
Redes Neuronales

Abel José Sánchez Alba y Arturo Cortés
Sánchez



1. Índice

- Historia de las redes neuronales
- Estructura de la neurona
- Perceptron
- Estructura de la red neuronal
- Función de error
- Backward propagation
- Otras redes
- Demostración

Historia

- 1949 - Donald Hebb crea la hipótesis del aprendizaje basado en plasticidad neuronal.
- 1958 - Frank Rosenblatt crea el perceptrón.
- 1962 - Novikoff comprueba que el perceptrón encuentra la solución si los datos son linealmente separables.
- 1965 - Ivakhnenko y Lapa publican la primera red neuronal multicapa.

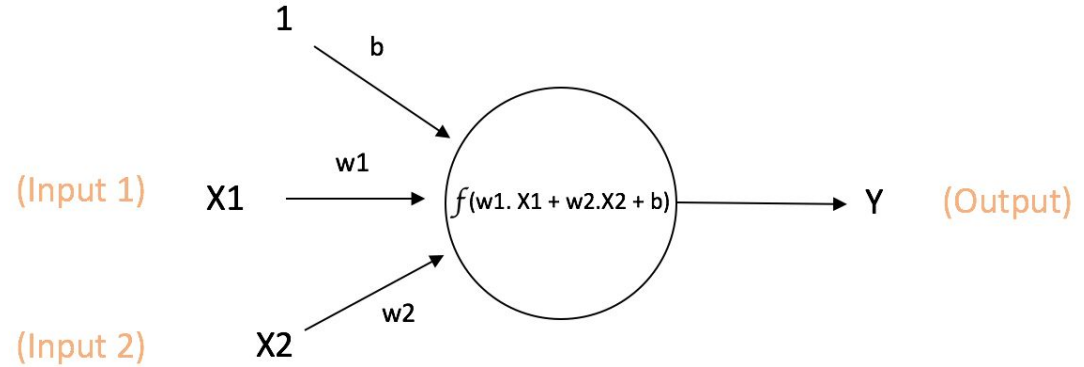
Historia

- 1969 - Minsky y Papert descubren que el perceptrón no puede clasificar la función XOR.
- 1975 - Paul John Werbos inventa el algoritmo de backpropagation.
- 1992 - Se inventa el algoritmo de max-pooling.

—

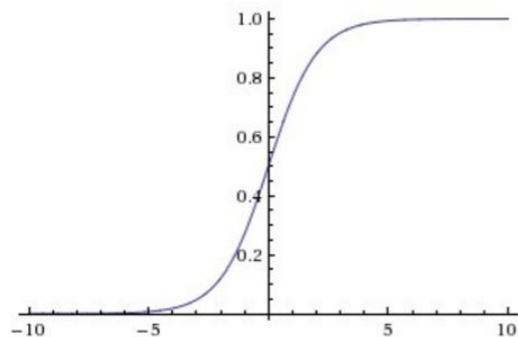
Estructura de una neurona

- Conjunto de entradas
- Parámetro bias
- Pesos
- Función de activación



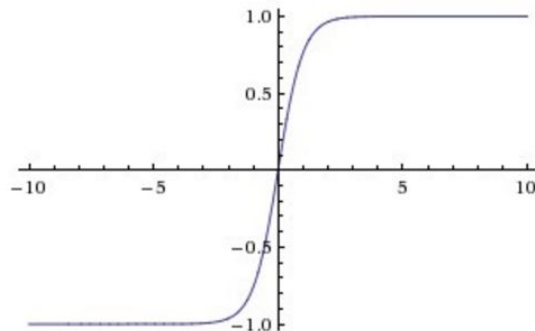
$$\text{Output of neuron} = Y = f(w1 \cdot X1 + w2 \cdot X2 + b)$$

