

**МИНИСТЕРСТВО ВЫСШЕГО И СРЕДНЕГО СПЕЦИАЛЬНОГО
ОБРАЗОВАНИЯ РЕСПУБЛИКИ УЗБЕКИСТАН**

ТАШКЕНТСКИЙ ХИМИКО-ТЕХНОЛОГИЧЕСКИЙ ИНСТИТУТ

кафедра “Информатика, автоматизация и управления”

МЕТОДИЧЕСКОЕ ПОСОБИЕ

**к лабораторной работе
«Аппроксимация функций с помощью
нейронных сетей»**

**по курсу «Моделирования и оптимизация
технологических процессов»**

Ташкент – 2012

**МИНИСТЕРСТВО ВЫСШЕГО И СРЕДНЕГО СПЕЦИАЛЬНОГО
ОБРАЗОВАНИЯ РЕСПУБЛИКИ УЗБЕКИСТАН**

ТАШКЕНТСКИЙ ХИМИКО-ТЕХНОЛОГИЧЕСКИЙ ИНСТИТУТ

кафедра “Информатика, автоматизация и управления”

МЕТОДИЧЕСКОЕ ПОСОБИЕ

**к лабораторной работе
«Аппроксимация функций с помощью
нейронных сетей»**

**по курсу «Моделирования и оптимизация
технологических процессов»**
для магистрантов по специальности 5А521802
«Автоматизация технологических процессов и производств» (по
отраслям)

Ташкент – 2012

Методическое пособие предназначено по курсу «Моделирования и оптимизация технологических процессов» для изучения принципов функционирования нейронных сетей и их обучения по методу обратного распространения. Пособие содержит описания структуры нейронных сетей, алгоритма обучения, задание на лабораторную работу и порядок ее выполнения.

ЛАБОРАТОРНАЯ РАБОТА №1

АППРОКСИМАЦИЯ ФУНКЦИЙ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

1. ЦЕЛЬ РАБОТЫ

Целью работы является изучение принципов функционирования нейронных сетей (НС) в рабочем режиме и режиме обучения.

2. КРАТКИЕ ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

2.1. Свойства и назначение нейронных сетей

Под методами **искусственного интеллекта** понимают все методы и математические модели, которые имитируют мыслительную (интеллектуальную) деятельность человека. В этом отношении можно выделить два подхода:

- моделирование **процесса** (алгоритма) мышления при принятии какого-либо решения,
- моделирование **работы мозга** на основе изучения биологических принципов его функционирования.

Первый подход реализуется в виде нечетких логических функций (Fuzzy Logic), экспертных систем, естественно-языковых систем. Вторым подходом представлен нейронными сетями, которые фактически являются математическими моделями мозга.

Основные свойства НС:

- массивно-параллельная обработка,
- высокая отказоустойчивость,
- использование неалгоритмических вычислений,
- способность к обобщению и классификации данных,
- способность к обучению с учителем или без учителя,
- возможность использования слабоформализованной информации.

НС получили распространение благодаря тому, что они:

- дают стандартный способ решения многих нестандартных задач;
- обладают способностью к обучению - необходимо только формировать учебные задачки, то есть труд программиста замещается трудом учителя;
- особенно эффективны там, где нужен аналог человеческой интуиции;
- позволяют создавать эффективное программное обеспечение для компьютеров с высокой степенью параллельной обработки.

В литературе [1] встречается значительное число признаков, которыми должна обладать задача, чтобы применение НС было оправдано и НС могла бы ее решить:

- отсутствует алгоритм или не известны принципы решения задач, но накоплено достаточное число примеров;
- проблема характеризуется большими объемами входной информации;
- данные неполны или избыточны, зашумлены, частично противоречивы.

Таким образом, НС хорошо подходят для распознавания образов и решения задач классификации, оптимизации и прогнозирования. В нефтяной и химической

промышленности НС могут применяться для анализа геологической информации, идентификация неисправностей оборудования, разведки залежей минералов по данным аэрофотосъемок, анализа составов примесей, при управлении процессами. Кроме того, с помощью НС может быть реализовано управление манипуляторами, управление качеством, обнаружение неисправностей, адаптивная робототехника, управление голосом и т.д.

