

Федеральное агентство связи
Федеральное государственное бюджетное образовательное учреждение высшего образования
«Поволжский государственный университет телекоммуникаций и информатики»

Кафедра Экономических и информационных систем

Методические указания по выполнению практического задания
по дисциплине

Инженерия знаний

для студентов 2 и 2у курсов заочной и дистанционной форм
обучения

по направлению подготовки 09.03.01 «Информатика и
вычислительная техника»

Самара
2015

Тема практического задания: Нейронные сети. Алгоритмы обучения нейронных сетей.

Для выполнения практического задания необходимо решить две задачи.

Нейронные сети

Искусственная нейронная сеть – математическая модель, реализуемая программно, построенная по подобию естественных нейронных сетей (сетей нервных клеток живого организма), представляющая собой соединение простых взаимодействующих между собой процессоров – искусственных нейронов.

Схема искусственного нейрона представлена на рис.1, где x_1, \dots, x_n – входы нейрона, w_1, \dots, w_n – синапсические веса связей нейрона, $\sum w_i x_i = S$ – взвешенная сумма входных значений нейрона, $f(\sum w_i x_i) = F(S)$ – функция активации, значением которой является Y – выходное значение нейрона.

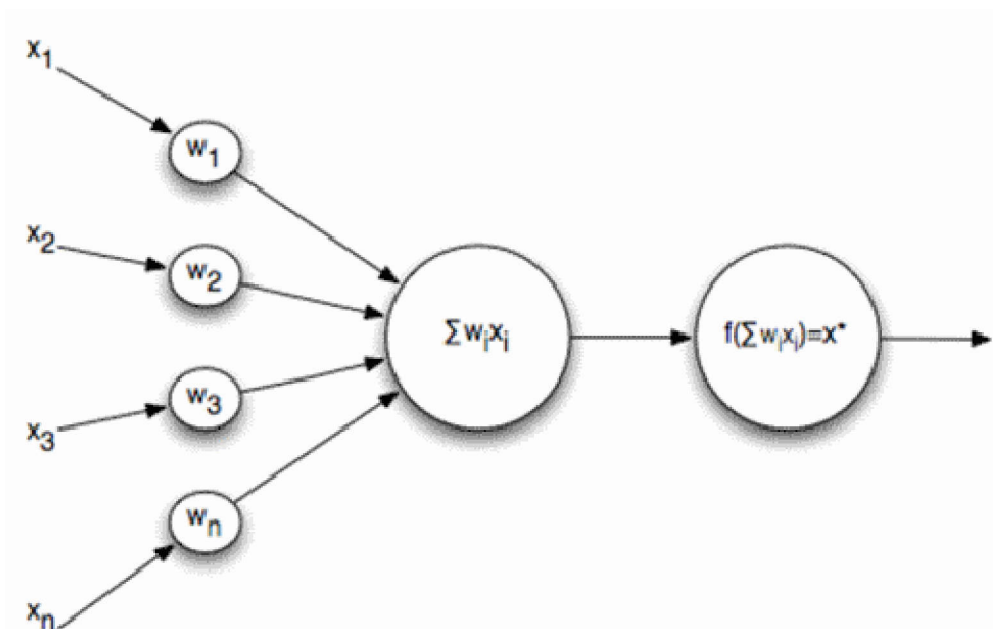


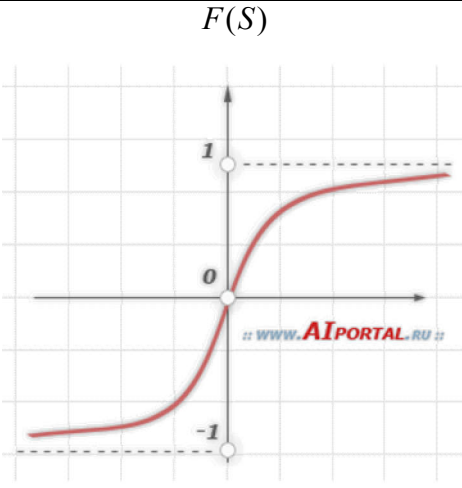
Рис.1 – Формальный нейрон (математическая модель)

Функция активации (активационная функция, функция возбуждения) – функция, вычисляющая выходной сигнал искусственного нейрона. В качестве аргумента принимает сигнал Y , получаемый на выходе входного сумматора Σ .