Funciones de activación



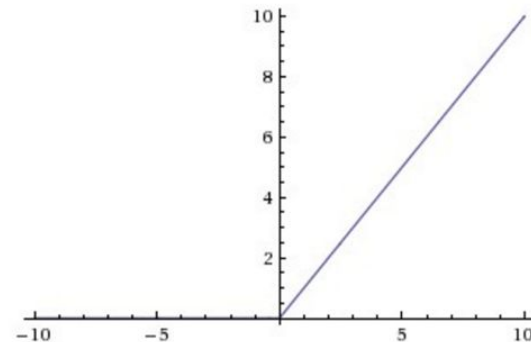
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



tanh

$$\tanh(x)$$



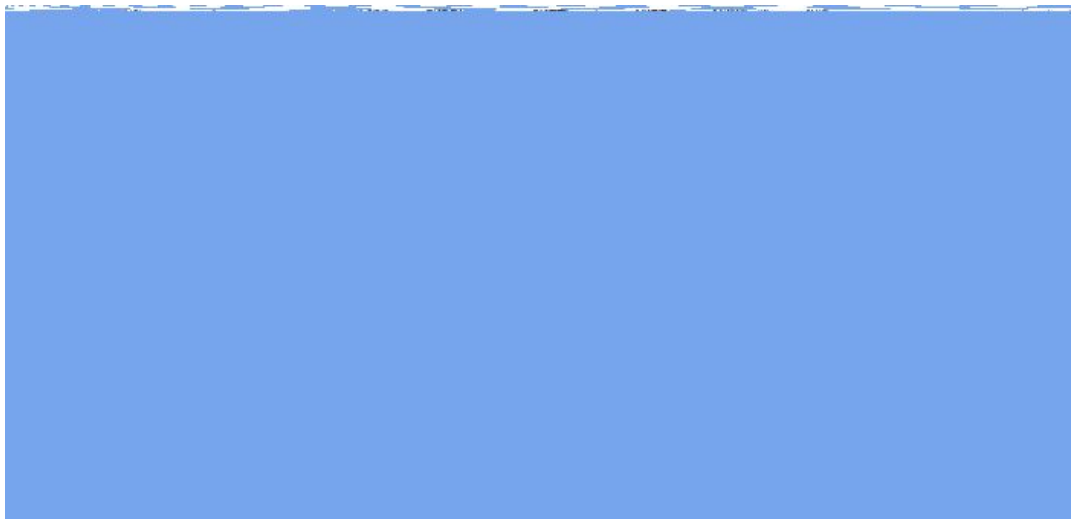
ReLU

$$\max(0, x)$$

Aprendizaje

Para cada elemento del conjunto de entrenamiento:

- Si la salida predecida es correcta, los pesos se dejan intactos
- Si se predice incorrectamente un 0, se le suma la entrada a los pesos.
- Si se predice incorrectamente un 1, se le resta la entrada a los pesos.

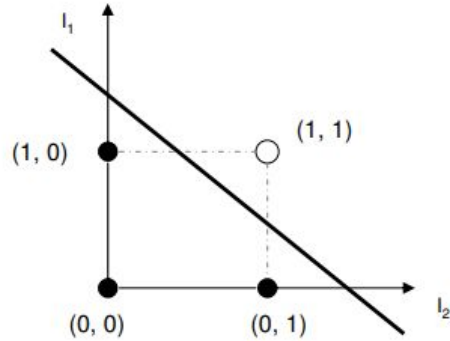


Limitaciones

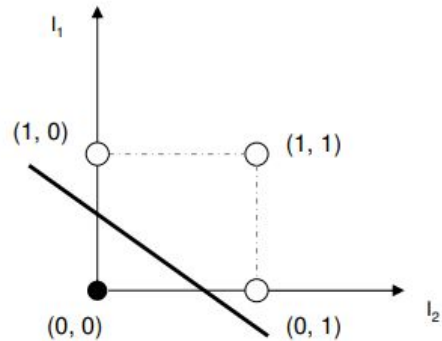


Limitaciones

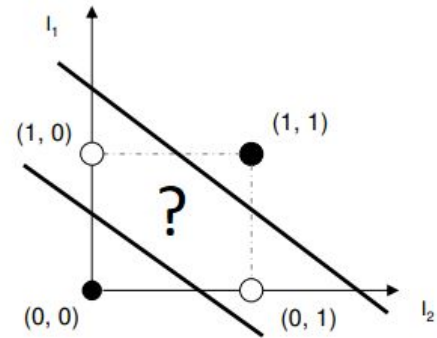
AND		
I_1	I_2	out
0	0	0
0	1	0
1	0	0
1	1	1



