

НЕЙРОННЫЕ СЕТИ. ОБУЧЕНИЕ ПЕРСЕПТРОНА.

Персептроны являются основной элементарной базой современных нейронных сетей. Исторически их появление относят к 1957 году, когда Р. Розенблатт разработал модель, которая вызвала большой интерес у исследователей. Несмотря на некоторые ограничения её исходной формы, она стала основой для многих современных, наиболее сложных алгоритмов обучения с учителем.

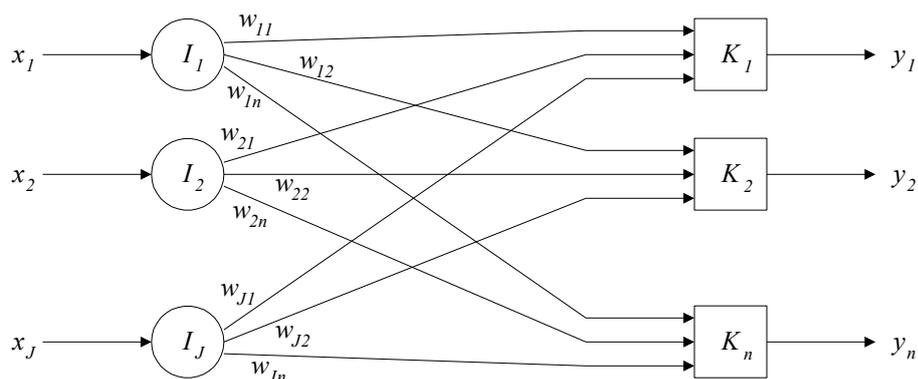


Рисунок 1

Персептрон является двухуровневой нерекуррентной сетью, вид которой показан на рисунке 1. Она использует алгоритм обучения с учителем. Обучающая выборка состоит из множества входных векторов, для каждого из которых указан свой требуемый вектор цели. Компоненты входного вектора X представлены непрерывным диапазоном значений. Компоненты вектора цели Y являются двоичными величинами (0 или 1). После обучения сеть получает на входе набор непрерывных входов и вырабатывает требуемый выход в виде вектора с бинарными компонентами.

Обучение осуществляется следующим образом:

1. Все веса сети случайно выбираются как малые величины;
2. На вход сети подаётся входной обучающий вектор X и вычисляется сигнал NET от каждого нейрона, используя стандартное выражение

$$NET_j = \sum_i x_i w_{ij}$$

3. Вычисляется значение пороговой функции активации для сигнала NET от каждого нейрона следующим образом:

$$OUT_j = \begin{cases} 1, & \text{если } NET_j > \theta_j \\ 0, & \text{в противном случае} \end{cases}$$

Здесь θ_j представляет собой порог, соответствующий нейрону j (в простейшем случае все нейроны имеют один и тот же порог).

- Вычисляется ошибка для каждого нейрона посредством вычитания полученного выхода из требуемого выхода:

$$error_j = target_j - OUT_j$$

- Каждый вес модифицируется следующим образом:

$$w_{ij}(t+1) = w_{ij}(t) + \alpha \cdot x_i \cdot error_j.$$

- Шаги 2-5 повторяются, пока ошибка не станет достаточно малой.

Рассмотрим обсуждаемый метод для решения задачи классификации. Пусть требуется построить сеть для решения задачи о принадлежности точки к тому или иному квадранту декартовой плоскости (от 1 до 4). Вход сети – координаты x и y точки, выход – вектор из 4-х элементов, причем элемент, соответствующий найденному квадранту точки равен 1, остальные равны 0. Таким образом, сеть будет состоять из 2 входов и 4 выходов. Требуется настроить веса сети так, чтобы она могла корректно распознавать принадлежность точки к тому или иному квадранту.

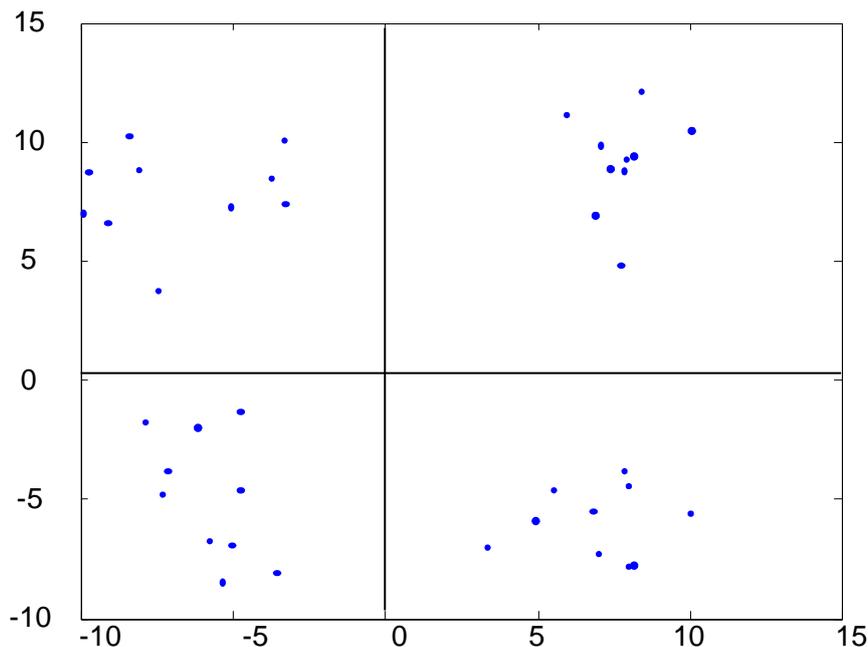


Рисунок 2.

На рисунке 2 показано расположение 40 случайно выбранных точек – по 10 в каждом квадранте. Это будет обучающей выборкой. Точки «вразнобой» проходят через процедуру обучения 1-5. Коэффициент α выбран равным 1. В результате, весовые коэффициенты приобретают следующие значения

$$\begin{aligned}w_{11} &= 11.7035 & w_{12} &= 5.8339 \\w_{21} &= -4.7798 & w_{22} &= -3.4843 \\w_{31} &= -3.5149 & w_{32} &= -11.5192\end{aligned}$$

$$w_{41} = 8.4805 \quad w_{42} = -10.1324$$

После этого сеть становится пригодной к работе над задачей по распознаванию принадлежности точки к квадранту.

ЗАДАНИЕ К ЛАБОРАТОРНОЙ РАБОТЕ

Задать случайно значения точек в пространстве – по 10 в каждой из восьми областей, образуемых осями координат. Обучить нейронную сеть типа персептрон, состоящую из 3 входов и 8 выходов так, чтобы она могла распознавать в какой из 8 областей пространства находится точка (аналогично рассмотренному примеру для плоскости).

