

Лабораторная работа № 1

Исследование однослойных нейронных сетей на примере моделирования булевых выражений

Цель работы — исследовать функционирование простейшей нейронной сети (НС) на базе нейрона с нелинейной функцией активации и обучить ее по правилу Видроу — Хоффа.

Теоретическая часть

Постановка задачи. Получить модель булевой функции (БФ) на основе однослойной НС (единичный нейрон) с двоичными входами $x_1, x_2, x_3, x_4 \in \{0, 1\}$, единичным входом смещения $x_0 = 1$, синаптическими весами w_0, w_1, w_2, w_3, w_4 , двоичным выходом $y \in \{0, 1\}$ и заданной нелинейной функцией активации (ФА) $f: R \rightarrow (0, 1)$ (рис. 1.1).

Для заданной БФ (варианты см. в табл. 1.4) реализовать обучение НС с использованием:

- 1) всех комбинаций переменных x_1, x_2, x_3, x_4 ;
 - 2) части возможных комбинаций переменных x_1, x_2, x_3, x_4 ;
- остальные комбинации являются тестовыми.

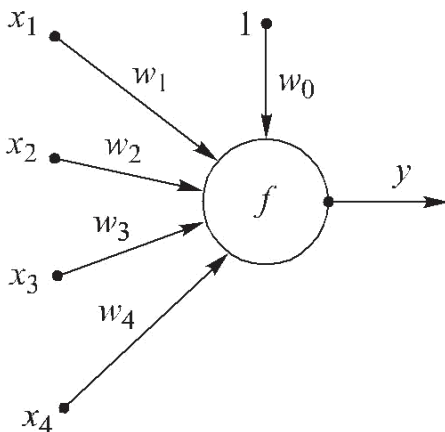


Рис. 1.1. Однослойная НС

Рабочий режим НС. Алгоритм функционирования НС с пороговой ФА имеет вид

$$\text{net} = \sum_{i=1}^4 w_i x_i + w_0; \quad (1.1)$$
$$y(\text{net}) = \begin{cases} 1, & \text{net} \geq 0, \\ 0, & \text{net} < 0, \end{cases}$$

где net — сетевой (комбинированный) вход; y — реальный выход НС.

Алгоритм функционирования НС с *логистической* ФА выглядит следующим образом:

$$\begin{aligned} \text{net} &= \sum_{i=1}^4 w_i x_i + w_0; \\ \text{out} &= f(\text{net}), \\ y(\text{out}) &= \begin{cases} 1, & \text{out} \geq 0,5, \\ 0, & \text{out} < 0,5, \end{cases} \end{aligned} \quad (1.2)$$

где out — сетевой (недискретизированный) выход НС.

Режим обучения НС. Для необученной НС ее реальный выход y в общем случае отличается от целевого выхода t , представляющего собой значения заданной БФ нескольких переменных $F(x_1, x_2, x_3, x_4): \{0, 1\}^4 \rightarrow \{0, 1\}$, т. е. имеется хотя бы один набор сигналов (x_1, x_2, x_3, x_4) , для которого ошибка

$$\delta = t - y \neq 0. \quad (1.3)$$

Каждая эпоха обучения $k = 1, 2, \dots$ включает в себя цикл последовательного предъявления всех образцов обучающей выборки на вход НС. Предъявление одного обучающего образца внутри эпохи является элементарным *шагом обучения* $l = 0, 1, 2, \dots$, во время которого вектор весовых коэффициентов $\mathbf{w} = (w_i)_{i=0, \dots, 4}$ корректируется согласно *правилу Видроу — Хоффа (дельта-правило)*:

$$\begin{aligned} w_i^{(l+1)} &= w_i^{(l)} + \Delta w_i^{(l)}, \\ \Delta w_i^{(l)} &= \eta \delta^{(l)} \frac{df(\text{net})}{d \text{net}} x_i^{(l)}, \end{aligned} \quad (1.4)$$

где $\Delta w_i^{(l)}$ — коррекция веса на l -м шаге; $\eta \in (0, 1]$ — норма обучения; $\delta^{(l)}$ — ошибка (1.3) на l -м шаге; $x_i^{(l)}$ ($i = 0, \dots, 4$) — компоненты обучающего вектора на l -м шаге обучения (при этом $x_0^{(l)} \equiv 1$).

Для наиболее популярных сигмоидальных ФА производная $df(\text{net})/d \text{net}$ легко выражается через значения функции $f(\text{net})$.

В случае НС с пороговой ФА (1.1) коррекцию веса в (1.4) следует брать в виде

$$\Delta w_i^{(k)} = \eta \delta^{(l)} x_i^{(l)}. \quad (1.5)$$

При этом реальный выход совпадает с сетевым: $y = \text{out}$.