OR		
I_1	I_2	out
0	0	0
0	1	1
1	0	1
1	1	1

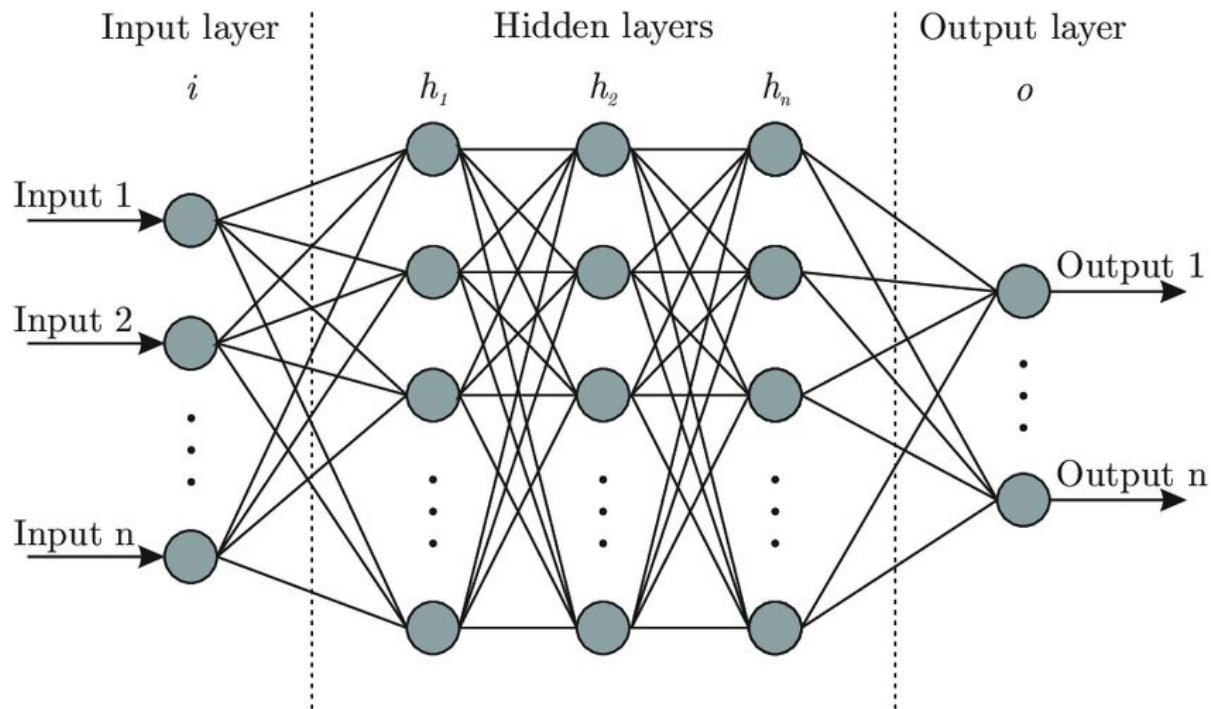


XOR		
I_1	I_2	out
0	0	0
0	1	1
1	0	1
1	1	0



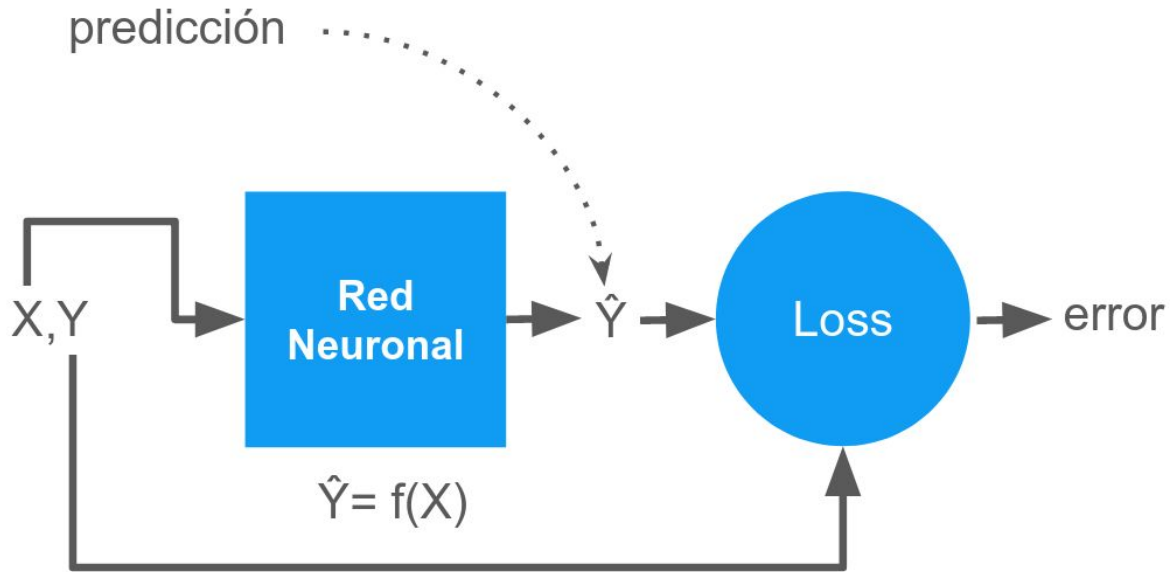
Estructura de una red neuronal

- Neuronas de entrada
- Neuronas ocultas
- Neuronas de salida



La función de error

- Indica la diferencia entre la predicción y el valor real de la variable.
- Es la función a minimizar
- Su elección depende del problema



Funciones de error y problemas

- Problema de regresión

- Error cuadrático medio

$$\frac{1}{n} \sum_i (\hat{Y}_i - Y_i)^2$$

- Clasificación Binaria

- Entropía cruzada o error

logarítmico

$$- \sum_{c=1}^M Y_{o,c} \log(\hat{Y}_{o,c})$$

