{\left(d - c + h_{\alpha,\beta}\right) \cdot \left(d - c + 2h_{\alpha,\beta}\right)}{2h^2}. \tag{5}$$

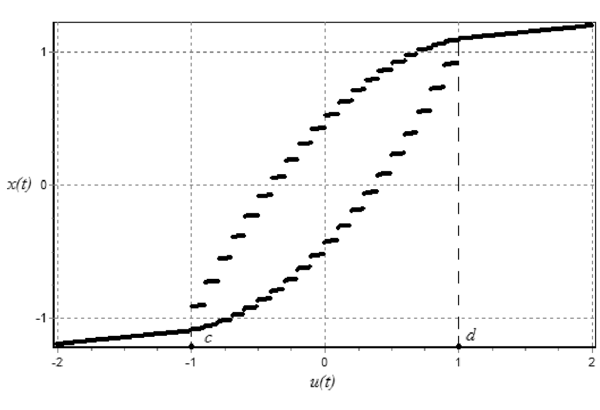

Рис. 4. Характеристика преобразователя Прейсаха

Таким образом, в рассматриваемом случае $n' = 231$.

ОДНОСЛОЙНАЯ ИНС С ГИСТЕРЕЗИСОМ

Рассмотрим однослойную ИНС с гистерезисной функцией активации на основе преобразователя Прейсаха. Архитектура ИНС приведена на рис. 5.

Здесь Y^m – вектор входных значений ;W^{mn} – матрица весовых коэффициентов; U^n – вектор воздействий; X^n – целевой вектор; $G = f(U, X, R_{n'})$ – функция активации.

В рассматриваемой нейронной сети будем считать, что $m = 2688$, $n = 10$, $n' = 231$, входной вектор Y имеет область значений {-1,1}, а вектор воздействий U^n определяется следующим образом:

$$U^n = Y^m \cdot W^{mn}, \tag{6}$$

$$u_k = \sum_{i=1}^{m} y_i \cdot w_{ik} \ \left(k = 1, 2, \ldots, n\right) \tag{7}$$

В качестве функции активации G используем описанную выше ГФА.

Алгоритм обучения:

Так как рассматриваемая ИНС является однослойной, для её обучения воспользуемся дельта-правилом, являющимся следствием первого и второго правил Хебба.

Данный метод является разновидностью методов обучения с учителем, то есть должен существовать набор векторов (Y^k,T^k), $k = \overline{1,p}$, называемый обучающей выборкой. Здесь $Y^k = (y_1^{\,k}, y_2^{\,k}, \ldots, y_m^{\,k})$, $k = \overline{1,p}$ – примеры входных образов, для которых заранее известна их принадлежность к одному из 10-ти классов (числа от 0 до 9) и соответствующий им набор целевых векторов $T^k = (t^k_{\ 1}, t^k_{\ 2}, \ldots, t^k_{\ n})$, $k = \overline{1,p}$.

Суть метода заключается в итерационной подстройке весовых коэффициентов W, последовательно уменьшающей выходную ошибку. Алгоритм обучения включает несколько шагов:

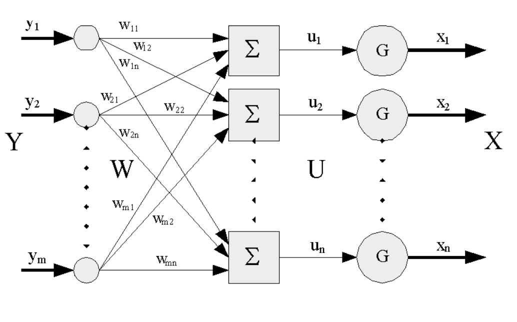

Рис. 5. Однослойная ИНС

ДВУХСЛОЙНАЯ ИНС С ГИСТЕРЕЗИСОМ

Рассмотрим двухслойную ИНС с гистерезисной функцией активации. Архитектура сети приведена на рис. 6.

Принцип построения ИНС оставим тот же что и в предыдущем пункте. Добавим еще один слой нейронов с ГФА, при этом слой 1 будем называть скрытым, а слой 2 – выходным слоем.

Обучение ИНС будем проводить с помощью процедуры обратного распространения ошибки. Таким образом, рассматриваемая ИНС можно классифицировать как сеть обратного распространения.

