

Лабораторная работа № 5

СЕТЬ ОБРАТНОГО РАСПРОСТРАНЕНИЯ ОШИБКИ

1. Цель работы

Изучение топологии, алгоритма функционирования многослойного персептрона.

1. Теоретические сведения

Нейронная сеть обратного распространения состоит из нескольких слоев нейронов, причем каждый нейрон предыдущего слоя связан с каждым нейроном последующего слоя. В большинстве практических приложений оказывается достаточно рассмотрения двухслойной нейронной сети, имеющей входной (скрытый) слой нейронов и выходной слой (рис 1).

Матрицу весовых коэффициентов от входов к скрытому слою обозначим W , а матрицу весов, соединяющих скрытый и выходной слой, – как V .

Для индексов примем следующие обозначения: входы будем нумеровать только индексом i , элементы скрытого слоя – индексом j , а выходы, соответственно, индексом k . Число входов сети равно n , число нейронов в скрытом слое – m , число нейронов в выходном слое – p . Пусть сеть обучается на выборке (X^t, D^t) , $t = 1, 2, \dots, T$

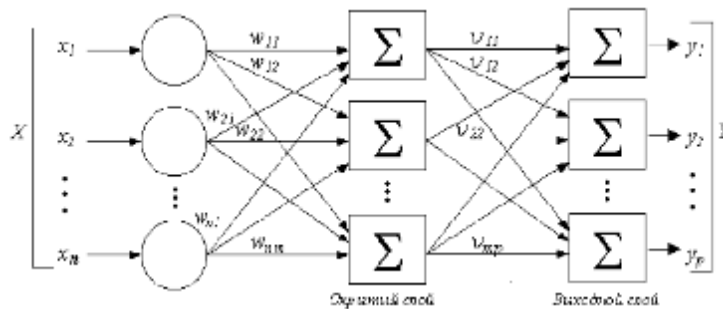


Рис. 1. Нейронная сеть обратного распространения

В качестве активационной функции в сети обратного распространения обычно используется логистическая функция $f(s) = \frac{1}{1 + e^{-s}}$, где s – взвешенная сумма входов нейрона.

При обучении нейронной сети ставится задача минимизации целевой функции ошибки, которая находится по методу наименьших квадратов:

$$E = \frac{1}{2} \sum_{k=1}^p (d_k - y_k)^2$$

где y_k – полученное реальное значение k -го выхода нейросети при подаче на нее одного из входных образов обучающей выборки; d_k – требуемое (целевое) значение k -го выхода для этого образа.

Алгоритм обучения сети обратного распространения

Рассмотрим теперь полный алгоритм обучения нейросети.

Шаг 1. Инициализация сети.

Весовым коэффициентам присваиваются малые случайные значения, например, из диапазона $(-0.3, 0.3)$; задаются ε – параметр точности обучения, α – параметр скорости обучения (как правило $\alpha \approx 0,1$ и может еще уменьшаться в процессе обучения), N_{\max} – максимально допустимое число итераций.

Шаг 2. Вычисление текущего выходного сигнала.

На вход сети подается один из образов обучающей выборки и определяются значения выходов всех нейронов нейросети.

Шаг 3. Настройка синаптических весов.

Рассчитать изменение весов для выходного слоя нейронной сети по формулам

$$v_{jk}^{N+1} = v_{jk}^N - \alpha \frac{\partial E}{\partial v_{jk}}, \text{ где } \frac{\partial E}{\partial v_{jk}} = \delta_k y_j^c, \delta_k = (y_k - d_k) y_k (1 - y_k).$$

Рассчитать изменение весов для скрытого слоя по формулам

$$w_{ij}^{N+1} = w_{ij}^N - \alpha \frac{\partial E}{\partial w_{ij}}, \text{ где } \frac{\partial E}{\partial w_{ij}} = \left(\sum_{k=1}^p \delta_k v_{jk}^{N+1} \right) y_j^c (1 - y_j^c) x_i.$$

Шаг 4.

Шаги 2–3 повторяются для всех обучающих векторов. Обучение завершается по достижении для каждого из обучающих образов значения функции ошибки, не превосходящего ε или после максимально допустимого числа итераций.

Замечание 1. На шаге 2 векторы из обучающей последовательности лучше предъявлять на вход в случайном порядке.

Замечание 2. Во многих случаях желательно наделять каждый нейрон обучаемым смещением. Это позволяет сдвигать начало отсчета логистической функции, давая эффект, аналогичный подстройке порога персептронного нейрона, и приводит к ускорению процесса обучения. Эта возможность может быть легко введена в обучающий алгоритм с помощью добавления к каждому нейрону дополнительного входа. Его вес обучается так же, как и все остальные веса, за исключением того, что подаваемый на него сигнал всегда равен +1, а не выходу нейрона предыдущего слоя.