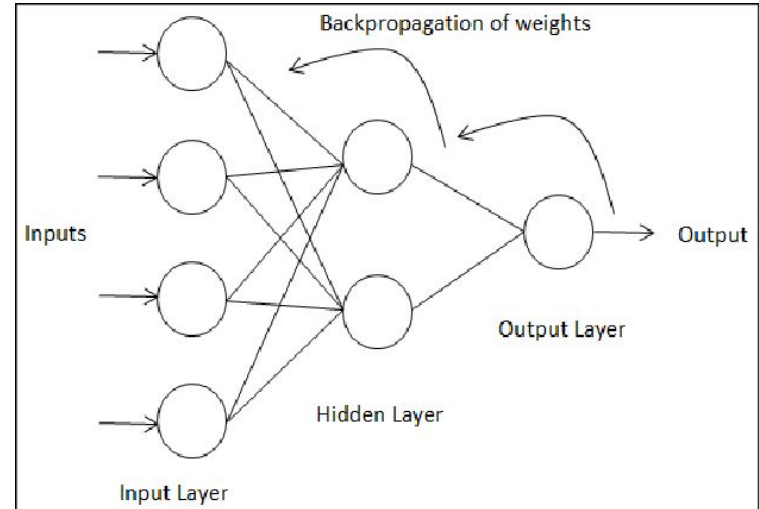
- Clasificación Multilabel

- Entropía cruzada o error

logarítmico

Back propagation

- Propagación del error hacia atrás
- Depende de la función de error
- Esencial para el descenso del gradiente



¿Cómo se calcula?

- En la última capa:

$$\delta^l = \frac{\partial C}{\partial a^l} \frac{\partial a^l}{\partial z^l}$$

- En el resto de capas:

$$\delta^{l-1} = W^l \delta^l \frac{\partial a^{l-1}}{\partial z^{l-1}}$$

—

Descenso del gradiente

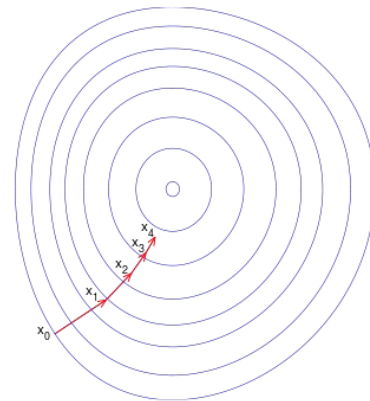
- Learning rate

- Actualización de sesgos: $\frac{\partial C}{\partial b^{l-1}} = \delta^{l-1} \quad b^l = b^l - \text{media}(\text{Errores}) * lr$