2.2. Принцип функционирования нейронных сетей

Нейронная сеть представляет собой совокупность нейронов, связанных между собой соответствующим образом.

Нейрон – преобразовательный элемент, имеющий некоторое количество входов (синапсов), на которые поступают входные сигналы x_i и один выход (аксон), с которого снимается выходной сигнал y . Каждый синапс имеет вес w_i , на который умножается входной сигнал x_i .

Структура нейрона представлена на рис. 1. Внутри нейрона можно выделить блок суммирования, определяющий взвешенную сумму всех входных сигналов

$$U = \sum_{i=1}^n w_i \cdot x_i$$

и блок функции активации $Y = F(U)$. Таким образом, нейрон функционирует за два такта: 1) суммирование входных сигналов и 2) вычисление Y по функции активации.

Функция активации должна удовлетворять двум условиям: 1) $|F(U)| < 1$ при любом U , 2) функция должна быть монотонной неубывающей.

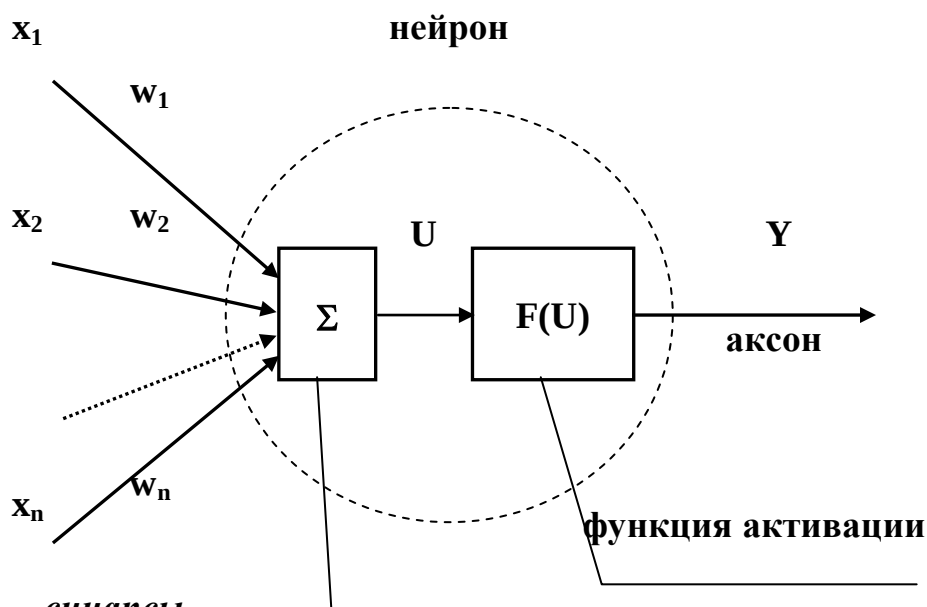


Рис.1. Структура нейрона

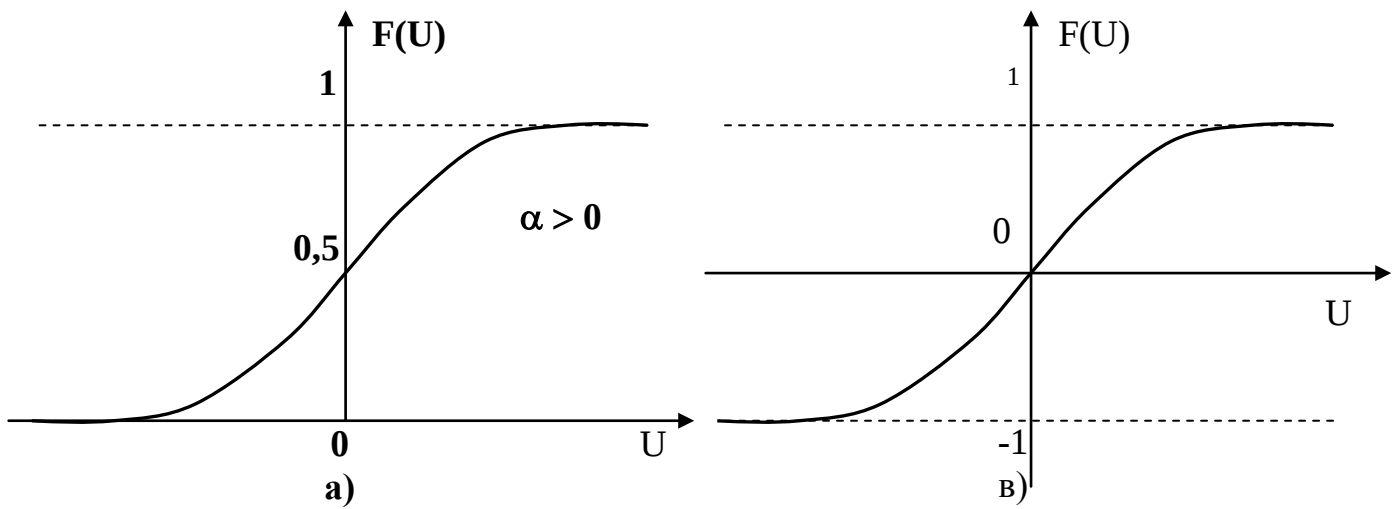


Рис. 2.

Наиболее часто в качестве функций активации используются следующие функции.

1) Ступенчатая функция

$$F(U) = \begin{cases} 1, & \text{если } U \geq \alpha \\ 0, & \text{если } U < \alpha ; \end{cases}$$

2) Сигмоидная функция

$$F(U) = \frac{1}{1 + e^{-\alpha U}} \quad (\text{см. рис. 2,а});$$

3) Гиперболический тангенс $F(U) = \text{th}(U)$ (см. рис. 2,б);