Наиболее распространенные функции активации представлены в табл.1.

Табл.1 – Основные функции активации

Название	Формула	График
Жесткая пороговая ступенька	$F(S) = \begin{cases} 1, & \text{при } S \geq \theta, \\ 0, & \text{при } S < \theta, \end{cases}$ $\theta = const$	
Линейная	$F(S) = \alpha \cdot S$ $\alpha = const$	
Сигмоидальная логистическая	$F(S) = \frac{1}{1 + e^{-S \cdot \alpha}}$ $\alpha = const$	

Гиперболический тангенс	$F(S) = th\left(\frac{S}{\alpha}\right) = \frac{e^{\frac{S}{\alpha}} + e^{-\frac{S}{\alpha}}}{e^{\frac{S}{\alpha}} - e^{-\frac{S}{\alpha}}}$ $\alpha = const$	
-------------------------	---	--

Объединение нейронов в общую сеть осуществляется различными способами. Существует множество нейронных сетей, которые классифицируются по нескольким признакам (табл.2).

Тип	Описание
По топологии	
Полносвязные	Каждый нейрон связан с другим нейроном в сети (из-за высокой сложности обучения не нашли широкого использования)
Слоистые	Нейроны располагаются слоями, каждый нейрон последующего слоя связан с нейронами предыдущего. Различают однослойные и многослойные сети.
По типу связей	
Прямого распространения	Все связи между нейронами идут от выходов нейронов предыдущего слоя к входам нейронов последующего
Рекуррентные	Допускаются связи выходов нейронов последующих слоев с входами нейронов предыдущих
По организации обучения	
С учителем	При обучении используются обучающие выборки, в которых определены требуемые от сети выходные значения, такие сети используют для решения задач классификации
Без учителя	Нейронная сеть сама в процессе работы выделяет классы объектов и относит объект к определенному классу, такие сети используют для задач кластеризации
По типу сигнала	
Бинарные	На вход нейронных сетей подают только нули и единицы

Аналоговые	Подаваемые на входы нейронов сигналы могут быть произвольными (вещественными числами)
<i>По типу структур</i>	
Однородная	Все нейроны в нейронной сети используют одну функцию активации
Неоднородная	Нейроны в нейронной сети имеют разные функции активации

Для решения конкретной задачи нужно выбрать подходящую нейронную сеть. При этом нужно учитывать не только перечисленные в таблице критерии, но и архитектуру сети. Выбор архитектуры подразумевает определение количества слоев и нейронов в этих слоях. Не существует формального алгоритма по определению нужной архитектуры, поэтому на практике выбирают или заведомо маленькую сеть и постепенно ее наращивают или заведомо большую и постепенно выявляют неиспользуемые связи и сокращают сеть.

Нейронная сеть, прежде чем использоваться на практике для решения какой-либо задачи должна быть обучена. Обучение нейронной сети – это процесс настройки синаптических весов. Существует множество алгоритмов, ориентированных на определенные типы сетей и на конкретные задачи. Рассмотрим алгоритмы для однослойной и многослойной сетей.

Алгоритм обучения сети по Δ -правилу

Простейшая нейронная сеть – однослойная (рис.2), представляющая из себя расположенные параллельно нейроны, получающие на входы одинаковые сигналы, но имеющие различные синаптические связи. Количество входов и выходов такой нейронной сети соответствует количеству нейронов.