Как видно из рис. 6: Y^m – вектор входных значений; W^{mn} – матрица весовых коэффициентов скрытого слоя; U^n – вектор воздействий скрытого слоя; Q^n – выходной вектор скрытого слоя; V^{np} – матрица весовых коэффициентов выходного слоя; E^p – вектор воздействий выходного слоя; X^p – целевой вектор.

Количество нейронов n в скрытом слое должно быть 30–50 % от числа m. Это число определяется экспериментальным путем. При слишком большом n может наблюдаться эффект переобучения сети, что отрицательно повлияет на способность сети к обобщению. При слишком малом n ИНС может потерять способность к обучению.

Как было показано выше, для решения задачи классификации чисел 0–9, будем считать, что $m = 2688$. Выберем $n = 500$, $n' = 231$, $p = 10$. Входной вектор Y имеет область значений {-1,1}.

Для индексов примем следующие обозначения: входы нумеруются индексом $i = \overline{1,m}$, элементы скрытого слоя индексом $j = \overline{1,n}$, а выхо-

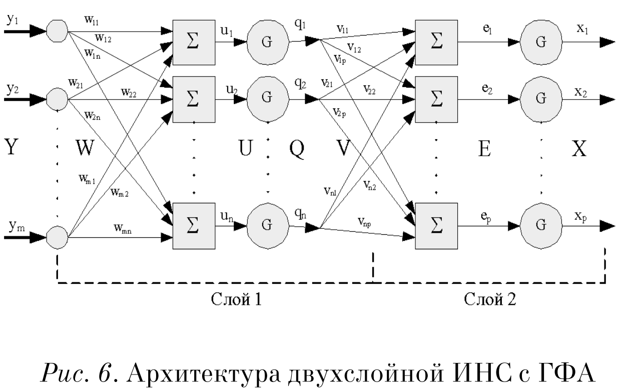

Рис. 6. Архитектура двухслойной ИНС с ГФА

ды соответственно индексом $k = \overline{1, p}$. Индексом l будем обозначать номер итерации работы ИНС.

Прямой проход через сеть осуществляется следующим образом:

$$U^n = Y^m \cdot W^{mn}, \tag{8}$$

$$Q^n = G\left(U^n\right), \tag{9}$$

$$E^p = Q^n \cdot V^{np}, \tag{10}$$

$$X^p = G\left(E^p\right). \tag{11}$$

Функция активации G имеет ту же реализацию, что и в предыдущем пункте.

Алгоритм обучения:

Рассматриваемая ИНС является сетью обратного распространения. Для обучения такой нейронной сети удобно использовать алгоритм минимизации целевой функции ошибки, которая находится по формуле:

$$E\left(W, V\right) = \frac{1}{2} \cdot \sum_{k=1}^{p} \left(x_k - d_k\right)^2, \tag{12}$$

где x_k – полученное реальное значение k-го выхода нейросети при подаче на нее одного из входных образов обучающей выборки (Y^t, D^t), $t = \overline{1, T}$;

d_k – требуемое (целевое) значение k-го выхода для этого образа.

Обучение нейросети проводилось методом градиентного спуска, то есть на каждой итерации веса перенастраиваются по формулам:

$$w_{ij}^{\ N+1} = w_{ij}^{\ N} - \sigma \cdot \frac{\partial E}{\partial w_{ij}}, \tag{13}$$

$$v_{jk}^{\ N+1} = v_{jk}^{\ N} - \sigma \cdot \frac{\partial E}{\partial v_{jk}}, \tag{14}$$

где σ – коэффициент, определяемый в процессе обучения.

Для метода градиентного спуска необходимо, чтобы функция активации была дифференцируемой на всей области допустимых значений. В нашем случае это не так, однако, зная в каждый момент времени величины $x_k^{(l+1)}$ и $x_k^{(l)}$, $q_j^{(l+1)}$ и $q_j^{(l)}$ легко вычислить производную активационной функции нейронов скрытого слоя $\frac{\partial q}{\partial u}$ и выходного слоя $\frac{\partial x}{\partial e}$ численным методом.