На каждой эпохе k суммарная квадратичная ошибка $E(k)$ равна расстоянию Хемминга между векторами целевого и реального выходов (размерности $N = 2^4 = 16$) по всем входным векторам $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$.

Практическая часть

Пример выполнения работы. Получим нейросетевую модель БФ (табл. 1.1):

$$F(x_1, x_2, x_3, x_4) = \overline{x_1 x_2 x_3 x_4}.$$

Таблица 1.1

Таблица истинности БФ

x_1	x_2	x_3	x_4	F
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	0	1	1	1
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	1
1	0	0	0	0
1	0	0	1	0
1	0	1	0	0
1	0	1	1	1
1	1	0	0	0
1	1	0	1	0
1	1	1	0	0
1	1	1	1	0

На начальном шаге $l = 0$ (эпоха $k = 0$) весовые коэффициенты возьмем в виде

$$w_0^{(0)} = w_1^{(0)} = w_2^{(0)} = w_3^{(0)} = w_4^{(0)} = 0.$$

1. Используем пороговую ФА (1.1). Динамика НС (при норме обучения $\eta = 0,3$) представлена в табл. 1.2. График суммарной ошибки приведен на рис. 1.2.

Таблица 1.2

Параметры НС на последовательных эпохах (пороговая ФА)

Но- мер эпохи k	Вектор весов w	Выходной вектор y	Сум- марная ошиб- ка E
0	(0, 0, 0, 0, 0)	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	13
1	(-0,6, -0,3, -0,6, 0,3, 0,3)	(0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	2
...
24	(-2,1, -1,2, -0,9, 1,8, 1,5)	(0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0)	0

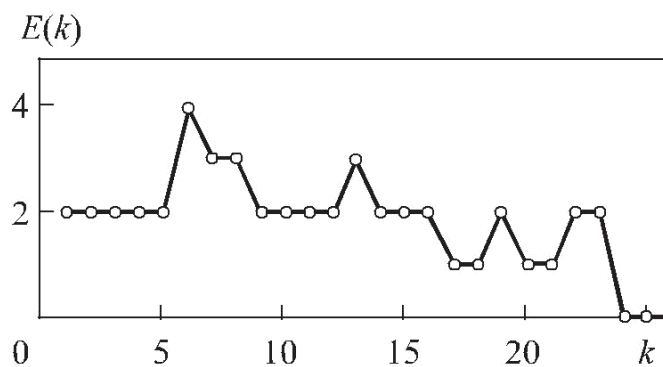


Рис. 1.2. График суммарной ошибки НС по эпохам обучения (пороговая ФА)

2. Используем сигмоидальную (логистическую) ФА:

$$f(\text{net}) = \frac{1}{1 + \exp(-\text{net})},$$

производная которой выражается через значения самой функции как

$$\frac{df(\text{net})}{d \text{net}} = f(\text{net})[1 - f(\text{net})].$$

Для полного обучения (при нулевых начальных весах и норме обучения $\eta = 0,3$) потребовалось 34 эпохи (табл. 1.3). График ошибки приведен на рис. 1.3.

**Параметры НС на последовательных эпохах
(логистическая ФА)**

Номер эпохи k	Вектор весов \mathbf{w}	Выходной вектор \mathbf{y}	Суммарная ошибка E
0	(0, 0, 0, 0, 0)	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	13
1	(-0,302, -0,170, -0,174, -0,028, -0,033)	(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	3
...
34	(-1,811, -0,918, -0,883, 1,379, 1,356)	(0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0)	0

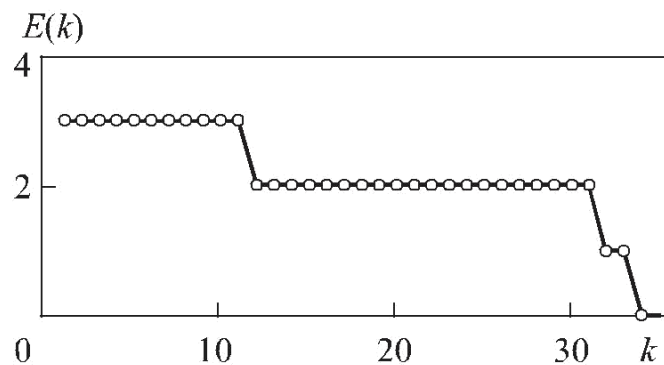


Рис. 1.3. График суммарной ошибки НС по эпохам обучения (логистическая ФА)

