

Обучение искусственных нейронных сетей.

Мищенко В.А.

Воронежский государственный педагогический университет

Одним из видов искусственных нейронных сетей является персептрон- однослойная нейронная сеть, которая в качестве активационной функции использует пороговую функцию. Обученный персептрон способен разделить множество входных образов на два класса: к первому будут относиться образы, для которых на выходе нейронной сети будет ноль, ко второму – образы с выходом персептрона равным единице. Рассмотрим подробнее процедуру обучения персептрона, состоящего из одного нейрона.

Процесс обучения такой нейронной сети состоит в подстройке весовых коэффициентов w_{*i} , где $i=0,n$, последовательно уменьшающей выходные ошибки. Для этого необходимо использовать обучение с учителем, т.е. для успешного обучения должен существовать набор входных образов $X_{k*} = (x_{k_1}, x_{k_2}, \dots, x_{k_n})$, где $k=1,p$, для которых заранее известно к какому из классов принадлежит образ. Персептрон будем считать обученным, если для каждого вектора X на выходе получается соответствующее значение y равное 0 или 1..

В алгоритме обучения можно выделить пять последовательных шагов:

- 1) Проинициализировать весовые коэффициенты w_{*i} $i=0,n$ небольшими случайными значениями из диапазона $[-0.3;0.3]$;
- 2) Подать на вход один из обучающих векторов и вычислить для него выход сети
- 3) Если полученный выход верный - перейти к пункту 5, в противном случае необходимо вычислить ошибку, допущенную сетью по формуле:

$$b = y_{k*} - y$$

- 4) Изменить весовые коэффициенты по следующему правилу:

$$w_{t+1_i} = w_{t_i} + \eta * b * x_{k_i}$$

Здесь t и $t+1$ - номера соответственно текущей и следующей итераций; η - коэффициент скорости обучения сети.

- 5) Шаги 2-4 повторяются для всех обучающих векторов до тех пор, пока сеть не перестает ошибаться.

Стоит отметить, что коэффициент скорости обучения, как правило, выбирают из диапазона $[0,5;0,7]$, однако, в некоторых случаях, например, при достаточно большой обучающей выборке возможно постепенное уменьшение данного коэффициента, начиная с единицы.

Таким образом, если объединить несколько персептронов в однослойную нейронную сеть, то полученная сеть сможет вести разделение входных образов на произвольное число групп.

Алгоритм обучения такой сети, аналогичен описанному выше алгоритму: необходимо произвести итерационную подстройку матрицы весов W , постепенно уменьшая ошибку в векторах на выходе.

Примечание. Запись вида z_a_b понимать как: z – наименование переменной или вектора, a – надстрочный знак, b – подстрочный знак. Символ $*$ - понимать как отсутствие одного из знаков.