4) Гладкие сжимающие функции

$$F(U) = \frac{U + Q}{|U + Q| + \alpha}$$

где Q – порог (смещение), α - параметр, определяющий крутизну статической характеристики нейрона.

Нейроны образуют нейронные сети путем соединения синапсов с аксонами.

Наиболее распространенными и хорошо изученными являются трехслойные НС, состоящие из трех слоев нейронов: входного, скрытого и выходного (см. рис. 3). Нейроны входного слоя имеют только по одному синапсу. Количество нейронов входного слоя соответствует количеству входных переменных сети X . Задачей нейронов этого слоя является только распределение входных сигналов по нейронам скрытого слоя, суммирования и вычисления функции активации в них не происходит. Количество нейронов в скрытом слое может быть различным и часто подбирается экспериментально. Недостаточное или избыточное количество нейронов в скрытом слое приводит к ухудшению точности аппроксимации. Кроме того, избыточное количество усложняет сеть и уменьшает быстродействие. Нейроны выходного слоя формируют выходные сигналы, их количество соответствует количеству выходов Y . Пример НС с 3 входными, 4 скрытыми и 2 выходными нейронами приведен на рисунке 3. Такая НС для краткости обозначается как (3-4-2). H_{ij} – нейроны.

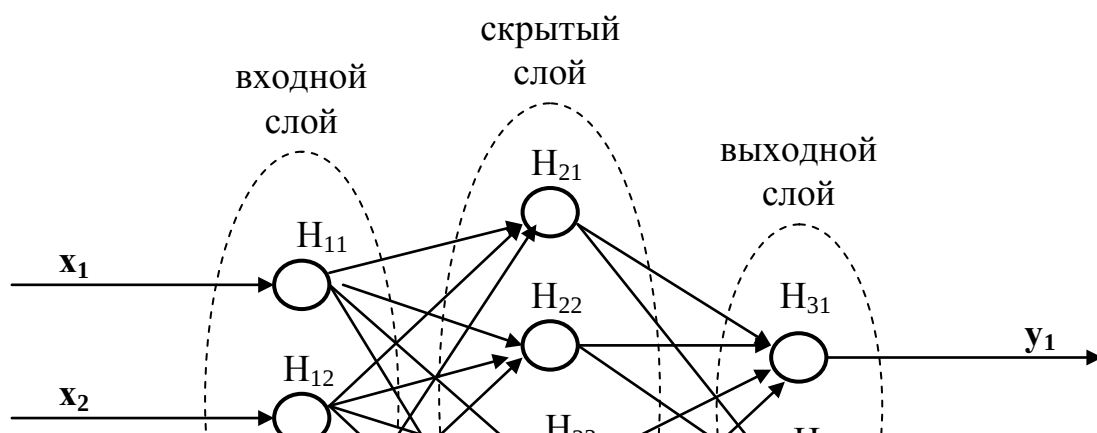


Рис. 3. Нейронная сеть вида (3-4-2)