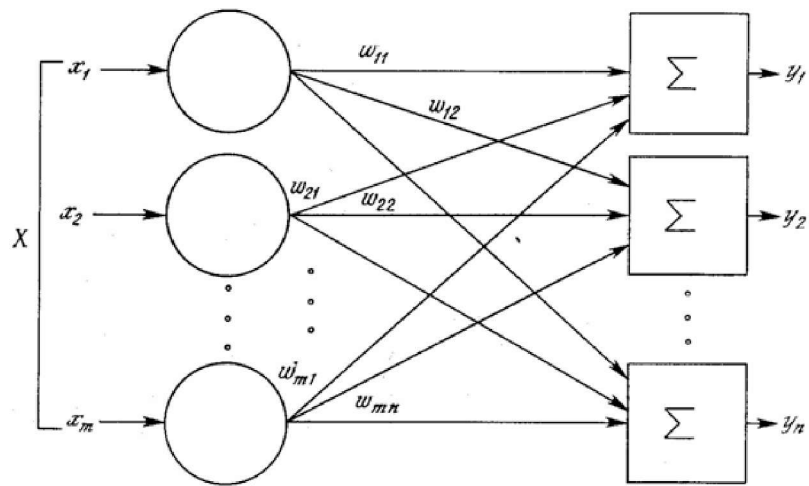


Рис.2 – Однослойная нейронная сеть

Такие нейронные сети можно обучать с помощью алгоритма обучения по Δ -правилу.

Алгоритм обучения по Δ -правилу:

1 шаг: инициализация матрицы весов (и порогов, в случае использования пороговой функции активации) случайным образом.

2 шаг: предъявление нейронной сети образа (на вход подаются значения из обучающей выборки – вектор X), берется соответствующий выход (вектор D).

3 шаг: вычисление выходных значений нейронной сети (вектор Y).

4 шаг: вычисление для каждого нейрона величины расхождения реального результата с желаемым:

$$\varepsilon_i = (d_i - y_i),$$

где d_i - желаемое выходное значение на i -м нейроне,

y_i - реальное значение на i -м нейроне.

5 шаг: изменение весов (и порогов при использовании пороговой функции) по формулам:

$$w_{ij}(t+1) = w_{ij}(t) - \eta \cdot \varepsilon_i \cdot x_j$$

$$\theta_i(t+1) = \theta_i(t) - \eta \cdot \varepsilon_i$$

где t - номер текущей итерации цикла обучения,

w_{ij} - вес связи j -го входа с i -ым нейроном,

η - коэффициент обучения, задается от 0 до 1,

x_j - входное значение,

θ_i - пороговое значение i -го нейрона.

6 шаг: проверка условия продолжения обучения (вычисление значения ошибки и/или проверки заданного количества итераций). Если обучение не завершено, то 2 шаг, иначе – заканчиваем обучение.

Пример решения задачи

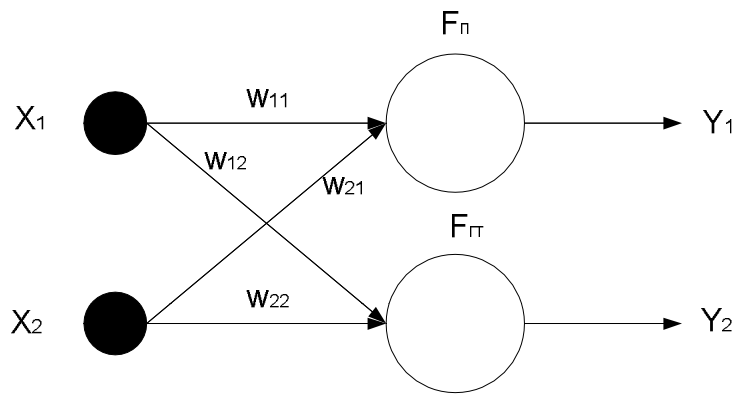
Задача. Просчитать одну итерацию цикла обучения по Δ -правилу однослойной бинарной неоднородной нейронной сети, состоящей из 2 нейронов и имеющей функции активации: жесткая ступенька ($\theta = 0,7$) и гиперболический тангенс ($\alpha = 1$). В качестве обучающей выборки использовать таблицу истинности для операций эквивалентности и дизъюнкции (не использовать первую строчку таблицы). Синапсические веса задать случайным образом.

Описание процесса решения. Для обучения нейронной сети по Δ -правилу необходимо:

- 1) Графически отобразить структуру нейронной сети. Определить размерность матрицы синапсических весов.
- 2) Определить обучающую выборку, представив ее в табличном виде.
- 3) Выбрать входные данные, на которых будет рассматриваться итерация цикла обучения.
- 4) Следуя алгоритму обучения по Δ -правилу, просчитать одну итерацию цикла и представить новые синапсические веса в матричном виде.

Решение.

