



Trabajo Práctico 2

Aprendizaje Profundo
Alejo Ordoñez

4 de diciembre de 2024

Resumen

En este trabajo se abordan el diseño, implementación y análisis de varias arquitecturas de redes neuronales. En particular, se cubren un perceptrón simple, un perceptrón multicapa, una máquina restringida de Boltzmann, una red convolucional y un auto-encoder. Se cubre el algoritmo de aprendizaje gradient descent, usando backpropagation para el cálculo de los gradientes, y se usa el algoritmo de preentrenamiento introducido en el trabajo de Hinton y Salakhutdinov (2006) para la de la máquina de Boltzmann.

1. Perceptrón simple

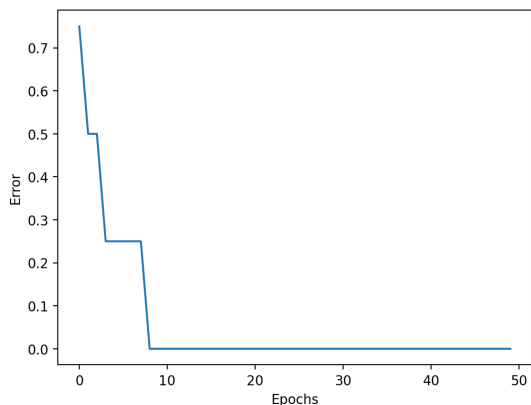
El perceptrón simple es la unidad básica de una red neuronal. Es una neurona artificial que usa una función de activación escalón. La función mapea la entrada \mathbf{x} (un vector de números reales) a una salida binaria $f(\mathbf{x})$:

$$f(\mathbf{x}) = h(\mathbf{w} \cdot \mathbf{x} + b),$$

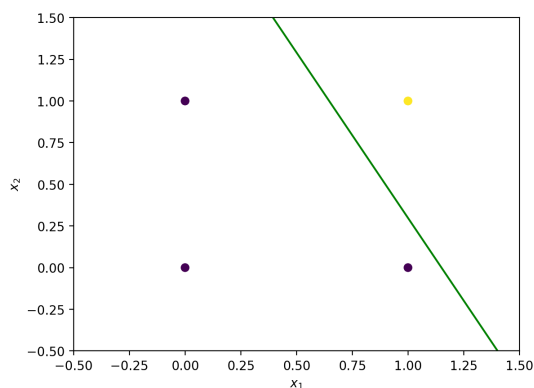
donde h es la función escalón, \mathbf{w} es el vector de pesos, \mathbf{x} es el vector de entradas y b es el sesgo. El sesgo genera una traslación del límite de decisión, que es el hiperplano definido por $\mathbf{w} \cdot \mathbf{x} + b = 0$. El entrenamiento de un perceptrón simple, consiste en la actualizar los pesos y el sesgo de acuerdo a

$$w_i = \alpha(y - \hat{y})x_i, \quad b = \alpha - (y - \hat{y}),$$

donde α es la tasa de aprendizaje, el par x, y es un ejemplo de entrenamiento y \hat{y} es la salida del perceptrón. En primer lugar se implementó un perceptrón simple que aprendió la función lógica AND de dos entradas. La tabla de verdad de la función AND es: La evolución del error durante el entrenamiento y la recta discriminadora aprendida por el perceptrón, $x_2 = -(w_1x_1 + b)/w_2$, se muestran a continuación:

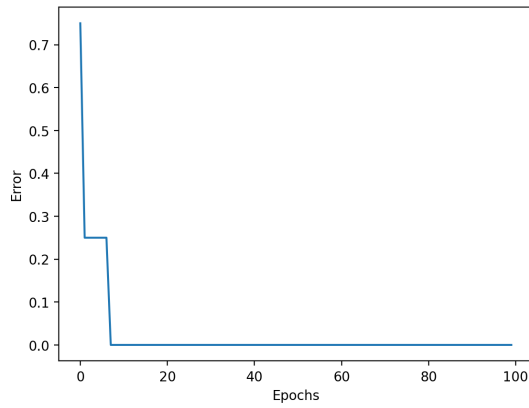


Evolución del error

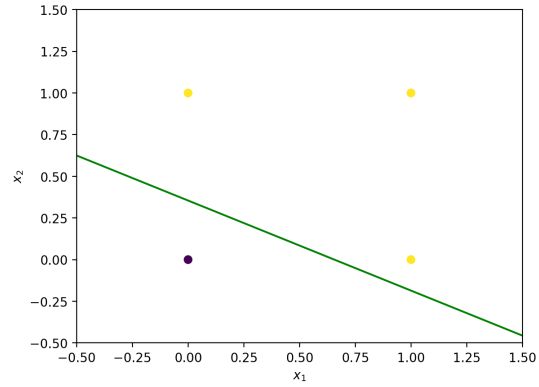


Recta discriminadora

Seguidamente, se entrenaron perceptrones para que aprendieran las funciones lógicas OR de dos entradas, y AND y OR de cuatro entradas. Los resultados se muestran a continuación:

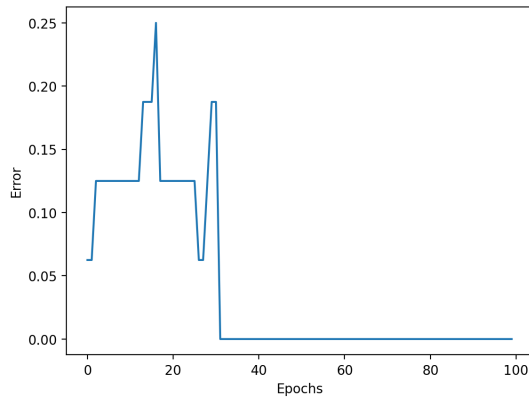


Evolución del error

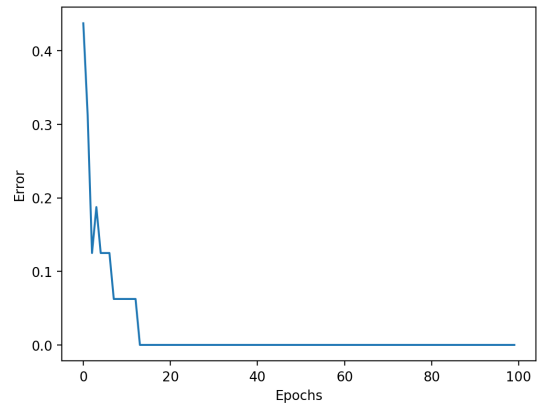


Recta discriminadora

Figura 2: OR de dos entradas



AND de cuatro entradas



OR de cuatro entradas

Figura 3: Evolución del error

Referencias

- [1] Ian Goodfellow. Yoshua Bengio. Aaron Courville, *Deep Learning*.
- [2] Michael Nielsen, *Neural Networks and Deep Learning*.
- [3] G. E. Hinton. R. R. Salakhutdinov, *Reducing the Dimensionality of Data with Neural Networks*.
- [4] Yann LeCun. Patrick Haffner. Léon Bottou. Yoshua Bengio, *Object Recognition with Gradient-Based Learning*.