Данные сети относятся к сетям **прямого распространения**, поскольку в них входные сигналы последовательно проходят через все нейроны и после преобразований напрямую подаются на выходы. Выходной сигнал y_{ij} каждого j -го нейрона в i -м слое определяется как

$$y_{ij} = F \left(\sum_{k=1}^{n(i-1)} w_{ij}^k \cdot y_{i-1,k} \right)$$

где $n(i)$ – число нейронов в i -м слое.

Выходные сигналы НС соответствуют $y_{N_{out}j}$, где N_{out} – число нейронов выходного слоя.

2.3. Методы обучения нейронных сетей

Нейронные сети относятся к классу аппроксиматоров и «черных ящиков», аппроксимирующих некоторые функции вида $Y = F(X)$, где Y – вектор выходных переменных, X – вектор входных [5, 6, 9].

Процесс аппроксимации заключается в подборе весовых коэффициентов w_{ij} и называется **обучением** НС. То есть НС может функционировать в двух режимах:

- эксплуатации, когда на вход подаются сигналы, а на выходе снимаются результаты вычислений;
- обучения, когда происходит корректировка весов таким образом, чтобы выходные сигналы наиболее точно соответствовали желаемым.

От качества обучения НС зависит точность ее работы в режиме эксплуатации.

Структура процесса обучения представлена на рисунке 4, где обозначены: $Y_{\text{жел}}$ – желаемые значения выходных сигналов, E – ошибка обучения ($E = Y_{\text{жел}} - Y$), K – корректирующие воздействия (обычно изменения весов Δw_{ij}).



Рис.4. Процесс обучения НС

Для обучения НС составляется **обучающая выборка** входных сигналов и соответствующих им выходных. Выборка может быть разделена на две части: **рабочую выборку** (на основе которой производится собственно обучение) и **тестирующую выборку** (для проверки качества обучения).

Далее определяется структура НС. Для трехслойной НС количества входных и выходных нейронов определяются по количествам входных и выходных переменных. Количество нейронов в скрытом слое N_c может быть взято из условия:

$$N_c \leq \frac{(N_p - 1) \cdot N_{out}}{N_{in} + N_{out} + 1},$$

где N_{in} и N_{out} – количества нейронов во входном и выходном слоях, N_p – количество обучающих примеров (объем выборки).

Весам синапсов необученной НС изначально присваиваются произвольные значения. Далее на вход НС подается первый вектор X из рабочей выборки, определяется вектор Y и ошибка обучения E . Исходя из значений вектора E корректируются веса синапсов. Затем подается следующий вектор X из выборки и т.д. Циклы обучения повторяются многократно, пока качество обучения не станет удовлетворительным (это можно проверить по тестирующей выборке).

Существует несколько методов обучения, которые можно классифицировать по способам использования учителя:

- обучение с учителем (коррекция весов производится исходя из сравнения текущего и желаемого выходных векторов);
- обучение с последовательным подкреплением знаний (сети не даются желаемые значения выходов, а ставится оценка «хорошо» или «плохо»);

- обучение без учителя (сеть сама вырабатывает правила обучения путем выделения особенностей из набора входных данных).

По использованию элементов случайности методы обучения подразделяются на:

- детерминистские (коррекция на основе анализа входных и выходных сигналов, а также дополнительной информации, например, желаемых выходов);
- стохастические (случайное изменение весов в ходе обучения – Больцмановское обучение).

К детерминистским правилам обучения относятся правило Хебба, дельта-правило, правило Кохонена, ART-правило, правило обратного распространения [5].

Наиболее распространенным правилом для сетей прямого распространения является правило обратного распространения (back propagation).