- Actualización de pesos: $\frac{\partial C}{\partial W^{l-1}} = \delta^{l-1} a^{l-2} \quad W^l = W^l - [W^l * a^{l-1} * lr]$

Descenso del gradiente

- Repetidos pasos en dirección opuesta al gradiente



- Si la función es convexa, se puede alcanzar un mínimo global



A mostly complete chart of Neural Networks

©2019 Fjodor van Veen & Stefan Leijnen asimovinstitute.org

-  Input Cell
-  Backfed Input Cell
-  Noisy Input Cell
-  Hidden Cell
-  Probabilistic Hidden Cell
-  Spiking Hidden Cell
-  Capsule Cell
-  Output Cell
-  Match Input Output Cell
-  Recurrent Cell
-  Memory Cell
-  Gated Memory Cell
-  Kernel
-  Convolution or Pool

Perceptron (P)



Feed Forward (FF)



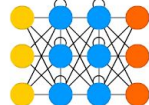
Radial Basis Network (RBF)



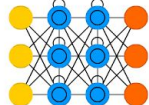
Deep Feed Forward (DFF)



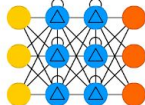
Recurrent Neural Network (RNN)



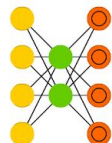
Long / Short Term Memory (LSTM)



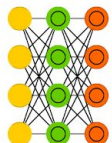
Gated Recurrent Unit (GRU)



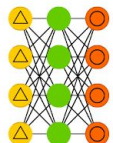
Auto Encoder (AE)



Variational AE (VAE)



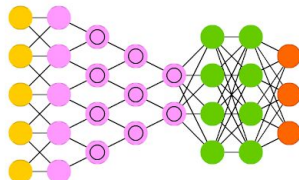
Denoising AE (DAE)



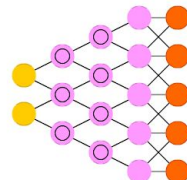
Sparse AE (SAE)



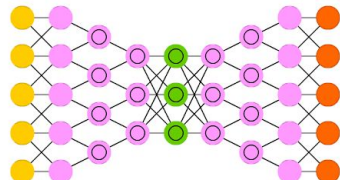
Deep Convolutional Network (DCN)



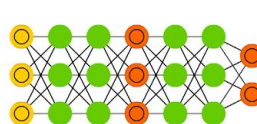
Deconvolutional Network (DN)



Deep Convolutional Inverse Graphics Network (DCIGN)



Generative Adversarial Network (GAN)



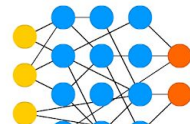
Liquid State Machine (LSM)



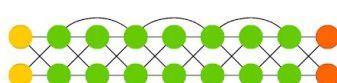
Extreme Learning Machine (ELM)



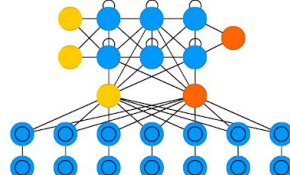
Echo State Network (ESN)



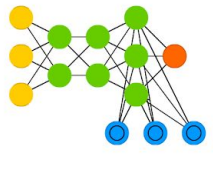
Deep Residual Network (DRN)



Differentiable Neural Computer (DNC)



Neural Turing Machine (NTM)



Markov Chain (MC)



Hopfield Network (HN)



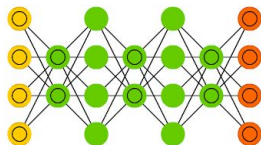
Boltzmann Machine (BM)



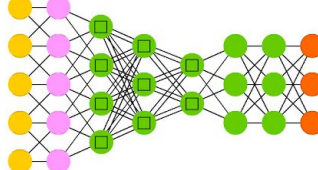
Restricted BM (RBM)



Deep Belief Network (DBN)



Capsule Network (CN)



Kohonen Network (KN)



Attention Network (AN)