- 1) По заданию нейронная сеть состоит из двух нейронов, значит, входов у однослойной нейронной сети будет 2 и выходов 2, а синапсических весов 4. Первый нейрон имеет функцию активации – жесткая ступенька, второй – гиперболический тангенс.



2) По заданию нейронная сеть бинарная, поэтому на ее входы могут подаваться только нули и единицы, так как входа 2, то возможных комбинаций входных значений будет 4 (обучающая выборка будет состоять из 4 векторов). Выход первого нейрона согласно заданию соответствует оператору эквивалентности, а второго – дизъюнкции. Поэтому таблица с обучающей выборкой будет выглядеть следующим образом:

x_1	x_2	d_1	d_2
0	0	1	0
0	1	0	1
1	0	0	1
1	1	1	1

3) Пусть в качестве вектора обучения будет рассматриваться третья строка таблицы.

4) Пошагово выполним алгоритм обучения по Δ -правилу.

1 шаг: зададим матрицу весов случайным образом из интервала $[0,1]$

$w_{ij}(1)$	1	2
1	0,7	0,9
2	0,5	0,2

2 шаг: вектор $X = \{1,0\}$, вектор $D = \{0,1\}$.

3 шаг: вычисление выходных значений нейронной сети (вектор Y).

$$\theta = 0,7;$$

$$S_1 = x_1 \cdot w_{11} + x_2 \cdot w_{21} = 1 \cdot 0,7 + 0 \cdot 0,5 = 0,7;$$

$$Y_1 = \begin{cases} 1, \text{ при } S_1 \geq \theta \\ 0, \text{ при } S_1 < \theta \end{cases} = \begin{cases} 1, \text{ при } 0,7 \geq 0,7 \\ 0, \text{ при } 0,7 < 0,7 \end{cases} = 1.$$

$$\alpha = 1,$$

$$S_2 = x_1 \cdot w_{12} + x_2 \cdot w_{22} = 1 \cdot 0,9 + 0 \cdot 0,2 = 0,9;$$

$$Y_2 = th\left(\frac{S}{\alpha}\right) = \frac{e^{0,9} + e^{-0,9}}{e^{0,9} - e^{-0,9}} \approx 1,39.$$

4 шаг:

$$\varepsilon_1 = (d_1 - y_1) = (0 - 1) = -1,$$

$$\varepsilon_2 = (d_2 - y_2) = (1 - 1,39) = -0,39.$$

5 шаг: задаем η - коэффициент обучения от 0 до 1 и изменяем веса:

$$\eta = 0,8;$$

$$w_{11}(2) = w_{11}(1) - 0,8 \cdot \varepsilon_1 \cdot x_1 = 0,7 - 0,8 \cdot (-1) \cdot 1 = 1,5;$$

$$w_{21}(2) = w_{21}(1) - 0,8 \cdot \varepsilon_1 \cdot x_2 = 0,5 - 0,8 \cdot (-1) \cdot 0 = 0,5;$$

$$\theta_1(2) = \theta_1(1) - 0,8 \cdot \varepsilon_1 = 0,7 - 0,8 \cdot (-1) = 1,5;$$

$$w_{12}(2) = w_{12}(1) - 0,8 \cdot \varepsilon_2 \cdot x_1 = 0,9 - 0,8 \cdot (-0,39) \cdot 1 = 1,212;$$

$$w_{22}(2) = w_{22}(1) - 0,8 \cdot \varepsilon_2 \cdot x_2 = 0,2 - 0,8 \cdot (-0,39) \cdot 0 = 0,2.$$

$w_{ij}(2)$	1	2
1	1,5	1,212
2	0,5	0,2

6 шаг: вычислим среднеквадратическую ошибку (можно выбрать другие методы оценки ошибки):

$$\varepsilon = \sum_{i=1}^N (d_i - y_i)^2 = \sum_{i=1}^2 \varepsilon_i^2 = \varepsilon_1^2 + \varepsilon_2^2 = (-1)^2 + (-0,39)^2 = 1,1521,$$

N - количество нейронов.

Так как необходимо было рассмотреть одну итерацию цикла обучения, то завершаем обучение.