2.4. Правило обратного распространения

Для обучения обычно используется НС с функциями активации сигмоидного типа. Целью обучения по правилу обратного распространения является минимизация ошибки обучения, которая определяется как

$$E = \frac{1}{2} \sum_{i=1}^{N_{out}} (Y_i - Y_{жел.i})^2$$

Для уменьшения ошибки веса изменяются по правилу

$$W_{ij} = W_{ij} - v \frac{\partial E}{\partial W_{ij}}$$

где v - константа, характеризующая скорость обучения. Данная формула описывает процесс градиентного спуска в пространстве весов.

Алгоритм обратного распространением состоит из следующих шагов.

Шаг 1. На вход НС подается вектор X из обучающей выборки и вычисляются выходы всех нейронов Y_{ij} .

Шаг 2. Определяется величина градиента ошибки EI для каждого нейрона выходного слоя:

$$EI_{Nj} = (Y_j - Y_{жел.j}) \cdot Y_j \cdot (1 - Y_j)$$

где Y_j – выход j -го нейрона выходного слоя.

Шаг 3. Двигаясь от последнего слоя к первому определяются градиенты EI_{ij} для каждого j -го нейрона каждого i -го слоя:

$$EI_{ij} = Y_{ij} \cdot (1 - Y_{ij}) \cdot \sum_{k=1}^{n(i+1)} EI_{i+1,k} \cdot w_{i+1,k}^j$$

где k – номер синапса, соединяющего нейрон N_{ij} с нейроном $N_{i+1,k}$ следующего слоя.

Шаг 4. Коррекция весов синапсов:

$$w_{ij}^k = w_{ij}^k - v \cdot EI_{ij} \cdot Y_{i-1,k}$$

Коррекция весов для входного слоя не производится.

Шаг 5. Если обучающая выборка не закончилась, то шаги 1 – 5 повторяются.

Шаг 6. Определяется величина ошибки E . Если она не удовлетворительна, то шаги 1 – 6 повторяются.

После успешного обучения НС может быть протестирована на тестовой выборке. Если Ошибка обучения на каждом примере из тестовой выборки удовлетворительна, то НС можно считать обученной и приступать к ее эксплуатации.

3. ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

3.1. Порядок выполнения работы

Лабораторная работа выполняется с использованием имитатора нейронной сети **Neural Network Wizard** в порядке, описанном ниже.

Необходимо:

1) Получить задание согласно варианту, указанному преподавателем (см. табл.).

В вариантах 1 – 3 необходимо с помощью нейронной сети провести аппроксимацию функции вида $y = f(x)$ на заданном интервале x (входная переменная сети – x , выходной – y).

Вариант 4: расшифровка двоичного 4-разрядного кода и представление его в 10-м виде, то есть, на вход сети подаются четыре логические переменные, кодирующие двоичное число от 0 до 16, сеть перекодирует его в десятичный вид и выдает в виде выходной переменной y .

Варианты 5 – 8: аппроксимация логических функций нескольких переменных X_i , выходная переменная Y – логическая.

Варианты 9 – 12: аппроксимация функций вида $y = f(X)$, где X – вектор входных переменных на заданном интервале X .

Для вариантов 3 и 9 диапазоны входных переменных выбираются самостоятельно. В вариантах 4 – 8 рассматривается вся область определения функции.

Вариант	Задание
1	$y = 2 \cdot \sin^2 x, x \in [0, 1]$
2	$y = 4 \frac{\cos x}{x}, x \in [-\pi, \pi]$
3	$y = 0.5 \cdot x^2 - 4.8 \cdot x + 3.5$
4	Представление двоичного 4-разрядного кода в 10-чном виде
5	$Y = X_1 \cdot X_2 + X_3 \cdot X_4$
6	$Y = \overline{X_1} \cdot \overline{X_2} + X_3 + X_4$
7	$Y = X_1 \cdot X_2 + \overline{X_3} \cdot X_4$
8	$Y = X_1 \cdot \overline{X_2} + X_3 \cdot \overline{X_4}$
9	$y = 2 \cdot x_1 + 5 \cdot x_1 \cdot x_2 + x_2$
10	$y = \sin x_1 \cdot \sin x_2, x_i \in [0, \pi]$
11	$y = 2 \cdot x_1 \cdot \cos x_2, x_1 \in [0, 1], x_2 \in [0, \pi]$
12	$y = x_1 + x_2 + x_3, x_i \in [0, 10]$