Функция ошибки в явном виде не содержит зависимости от весовых коэффициентов w_{ij} и v_{jk}, поэтому для вычисления соответствующих производных воспользуемся формулами:

$$\frac{\partial E}{\partial v_{jk}} = \frac{\partial E}{\partial x_k} \cdot \frac{\partial x_k}{\partial e_k} \cdot \frac{\partial e_k}{\partial v_{jk}}. \tag{15}$$

Исходя из архитектуры сети известно, что

$$e_k = \sum_{j=1}^{n} v_{jk} \cdot q_j, \tag{16}$$

следовательно

$$\frac{\partial e_k}{\partial v_{jk}} = q_j. \tag{17}$$

Значение производной функции активации $\frac{\partial x_k}{\partial e_k}$ вычисляется в процессе работы сети численным методом, а производная функции ошибки находится следующим образом:

$$\frac{\partial E}{\partial x_k} = x_k - d_k. \tag{18}$$

Таким образом, получим выражение соотношение:

$$\frac{\partial E}{\partial v_{jk}} = \left(x_k - d_k\right) \cdot q_j \cdot \frac{x_k^{(l+1)} - x_k^{(l)}}{e_k^{(l+1)} - e_k^{(l)}}. \tag{19}$$

Производную $\frac{\partial E}{\partial w_{ij}}$ находим аналогично:

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial q_j} \cdot \frac{\partial q_j}{\partial u_j} \cdot \frac{\partial u_j}{\partial w_{ij}}. \tag{20}$$

Здесь

$$\frac{\partial u_j}{\partial w_{ij}} = y_i. \tag{21}$$

Так как функция ошибки не зависит в явном виде от выходов скрытого слоя q_j, производная $\frac{\partial E}{\partial q_j}$ определяется соотношением

$$\frac{\partial E}{\partial q_j} = \sum_{k=1}^{p} \frac{\partial E}{\partial x_k} \cdot \frac{\partial x_k}{\partial e_k} \cdot \frac{\partial e_k}{\partial q_j}. \tag{22}$$

Используя выражения, полученные ранее, имеем:

$$\begin{aligned} \frac{\partial E}{\partial q_j} = \sum_{k=1}^{p} \left(x_k - d_k\right) \times \\ \times v_{jk} \cdot \frac{x_k^{(l+1)} - x_k^{(l)}}{e_k^{(l+1)} - e_k^{(l)}}. \end{aligned} \tag{23}$$

Введем обозначение

$$\delta_k = \left(x_k - d_k\right) \cdot \frac{x_k^{(l+1)} - x_k^{(l)}}{e_k^{(l+1)} - e_k^{(l)}}. \tag{24}$$

Тогда имеет место соотношение:

$$\frac{\partial E}{\partial v_{jk}} = \delta_k \cdot q_j, \qquad (25)$$

$$\frac{\partial E}{\partial w_{ij}} = y_i \cdot \frac{q_j^{(l+1)} - q_j^{(l)}}{u_j^{(l+1)} - u_j^{(l)}} \cdot \sum_{k=1}^{p} \left(\delta_k \cdot v_{jk}\right). \qquad (26)$$

Опишем теперь пошагово полный алгоритм обучения данной ИНС:

Шаг 1. Инициализация сети.

Инициализируем матрицы весовых коэффициентов случайными значениями. Задаем параметр точности обучения ε, параметр скорости обучения σ, создаем набор векторов обучающей выборки (Y^t, D^t).

Шаг 2. Прямой проход через сеть.

На вход сети подаем один из образов обучающей выборки, и определяем значения выходов всех нейронов ИНС.

Шаг 3. Настройка синоптических весов.

Корректируем весовые коэффициенты выходного слоя в соответствие с формулами:

$$v_{jk}^{N+1} = v_{jk}^{N} - \sigma \cdot \frac{\partial E}{\partial v_{jk}},$$

где $\frac{\partial E}{\partial v_{jk}} = \delta_k \cdot q_j$, $\delta_k = \left(x_k - d_k\right) \cdot \frac{x_k^{(l+1)} - x_k^{(l)}}{e_k^{(l+1)} - e_k^{(l)}}$.

Корректируем весовые коэффициенты скрытого слоя в соответствие с формулами:

$$w_{ij}^{N+1} = w_{ij}^{N} - \sigma \cdot \frac{\partial E}{\partial w_{ij}},$$

где $\frac{\partial E}{\partial w_{ij}} = y_i \cdot \frac{q_j^{(l+1)} - q_j^{(l)}}{u_j^{(l+1)} - u_j^{(l)}} \cdot \sum_{k=1}^{p} \left(\delta_k \cdot v_{jk}\right).$

Шаг 4. Шаги 2–3 повторяются для всех обучающих векторов. Обучение завершается по достижении для каждого из обучающих образов значения функции ошибки, удовлетворяющего условию $E(W,V) < \varepsilon$.

ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ

Рассмотрим работу ИНС с ГФА на примере решения задачи классификации чисел 0 – 9. В данном случае имеем 10 классов образов, поэтому в качестве целевого выходного вектора будем использовать вектор, состоящий из 10 элементов. Входными данными являются образы размером 48x56 пикселей. Таким образом, входной вектор будет состоять из 2688 элементов (по одному на каждый пиксель), при этом значение +1 соответствует черному пикселю, а -1 – белому.

ОДНОСЛОЙНАЯ ИНС С ГИСТЕРЕЗИСОМ

Перед обучением созданной ИНС, зададим начальные параметры:

Наборы входных образов Y^k должны иметь значения {-1,1}, при этом -1 соответствует черному пикселю, а 1 – белому. Каждый k-й образ содержит 2688 элементов. Количество образов обучающей выборки $k = 10$ (числа от 0 до 9).

Набор целевых векторов T^k, соответствующих входным образам Y^k, должны иметь значения {-1,2;1,2}, где -1,2 соответствует выводу «нет», а 1,2 – «да». Зададим скорость обучения $\sigma = 0{,}1$, значение допустимой ошибки $\varepsilon = 0{,}01$, начальные условия (u_k^0, x_k^0) значениями (-2;-1,2). Проинициализируем матрицу весовых коэффициентов W случайными значениями из диапазона [-0,05;0,05]. При заданных параметрах, процесс обучения созданной ИНС прошел за 240 общего количества итераций и 19 эпох.

Для дальнейшей работы данной ИНС зададим пороги принятия решений {-1,1;1,1}. Это значит, что по индексу элемента выходного вектора $x_i > 1{,}1$ можно определить нейрон, находящийся в активном состоянии («да»), $x_i < -1{,}1$ – в пассивном состоянии («нет»).

Рассмотренная ИНС с ГФА обладает большей помехоустойчивостью по сравнению с подобной ИНС построенной по стандартной схеме и может применяться в системах классификации образов с большим уровнем шумов и кратковременными помехами (например, при обработке видеопотока).

Также, как показало моделирование, данная ИНС обладает большей скоростью сходимости в процессе обучения по сравнению с подобной ИНС, ГФА которой построена на основе входо-выходных соответствий, формализуемых дифференциальным уравнением.

ДВУХСЛОЙНАЯ ИНС С ГИСТЕРЕЗИСОМ

Перед обучением созданной ИНС, зададим начальные параметры:

Наборы входных образов Y^k должны иметь значения {-1,1}, где -1 соответствует черному пикселю, 1 – белому. Каждый k-й образ содержит 2688 элементов. Количество образов обучающей выборки $k = 10$.

Набор целевых векторов D^t, соответствующих входным образам Y^t, должны иметь значения {-1,2;1,2}, где -1,2 соответствует выводу «нет», 1,2 – «да». Инициализируем матрицы W и V случайными значениями из диапазона (-0,1;0,1). Зададим параметры $\varepsilon = 0{,}1$, $\sigma = 0{,}01$. Зададим начальные условия (u_k^0, x_k^0) значениями (-2;-1,2). При заданных параметрах, процесс обучения созданной ИНС прошел за 156 общего количества итераций и 12 эпох. Для дальнейшей работы данной ИНС зададим пороги принятия решений {-1,1;1,1}. Это значит, что по индексу элемента выходного вектора $x_i > 1{,}1$ можно определить нейрон, находящийся в активном состоянии, $x_i < -1{,}1$ – в пассивном.

Как видно из результатов моделирования, двухслойная сеть обучилась за меньшее число итераций, чем сеть с одним слоем. Это можно объяснить тем, что для обучения двухслойной сети был использован более быстрый метод обучения (метод градиентного спуска). Двухслойная нейронная сеть с гистерезисом обладает повышенной стойкостью к шумам и кратковременным помехам при решении задачи классификации образов.

Моделирование двухслойной ИНС с гистерезисом показало, что при обучении сети с помощью одинаково повторяющейся последовательности образов, сеть способна запоминать эту последовательность и в процессе работы реагировать на нее активнее (с большим весом), чем на те же образы, поданные на вход в случайном порядке. Таким образом, данная ИНС, в отличие от однослойной сети, обладает не только способностью к распознаванию (классификации) образов, но и способностью к выделению из набора подаваемых на нее образов, искомой последовательности. Другими словами, двухслойная сеть обладает свойством ассоциативной памяти.

