

.....  
**Алгоритм обучения методом градиентного спуска**  
.....

Шаг 1. Подготовить обучающую выборку, каждый элемент которой будет состоять из пар  $(X, D)_m$  ( $m=1, \dots, q$ ) – обучающего вектора  $X_m = (x_1, \dots, x_n)$  ( $i=1, \dots, n$ ) с вектором желаемых значений  $D_m = (d_1, \dots, d_k)$  ( $j=1, \dots, k$ ) выходов персептрона.

Шаг 2. Генератором случайных чисел всем синаптическим весам  $w_{i,j}$  и нейронным смещениям  $w_{0,j}$  ( $i=0, \dots, n$ ;  $j=1, \dots, k$ ) присваиваются некоторые малые случайные значения.

Шаг 3. Общей ошибке  $E_{общ}$ , вычисляемой на всех обучающих образах и всем значениям коррекции синаптических весов  $\Delta w_{i,j}$  присвоить нулевое значение.

$$E_{общ} = 0 \quad (1.1)$$

$$\Delta w_{i,j} = 0 \quad (1.2)$$

Шаг 4. Из обучающей выборки  $(X, D)_1, \dots, (X, D)_q$ , взять следующий по счету вектор  $X_m = (x_1, \dots, x_n)$  и подать его на входы персептрона  $x_1, \dots, x_n$ . Сигналам нейронных входов смещения  $x_0$  присваиваются единичные значения:  $x_0 = 1$ .

Шаг 5. Для каждого  $j$ -го нейрона вычислить взвешенную сумму входных сигналов  $net_j$  и выходной сигнал  $y_j$  на основании функции активации  $f$ :

$$net_j = \sum_{i=0}^n x_i w_{i,j} \quad (1.3)$$

$$y_j = f(net_j) \quad (1.4)$$

Шаг 6. Вычислить ошибку  $E$  для текущего обучающего вектора и сложить полученное значение с общей ошибкой  $E_{общ}$ :

$$E = \sum_{j=1}^k \frac{1}{2} (d_j - y_j)^2 \quad (1.5)$$

где  $d_j$  – желаемое, а  $y_j$  – фактическое значение выхода  $j$ -го нейрона в соответствии с поданным  $m$ -ым входным вектором  $X_m$  из пары  $(X, D)_m$ ,  $k$  – количество выходов (классов) персептрона.

$$E_{общ} = E_{общ} + \frac{E}{q}, \quad (1.6)$$

где  $q$  – количество обучающих векторов.

Шаг 7. Посчитать величину коррекции синаптических весов  $j$ -го нейрона и

.....

нейронных смещений и запомнить их (накопление изменений):

$$\Delta w_{i,j}(t+1) = \Delta w_{i,j}(t) - \eta \delta_j x_i, \quad (1.7)$$

где  $\eta$  – коэффициент скорости обучения.

$$\delta_j = -(d_j - y_j) \cdot f'(net_j) \quad (1.8)$$

Шаги 4-7 выполняются последовательно для каждого входного образа, на котором обучается персептрон.

Шаг 8. После подачи последнего обучающего вектора (завершения эпохи), проверить критерий останова обучения, если он выполняется, то завершить обучение. В противном случае – выполнить шаг 9, после чего вернуться к шагу 3.

Шаг 9. Произвести коррекцию синаптических весов  $j$ -го нейрона и нейронных смещений по сохраненным (накопленным) значениям  $\Delta w_{i,j}$ :

$$w_{i,j}(t+1) = w_{i,j}(t) + \Delta w_{i,j}(t+1) \quad (1.9)$$

где  $t$  – номер итерации.

Шаги 5 и 9 повторяются для всех нейронов персептронного слоя при подаче конкретного образа.



Критерии останова алгоритма обучения могут быть следующими:

- 1) Значение общей ошибки  $E_{общ}$ , вычисляемое на основании всех  $q$  обучающих векторов

$$E_{общ} = \sum_{m=1}^q \frac{1}{q} \sum_{j=1}^k \frac{1}{2} (d_j - y_j)^2 = \frac{1}{2q} \sum_{m=1}^q \sum_{j=1}^k (d_j - y_j)^2 \quad (1.10)$$

не превышает заданного заранее установленного порогового значения  $E_{порог}$ , близкого к нулю:  $E_{общ} < E_{порог}$ .

- 2) Превышен установленный лимит количества эпох – самый широко используемый критерий останова, но необходимо знать достаточное количество эпох для обучения, что возможно путем экспериментов.
- 3) Значение  $E_{общ}$  меняется незначительно на протяжении нескольких эпох.