3. Рассматривая предыдущий случай, попытаемся последовательно уменьшать размер обучающей выборки, выбирая различные подмножества из 15, 14, 13 и т. д. обучающих векторов до тех пор, пока еще достижима нулевая ошибка. В рассматриваемом примере будет найден минимальный набор из пяти векторов

$$\mathbf{x}^{(1)} = (0, 0, 0, 1), \quad \mathbf{x}^{(2)} = (0, 1, 1, 1), \quad \mathbf{x}^{(3)} = (1, 0, 1, 0), \\ \mathbf{x}^{(4)} = (1, 0, 1, 1), \quad \mathbf{x}^{(5)} = (1, 1, 1, 0),$$

дающий следующие синаптические коэффициенты:

$$\mathbf{w} = (-0,225, -0,15, -0,075, 0,15, 0,224).$$

При этом для полного обучения потребовалось лишь 6 эпох.

Варианты заданий

Таблица 1.4

Варианты БФ и ФА

№ варианта	Моделируемая БФ	ФА*
1	$\overline{x_1 x_2 (x_3 + x_4)}$	1, 2
2	$(\overline{x_3 + x_4}) \overline{x_1} + x_2$	1, 3
3	$x_1 + \overline{x_2} + (\overline{x_3 + x_4})$	1, 4
4	$(\overline{x_1 + x_3}) x_2 + x_2 x_4$	1, 2
5	$(\overline{x_1 + x_2 + x_3})(x_2 + \overline{x_3 + x_4})$	1, 3
6	$x_3 x_4 + \overline{x_1} + \overline{x_2}$	1, 4
7	$\overline{(x_1 + x_2)} + x_3 + x_4$	1, 2
8	$(x_1 + x_2 + x_4) x_3$	1, 3
9	$(x_1 + x_2 + x_3)(x_2 + x_3 + x_4)$	1, 4
10	$x_1 x_2 + x_3 + x_4$	1, 2
11	$(x_1 + x_2) x_3 x_4$	1, 3
12	$(x_1 + x_2) x_3 + x_4$	1, 4
13	$(\overline{x_1 + x_2 + x_3})(\overline{x_2 + x_3 + x_4})$	1, 2
14	$\overline{\overline{(x_1 + x_2)} + x_3 + x_4}$	1, 3
15	$\overline{(x_1 + x_2) x_3 + x_3 x_4}$	1, 4
16	$\overline{(x_1 + x_2 + x_3)(x_2 + x_3 + x_4)}$	1, 2
17	$\overline{x_1 x_2 + x_3 + x_4}$	1, 3
18	$\overline{(x_1 + x_2) x_3 x_4}$	1, 4
19	$\overline{(x_1 + x_2) x_3 + x_4}$	1, 2
20	$\overline{x_3 x_4 (x_1 + x_2)}$	1, 3
21	$\overline{\overline{(x_1 + x_2) x_3 + x_4}}$	1, 4

№ варианта	Моделируемая БФ	ФА*
22	$\overline{x_3 + x_4 + (x_1 + x_2)}$	1, 2
23	$\overline{(x_2 + x_4)x_1 + x_1x_3}$	1, 3
24	$\overline{(x_4 + x_3 + x_2)(x_3 + x_2 + x_1)}$	1, 4
25	$\overline{x_1x_2 + x_3 + x_4}$	1, 2

* Функции активации:

1) $f(\text{net}) = \begin{cases} 1, & \text{net} \geq 0, \\ 0, & \text{net} < 0; \end{cases}$ 3) $f'(\text{net}) = \frac{1}{1 + \exp(-\text{net})}$;

2) $f(\text{net}) = \frac{1}{2} \left(\frac{\text{net}}{1 + |\text{net}|} + 1 \right)$; 4) $f(\text{net}) = \frac{1}{2} (\tanh(\text{net}) + 1)$.

Порядок защиты лабораторной работы. Для защиты лабораторной работы студент должен продемонстрировать программную реализацию НС, моделирующей заданную в варианте функцию, результаты обучения сети со всеми промежуточными шагами; предоставить подготовленный отчет согласно указанным ниже требованиям и ответить на контрольные вопросы по теоретической и практической частям работы.

Требования к отчету. В отчете должны содержаться: титульный лист; цель работы; постановка задачи; таблица истинности; график суммарной квадратичной ошибки в зависимости от эпохи обучения; целевые значения и реальный выход НС на каждой эпохе обучения; вектор синаптических коэффициентов на каждой эпохе обучения; минимальный набор обучающих векторов; выводы по результатам численного эксперимента.

Контрольные вопросы

1. Дайте определение персептрона и поясните алгоритм его функционирования.
2. Приведите функции активации НС и их производные.
3. Сформулируйте правило обучения Видроу — Хоффа.