ЗАКЛЮЧЕНИЕ

Были предложены принципы построения однослойной и двухслойной искусственных нейронных сетей с гистерезисной функцией активации нейронов на основе модели Прейсаха, а также алгоритмы их обучения. Проведен анализ и сравнение работы рассматриваемых искусственных нейронных сетей на примере задачи классификации образов (чисел от 0 до 9). Как показало моделирование, нейросети с гистерезисом обладают рядом преимуществ по сравнению с классическими искусственными нейтронными сетями.

СПИСОК ЛИТЕРАТУРЫ

1. *Churchland P.S* and *T.J. Sejnowski*. The Computational Brain, Cambridge, MA: MIT Press, 1992.

2. *Levine M.* Man and Machine Vision. New York: McGraw-Hill, 1985.

3. *Marr D.* Vision, New York: W.H. Freeman and Company, 1982.

4. *Guangpu Xia*, *Zheng Tang* and *Yong Li*. Hopfield Neural Network with Hysteresis for Maximum Cut Problem. Neural Information Processing – Letters and Reviews. Vol. 4, No. 2, August 2004.

5. *Chunguang Li, Juebang Yu, Xiaofeng Liao*. Chaos in a three-neuron hysteresis Hopfield-type neural network. Elsevier. 9 July 2001.

6. *Lipo Wang, John Ross*. Synchronous neural networks of nonlinear threshold elements with hysteresis. Neurobiology. Vol. 87, pp. 988–992, February 1990.

7. *K. Jin'no*, IEICE Trans. Fund. 79 (1996) 402.

8. *Красносельский М.А.* Системы с гистерезисом / М.А. Красносельский, А.В. Покровский. М.: Наука, 1983, – 272 с.

9. *M. Kuczmann*. Dynamic Preisach hysteresis model. Journal of Advanced Research in Physics 1(1), 011003 (2010).

10. *R. Venkataraman Iyer, Xiaobo Tan, P. S. Krishnaprasad*. Approximate Inversion of the Preisach Hysteresis Operator With Application to Control of Smart Actuators. IEEE Transactions On Automatic Control, Vol. 50, No. 6, June 2005.

11. *Isaak D. Mayergoyz*. Mathematical Models Of Hysteresis. IEEE Transactions On Magnetics, Vol, Mag-22, No. 5, September 1986.

12. *С. Хайкин*. Нейронные сети: полный курс 2-е изд. М.: «Вильямс», 2006 г., – 1104 с.

13. *Р. Калан*. Основные концепции нейронных сетей. М.: «Вильямс», 2001 г., – 291 с.

14. *Круглов В.В., Борисов В.В.* Искусственные нейронные сети. Теория и практика. 2-е изд. М.: Горячая линия – телеком, 2002 г., – 382 с.

15. *Соловьев А.М., Семенов М.Е., Мишин М.Ю., Кабулова Е.Г.* Искусственные нейронные сети с гистерезисной функцией активации. Журнал “Теория и техника радиосвязи” 2013 г. № 2, С. 52–59.

16. *Семенов М.Е., Канищева О.И., Гулин А.С., Прохоров Д.М.* Корректные периодические режимы в системах управления с монотонными гистерезисными нелинейностями. Наукоемкие технологии. – 2010, № 12, С. 67–72.

Семенов М. Е. – доктор физико-математических наук, профессор кафедры цифровых

Semenov M. E. – doctor of physical and mathematical sciences, professor of digital techтехнологий Воронежского государственного университета,e-mail: mkl150@mail.ru

Соловьев А. М. – начальник сектора, e-mail: amsolov82@gmail.com, ОАО «Концерн «Созвездие»

Матвеев М. Г. – доктор технических наук, профессор, заведующий кафедры информационных технологий управления Воронежского государственного университета, e-mail: mgmatveev@yandex.ru

Канищева О. И. – кандидат физико-математических наук, доцент 206 кафедры математики ВУНЦ ВВС «ВВА» (г. Воронеж), e-mail: oleka_olesya@mail.ru

nology, Voronezh State University, e-mail: mkl150@mail.ru

Solovyev A. M. – manager of sector, e-mail: amsolov82@gmail.com, JSC «Sozvezdie» Concern»

Matveev M. G. – doctor of technical Sciences, Professor, head of Department of information technologies of management, Voronezh state University, e-mail: mgmatveev@yandex.ru

Kanishcheva O. I. – p.h.d., 206-th department of mathematics, Military Educational Scientific Center of Air Forces (Voronezh), e-mail: oleka_olesya@mail.ru