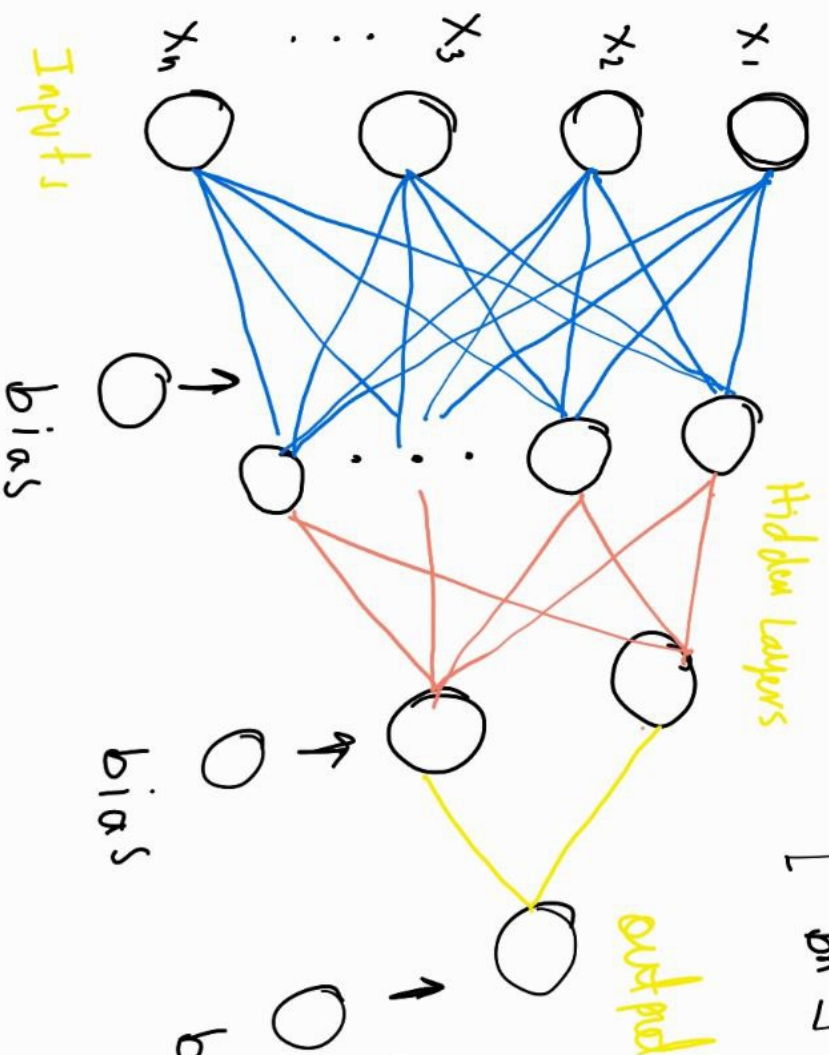


$$f\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} w_{11} & w_{21} & \dots & w_{n1} \\ w_{12} & \cdot & \cdot & \cdot \\ \vdots & \vdots & \cdot & \vdots \\ w_{1n} & w_{2n} & \dots & w_{nn} \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}\right) = f(x^T w + b)$$



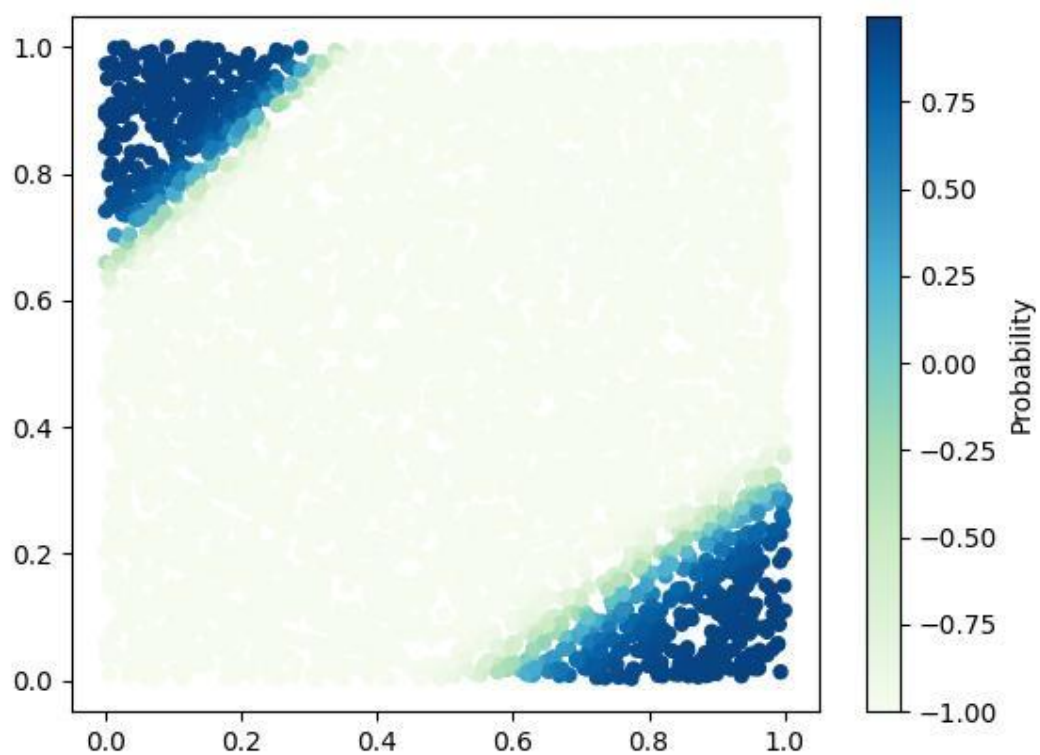
Los pesos y bias de la red se actualizan con un error proporcional al cuadrado de la diferencia entre el resultado deseado y el resultado obtenido.

↑ algoritmo del

○ backpropagation y

bias gradient descent.

XOR



AND