Задача №1 для самостоятельного решения

Просчитать одну итерацию цикла обучения по Δ -правилу однослойной бинарной неоднородной нейронной сети, состоящей из 2 нейронов и имеющей функции активации: линейную ($\alpha = 0,7$) и сигмоидальную логистическую ($\alpha = 1$). В качестве обучающей выборки использовать таблицу истинности для операций конъюнкции и импликации (не использовать первую строчку таблицы). Синапсические веса задать случайным образом.

Алгоритм обратного распространения ошибки

Многослойная искусственная нейронная сеть (рис.3) может содержать произвольное количество слоев (K), каждый слой состоит из нескольких нейронов, число которых также может быть произвольно (N_k – число нейронов в слое), количество входов n , количество выходов – $N = N_k$ нейронов в выходном (последнем) слое.

Слои между первым и последним называются промежуточными или скрытыми. Веса в такой сети имеют три индекса: i – номер нейрона следующего слоя, для которого связь входная, j – номер входа или нейрона текущего слоя, для которого связь входная, k – номер текущего слоя в нейронной сети (для входов вектора X , $k = 0$).

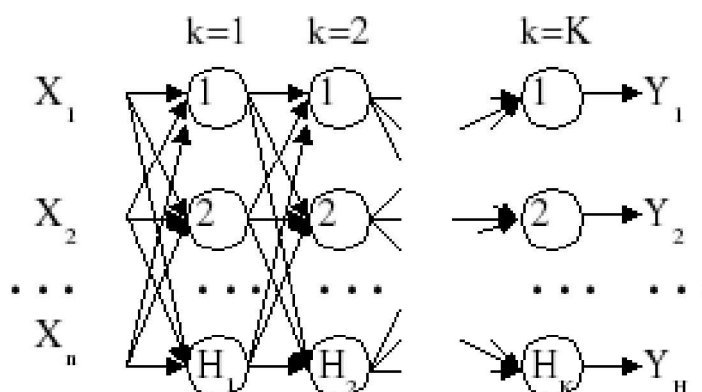


Рис.3 – Многослойная нейронная сеть прямого распространения

Многослойные нейронные сети прямого распространения обучаются методом обратного распространения ошибки.

Алгоритм обучения методом обратного распространения ошибки:

1 шаг: инициализация матриц весов случайным образом (в циклах).

2 шаг: предъявление нейронной сети образа (на вход подаются значения из обучающей выборки – вектор X) и берется соответствующий выход (вектор D).

3 шаг (прямой проход): вычисление в циклах выходов всех слоев и получение выходных значений нейронной сети (вектор Y).

$$y_i^k = f\left(\sum_{j=0}^{N_{k-1}} w_{ij}^k \cdot y_j^{k-1}\right),$$

$$y_j^0 = x_j ,$$

$$y_0^{k-1} = 1 ,$$

$$x_0 = 1 ,$$

где y_i^k - выход i -го нейрона k слоя,

f - функция активации,

w_{ij}^k - синапсическая связь между j -м нейроном слоя $k-1$ и i -м нейроном слоя k ,

x_j - входное значение.

4 шаг (обратный проход): изменение весов в циклах по формулам:

$$w_{ij}^k(t+1) = w_{ij}^k(t) + \eta \cdot \delta_i^k \cdot y_j^{k-1} .$$

$\delta_i^k = (d_i - y_i) \cdot y_i \cdot (1 - y_i)$ - для последнего слоя (выходного) слоя,

$$\delta_i^k = y_i \cdot (1 - y_i) \cdot \sum_{i=1}^{N_{k+1}} \delta_i^{k+1} \cdot w_i^{k+1} - \text{для промежуточных слоев,}$$

где t - номер текущей итерации цикла обучения (номер эпохи),

η - коэффициент обучения задается от 0 до 1,

y_i^k - выход i -го нейрона k слоя,

w_{ij}^k - синапсическая связь между j -ым нейроном слоя $k-1$ и i -ым нейроном слоя k ,

d_i - желаемое выходное значение на i -ом нейроне,

y_i - реальное значение на i -ом нейроне выходного слоя.