- 2) Подготовить обучающую выборку и оформить ее в виде текстового файла *.txt с колонками значений входных и выходных переменных.
- 3) Провести обучение нескольких нейронных сетей с помощью **Neural Network Wizard**. Рассматриваются пять вариантов трехслойной сети с количеством нейронов в скрытом слое, равном 1, 3, 5, 7 и 9. Обучение каждой сети вести, например, в течение 5 минут.
- 4) Проверить качество каждой обученной сети, для чего рассчитать несколько значений выходной переменной по заданной функции и по нейронной сети, затем на графике показать близость каждой обученной модели к исходной.
- 5) Результаты оформить в виде отчета. В отчете представить:
 - цель работы,
 - задание,
 - обучающую выборку,
 - параметры каждой обученной сети: числа нейронов во входном, скрытом и выходном слоях, числа эпох и т.д. (см. отчет, который выдает программа),
 - графики, построенные для заданной функции и обученных моделей (для вариантов 4 – 8 приводятся таблицы значений функций),
 - вывод о точности аппроксимации.

3.2. Порядок работы с Neural Network Wizard

Программным обеспечением лабораторной работы является **Neural Network Wizard**, который может быть использован без предварительной инсталляции.

Путь к программе: **Neural Network Wizard\Bin\Wizard.exe**.

Прежде, чем запускать программу, необходимо в любом текстовом редакторе подготовить текстовый файл с обучающей выборкой. Пример такого файла для функции $res = s1 + s2$ приведен ниже.

s1	s2	res
0		0
1		2
2		4
3		6
4		8
5		10
6		12

В первой строке файла указываются имена входных/выходных переменных. Далее идут их значения в колонках.

Файл сохраняется как текстовый с расширением TXT.

После запуска программы в первом окне задается имя файла. Кнопка «Далее» позволяет перейти к следующему окну (см. рис. 5). В новом окне перечислен список доступных полей, взятый из первой строки указанного файла. Для каждого из полей необходимо указать, чем является данная переменная. Если данная переменная входная, то в группе «Использовать поле как» выбирается вариант «Входное», если выходная – «Целевое». Кроме того, для каждой переменной можно указать вид нормализации (приведения к диапазону $[0,1]$) с соответствующими параметрами. Кнопка «Далее» переводит к следующему окну.

В данном окне (см. рис. 6) определяются структура и параметры НС: количество скрытых слоев, количество нейронов в каждом слое, а также вид сигмоидной функции.

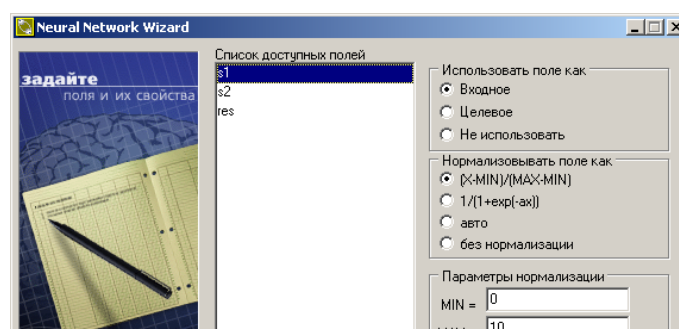
В следующем окне (см. рис. 7) задаются параметры обучения и критерии остановки обучения, если она требуется. Кнопка «Далее» показывает краткий предварительный отчет.

В следующем окне визуализирован процесс обучения (см. рис. 8). Чтобы запустить обучение нажимается кнопка «Запуск обучения». На верхней диаграмме показано распределение ошибки обучения: по горизонтали значение ошибки (чем правее столбец, тем больше ошибка), по вертикали количество примеров из выборки с данной ошибкой. Зеленые столбцы – ошибка на рабочей обучающей выборке, красные – на тестовой. В процессе обучения столбцы должны стремиться в левую часть диаграммы.

Ниже диаграммы отображается распределение примеров на рабочей и тестовой выборках. Каждый пример изображен здесь точкой. Чем ближе точка к диагонали, тем точнее НС предсказала ее значение.