Замечание 3. Количество входов и выходов сети, как правило, диктуется условиями задачи, а размер скрытого слоя находят экспериментально. Обычно число нейронов в нем составляет 30–50 % от числа входов. Слишком большое количество нейронов скрытого слоя приводит к тому, что сеть теряет способность к обобщению (она просто досконально запоминает элементы обучающей выборки и не реагирует на схожие образцы, что неприемлемо для задач распознавания). Если число нейронов в скрытом слое слишком мало, сеть оказывается просто не в состоянии обучиться.

Замечание 4. В результате слишком слишком близкой подгонки может возникнуть проблема переобучения. Если данные в обучающей выборке содержали шумы, то в результате переобучения сеть запоминает эти шумы и утрачивает способность к обобщению. Для контроля за переобучением сети вся доступная исходная выборка должна быть разделена на два множества – обучающее ($\approx 75\%$ данных) и тестовое ($\approx 25\%$ данных).

Замечание 5. Выходы каждого нейрона сети лежат в диапазоне $(0,1)$ – области значений логистической функции – это надо учитывать при формировании обучающей выборки. Желательно, чтобы и входы сети принадлежали этому диапазону. Если надо получить от сети бинарный выход, то, как правило, вместо 0 используют 0,1, а вместо «1» 0,9, так как границы интервала слишком недостижимы.

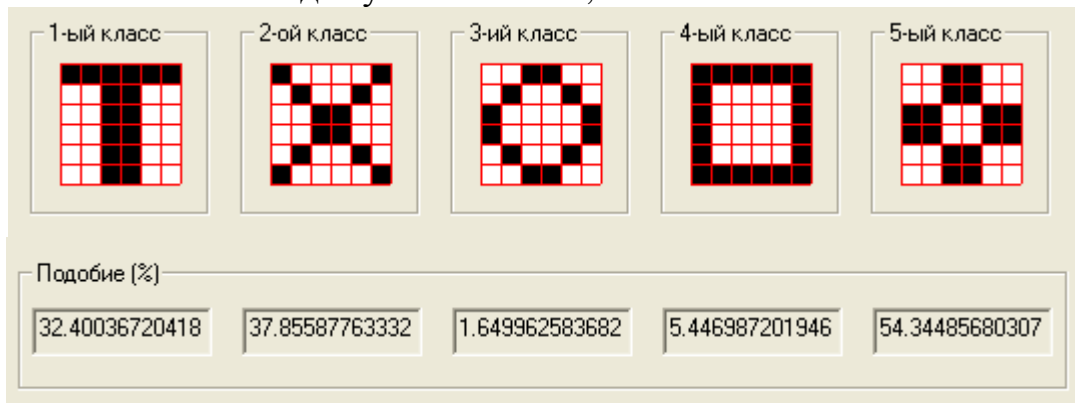
3. Задание

1. Ознакомьтесь с теоретической частью.
2. Напишите программу, реализующую алгоритм обучения сети обратного распространения.
3. Произведите обучение сети. Исходные данные – 5 классов образов, размер идеального образа 6×6 (для вариантов с 10 по 15 размер идеального образа необходимо увеличить).
4. Подайте на вход сети ряд тестовых образов, по 3 зашумленных образа каждого из 5 классов.
5. Проанализируйте результаты работы программы, которые должны иметь следующий вид:

- вывести распознаваемый зашумленный образ;



- вывести процент подобия распознаваемого зашумленного образа по отношению к каждому из 5 классов;



- вывести количество шагов, затраченных на обучение сети на заданное количество классов.

6. Напишите отчет.

Содержание отчета:

- топология многослойного персептрона;
- основные формулы обучения и воспроизведения;
- идеальные образы для обучения сети по алгоритму обратного распространения.
- тестовые зашумленные образы;
- результаты воспроизведения: процент подобия по отношению к каждому из классов, количество шагов, затраченных на обучение;
- выводы: преимущества и недостатки данного алгоритма

Таблица 1

Варианты задания

№ варианта	1-й класс	2-й класс	3-й класс	4-й класс	5-й класс
1	2	3	4	5	7
2	N	F	I	P	D
3	^	v	∃	⊂	⊃
4	⊗	⊕	×	÷	±
5	≤	≥	≠	≡	≈
6	L	U	T	O	K
7	→	↔	←	↓	↵
8	∩	∪	⊃	⊇	⊆
9	⊆	®	©	⊗	∅
10	♣	♦	♥	♠	•
11	J	K	L	>	⊗
12	0	⊖	⊗	Δ	‡
13	ý	þ	Ö	°	∴
14	V	X	Y	Z	{
15	з	‡	±	«	©