5 шаг: проверка условия продолжения обучения (вычисление значения ошибки и/или проверка заданного количества итераций). Если обучение не завершено, то 2 шаг, иначе заканчиваем обучение. Среднеквадратичная ошибка вычисляется следующим образом:

$$\varepsilon = \frac{1}{Q} \cdot \sum_{q=1}^Q \sum_{i=1}^N (d_i - y_i)^2 ,$$

где Q - общее число примеров,

N - количество нейронов в выходном слое,

d_i - желаемое выходное значение на i -ом нейроне,

y_i - реальное значение на i -ом нейроне выходного слоя.

Пример решения задачи

Задача. Просчитать одну итерацию цикла обучения методом обратного распространения ошибки многослойной бинарной неоднородной нейронной сети, состоящей из 2 слоев, причем в первом слое находится 2 нейрона и используется сигмоидальная логистическая функция активации ($\alpha = 0,9$), а во втором – 1, линейная ($\alpha = 0,7$) функция. В качестве обучающей выборки использовать таблицу истинности для операции «штрих Шеффера» (не использовать первую строчку таблицы). Синаптические веса задать случайным образом.

Описание процесса решения. Для обучения нейронной сети методом обратного распространения ошибки необходимо:

1) Графически отобразить структуру нейронной сети. Определить размерность и количество матриц синаптических весов (для каждого слоя своя матрица).

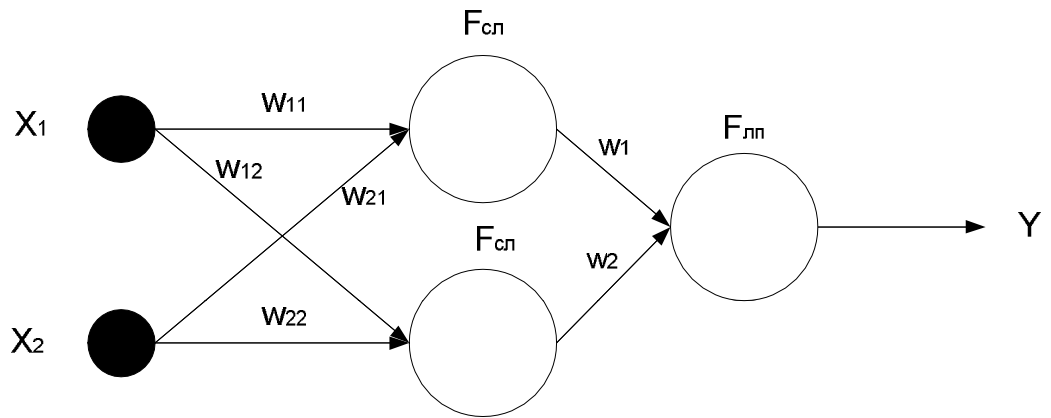
2) Определить обучающую выборку, представив ее в табличном виде.

3) Выбрать входные данные, на которых будет рассматриваться итерация цикла обучения.

4) Следуя алгоритму обучения методом обратного распространения ошибки просчитать одну итерацию цикла и представить новые синаптические веса в матричном виде.

Решение.

1) По заданию нейронная сеть состоит из трех нейронов: два входных, один выходной, следовательно синаптических весов – 6. Первый слой нейронов имеет сигмоидальную логистическую функцию активации, второй – линейную.



2) По заданию нейронная сеть бинарная, поэтому на ее входы могут подаваться только нули и единицы, так как входа 2, то возможных комбинаций входных значений будет 4 (обучающая выборка будет состоять из 4 векторов). Выход нейронной сети согласно заданию соответствует оператору «штрих Шеффера». Поэтому таблица с обучающей выборкой будет выглядеть следующим образом:

x_1	x_2	d_1
0	0	1
0	1	1
1	0	1
1	1	0

3) Пусть в качестве вектора обучения будет рассматриваться вторая строка таблицы.

4) Следуя алгоритму обучения методом обратного распространения ошибки, необходимо выполнить 5 шагов.

1 шаг: зададим матрицу весов случайным образом из интервала $[0,1]$

$w_{ij}(1)$	1	2
1	0,6	0,9
2	0,1	0,5

$w_g(1)$	1	2
1	0,3	0,8

2 шаг: вектор $X = \{0,1\}$, $D = \{1\}$.