Остановка обучения происходит либо по ранее указанному критерию, либо с помощью той же кнопки «Запуск обучения».

В следующем окне (см. рис. 9) представляется возможность оценить точность работы НС в эксплуатационном режиме. Для этого в левом поле указывается произвольное значение входного сигнала. После нажатия на кнопку «Расчет» в правом поле появляется рассчитанное сетью значение выходного.



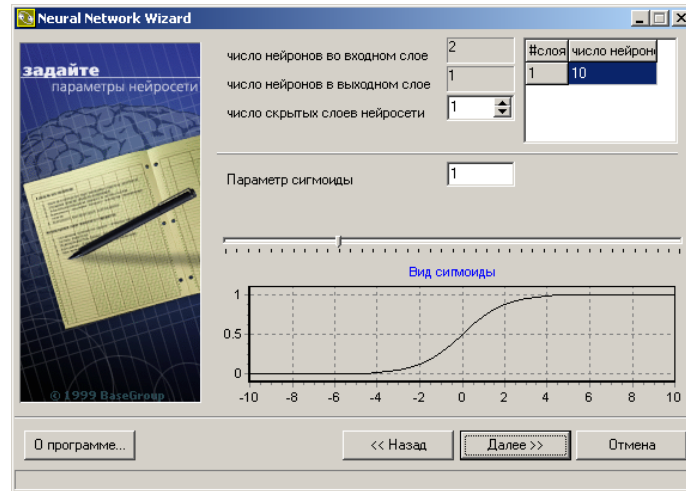


Рисунок 6

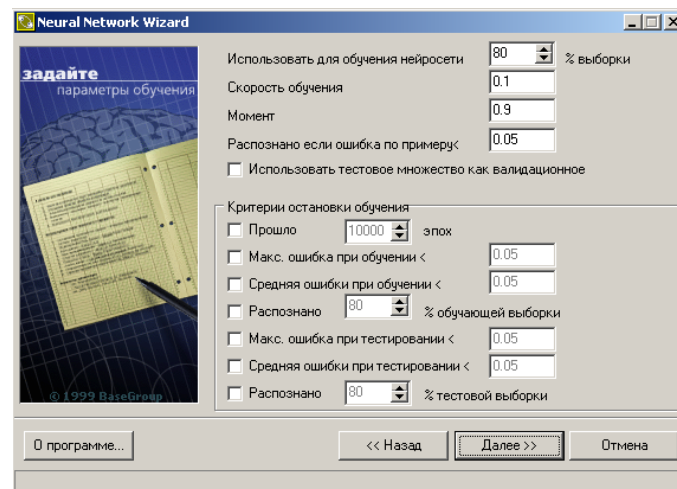


Рисунок 7

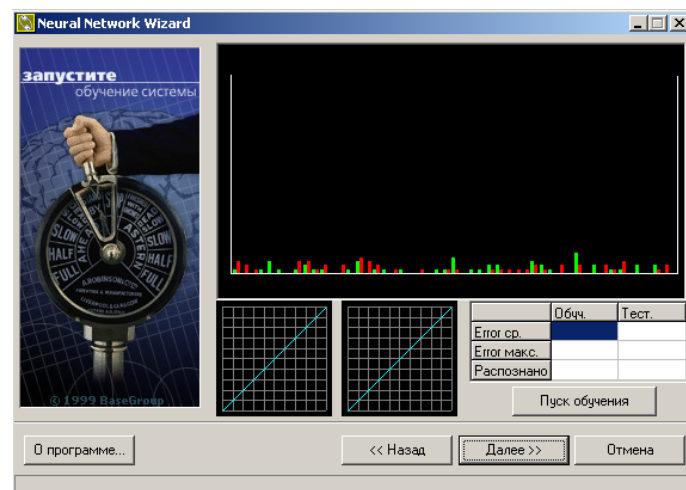


Рисунок 8

Кнопка «Сохранить» позволяет записать параметры обученной сети в виде файла (по умолчанию расширение файла NNW).

