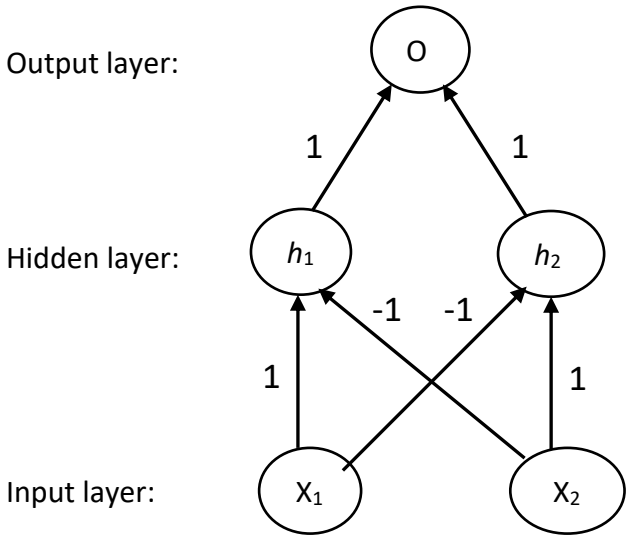

repeat
  for each e **in** *examples* **do**
    for each node j in the input layer **do** $a_j \leftarrow x_j[e]$
    for $\ell = 2$ **to** M **do**
      $in_i \leftarrow \sum_j W_{j,i}\ a_j$
      $a_i \leftarrow g(in_i)$
    for each node i in the output layer **do**
      $\Delta_i \leftarrow g'(in_i) \times (y_i[e] - a_i)$
    for $\ell = M - 1$ **to** 1 **do**
      for each node j in layer ℓ **do**
        $\Delta_j \leftarrow g'(in_j) \sum_i W_{j,i}\ \Delta_i$
        for each node i in layer $\ell + 1$ **do**
          $W_{j,i} \leftarrow W_{j,i} + \alpha \times a_j \times \Delta_i$
until some stopping criterion is satisfied

Sigmoid activation function
Learning rate 0.7
Trace the algorithm.
https://www.tinkershop.net/ml/sigmoid_calculator.html

XOR FFNet

X1	X2	y	h_1	h_2	h'_1	h'_2	O	O'	Δo	Δh1	Δh2	w'_{h1-O}	w'_{h2-O}
0	**0**	0											
0	**1**	1											
1	**0**	1											
1	**1**	0											

X1	X2	w_{h1-O}	w_{h2-O}	w'_{x1-h1}	w'_{x1-h2}	w_{x1-h1}	w_{x1-h2}	w'_{x2-h1}	w'_{x2-h2}	w_{x2-h1}	w_{x2-h2}
0	**0**										
0	**1**										
1	**0**										
1	**1**										