3 шаг (прямой проход): вычисление в циклах выходов всех слоев и получение выходных значений нейронной сети (вектор Y).

$$S_1 = x_1 \cdot w_{11} + x_2 \cdot w_{21} = 0 \cdot 0,6 + 1 \cdot 0,1 = 0,1;$$

$$S_2 = x_1 \cdot w_{12} + x_2 \cdot w_{22} = 0 \cdot 0,9 + 1 \cdot 0,5 = 0,5;$$

$$\alpha = 0,9;$$

$$Y_1 = \frac{1}{1 + e^{-S_1 \cdot \alpha}} = \frac{1}{1 + e^{-0,1 \cdot 0,9}} = 0,5224;$$

$$Y_2 = \frac{1}{1 + e^{-S_2 \cdot \alpha}} = \frac{1}{1 + e^{-0,5 \cdot 0,9}} = 0,61;$$

$$S_3 = Y_1 \cdot w_1 + Y_2 \cdot w_2 = 0,5224 \cdot 0,3 + 0,61 \cdot 0,8 = 0,64472;$$

$$\alpha = 0,7;$$

$$Y = \alpha \cdot S = 0,4513.$$

4 шаг (обратный проход): изменение весов

$$\eta = 0,7;$$

$$\delta^2 = (d - Y) \cdot Y \cdot (1 - Y) = (1 - 0,4513) \cdot 0,4513 \cdot (1 - 0,4513) = 0,3587;$$

$$w_1(2) = w_1(1) + \eta \cdot \delta^2 \cdot Y_1 = 0,3 + 0,7 \cdot 0,3587 \cdot 0,5224 = 0,431;$$

$$w_2(2) = w_2(1) + \eta \cdot \delta^2 \cdot Y_2 = 0,8 + 0,7 \cdot 0,3587 \cdot 0,61 = 0,953.$$

$$\delta_1^1 = Y_1(1 - Y_1) \cdot \sum_{i=1}^{N_{k+1}} \delta_i^{k+1} \cdot w_i^{k+1} = Y_1 \cdot (1 - Y_1) \cdot \delta^2 \cdot w_1 = 0,5224 \cdot (1 - 0,5224) \cdot 0,3587 \cdot 0,3 = 0,0268;$$

$$\delta_2^1 = Y_2(1 - Y_2) \cdot \sum_{i=1}^{N_{k+1}} \delta_i^{k+1} \cdot w_i^{k+1} = Y_2 \cdot (1 - Y_2) \cdot \delta^2 \cdot w_2 = 0,61 \cdot (1 - 0,61) \cdot 0,3587 \cdot 0,8 = 0,0682;$$

$$w_{11}(2) = w_{11}(1) + \eta \cdot \delta_1^1 \cdot x_1 = 0,6 + 0,7 \cdot 0,0268 \cdot 0 = 0,6;$$

$$w_{12}(2) = w_{12}(1) + \eta \cdot \delta_2^1 \cdot x_1 = 0,9 + 0,7 \cdot 0,0682 \cdot 0 = 0,9;$$

$$w_{21}(2) = w_{21}(1) + \eta \cdot \delta_1^1 \cdot x_2 = 0,1 + 0,7 \cdot 0,0268 \cdot 1 = 0,119;$$

$$w_{22}(2) = w_{22}(1) + \eta \cdot \delta_2^1 \cdot x_2 = 0,5 + 0,7 \cdot 0,0682 \cdot 1 = 0,548.$$

$w_{ij}(2)$	1	2
1	0,6	0,9
2	0,119	0,548

$w_g(2)$	1	2
1	0,431	0,953

5 шаг:

$$\varepsilon = \sum_{i=1}^N (d_i - y_i)^2 = (1 - 0,4513)^2 = 0,237.$$

Так как необходимо было рассмотреть одну итерацию цикла обучения, то завершаем обучение.

Задача №2 для самостоятельного решения

Просчитать одну итерацию цикла обучения методом обратного распространения ошибки многослойной бинарной неоднородной нейронной сети, состоящей из 2 слоев, причем в первом слое находится 2 нейрона и используется функция активации – гиперболический тангенс ($\alpha = 2$), а во втором – 1, функция - жесткая ступенька ($\theta = 0,6$). В качестве обучающей выборки использовать таблицу истинности для операции «штрих Шеффера» (не использовать первую строку таблицы). Синаптические веса задать случайным образом.