В данном файле кроме прочих параметров указаны:

EPOCH – количество эпох (циклов) обучения,

Layer_* - количество нейронов в соответствующем слое (нейроны нумеруются от 0 до N-1),

$W_{i_j_k}$ – веса синапсов (i = номер слоя - 2, j – номер нейрона, k – номер синапса данного нейрона).

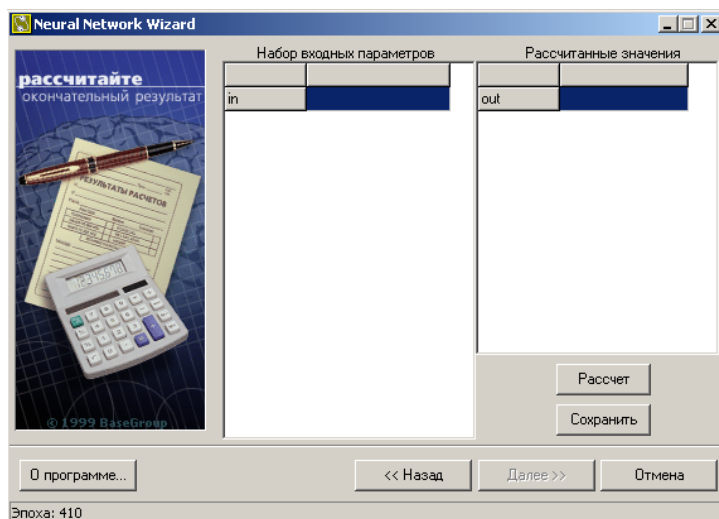


Рисунок 9

4. Контрольные вопросы

1. Что такое нейронная сеть и каковы ее основные свойства?
2. Какова структура нейрона?
3. Какие функции активации могут быть использованы в нейронных сетях? Какие требования предъявляются к функциям активации?
4. Какие функции выполняет входной слой в многослойной сети?
5. Можно ли обучить нейронную сеть без скрытого слоя?
6. В чем заключается обучение нейронных сетей?
7. Почему один из алгоритмов обучения получил название «алгоритм обратного распространения»?
8. Чем отличается обучение с учителем от обучения без учителя?
9. Почему входные и выходные сигналы нейронной сети должны быть нормированы, т.е. приведены к диапазону $[0,1]$?

Список литературы

1. Fogelman Soulie F. Neural networks, state of the art, neural computing.// London: IBC Technical Services, 1991.
2. Горбань А. Нейроинформатика и ее приложения. // Открытые системы. -№ 4-5, 1998. -С. 36 - 41.
3. Роберт Хехт-Нильсен. Нейрокомпьютинг: история, состояние, перспективы. // Открытые системы. -№ 4-5, 1998. -С. 23 - 28.
4. Розенблатт Ф. Принципы нейродинамики. Персептроны и теория механизмов мозга. -М.: Мир, 1965.
5. Гордиенко Е.К., Лукьяница А.А. Искусственные нейронные сети. I. Основные определения и модели.// Изв. РАН. Техническая кибернетика, 1994, № 5. -С. 79 - 92.
6. Короткий С.Г. Нейронные сети: алгоритм обратного распространения. // ВУТЕ/Россия, 2000, № 5, -С. 26-29.
7. Свешников С.В., Шквар А.М. Нейротехнические системы обработки информации. -Киев: Наукова думка, 1983. -222 с.

8. Интеллектуальные системы управления с использованием нейронных сетей: Учеб. пособие. / В.И. Васильев, Б.Г. Ильясов, С.С. Валеев и др.; Уфимск. гос. авиац. техн. ун-т. Уфа, 1997. -92 с.
9. Куликов Г.Г., Брейкин Т.В., Арьков В.Ю. Интеллектуальные информационные системы: Учеб. пособие / Уфимск. гос. авиац. техн. ун-т. -Уфа, 1999. -129 с.
10. Короткий С.Г. Нейронные сети: основные положения. // ВУТЕ/Россия, 2000, № 5, -С. 18-21.
11. Интеллектуальные системы обработки информации на основе нейросетевых технологий: Учеб. пособие. / Ю.И. Зозуля, Уфимск. гос. авиац. техн. ун-т. -Уфа 2000. -138 с.

