

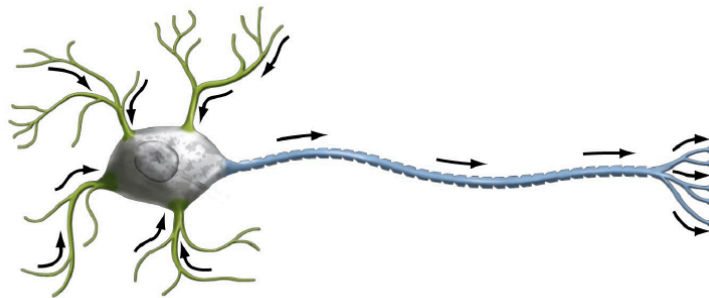
Aprendizaje Automático 2

César Martínez
30/11/2024

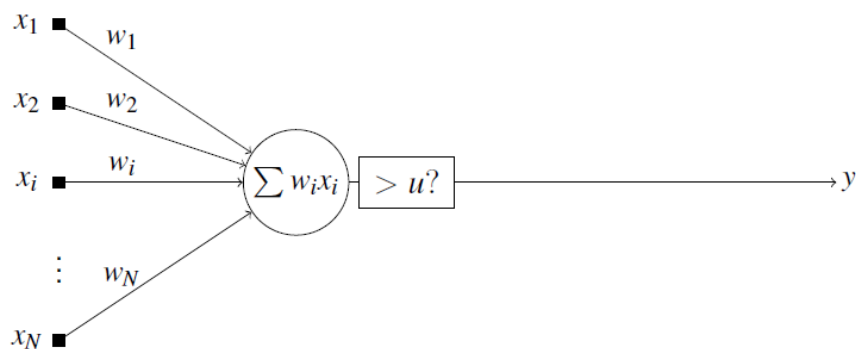
Redes neuronales tipo Perceptrón

Introducción

Las redes neuronales artificiales son un modelo computacional inspirado en la neurona biológica (imagen inferior), que es el fundamento de la estructura y funcionamiento del cerebro humano.



Aquí podemos observar con flechas el flujo de información: reciben influencia de otras neuronas, tienen un cuerpo que realiza una integración de estas entradas y produce una despolarización dependiendo del valor que tome esa integración. La despolarización se modela con una función no lineal que envía un pulso de información hacia otras neuronas posteriores en esta gran red neuronal. El modelo resultante es:



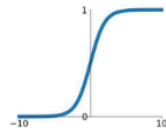
Puesto en fórmulas matemáticas, la salida se modela como:

$$y = \phi(v - u) = \phi \left(\sum_{i=1}^N w_i x_i - u \right)$$

donde $\phi(z)$ es la función de activación no lineal que puede tomar diferentes formas, como se muestra a continuación:

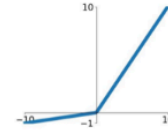
Sigmoid

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



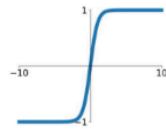
Leaky ReLU

$$\max(0.1x, x)$$



tanh

$$\tanh(x)$$

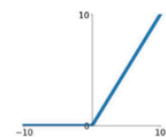


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

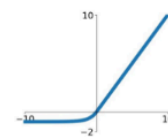
ReLU

$$\max(0, x)$$



ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$

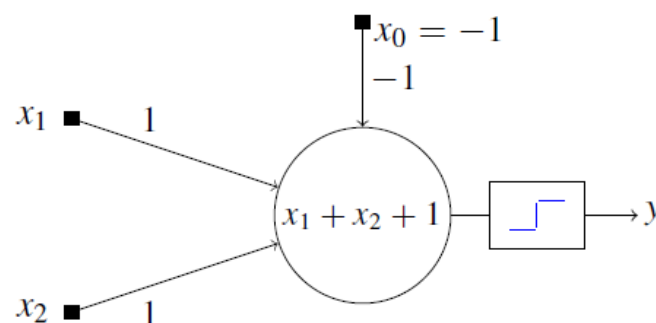


Las neuronas artificiales luego pueden ser organizadas en capas, para continuar con la idea de modelado biológico. Su principal objetivo es aprender patrones y relaciones complejas a partir de datos, lo que las hace fundamentales en tareas de clasificación, regresión y reconocimiento de patrones.

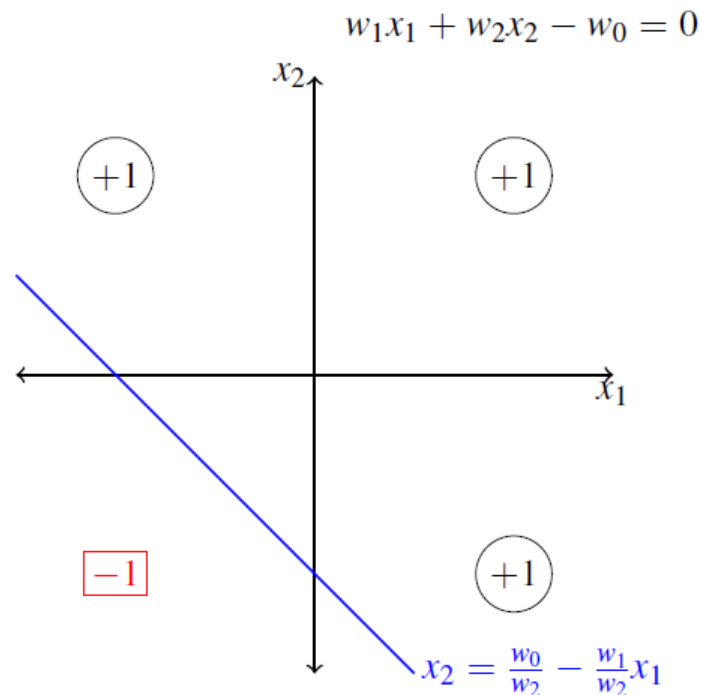
El Perceptrón es la forma más simple de una red neuronal, desarrollada por Frank Rosenblatt en 1958. Representa el punto de partida en la evolución de las redes neuronales y sienta las bases para modelos más complejos, como las redes neuronales multicapa (MLP).

Estructura del Perceptrón

El Perceptrón consiste en una sola capa de neuronas de salida conectadas a una capa de entrada. Cada entrada está asociada a un *peso* (w), que representa la importancia de esa entrada en la decisión final. Además, se añade un término adicional llamado *sesgo* (x_0), que permite ajustar la salida del modelo independientemente de las entradas. La fórmula básica del Perceptrón con 2 entradas es:

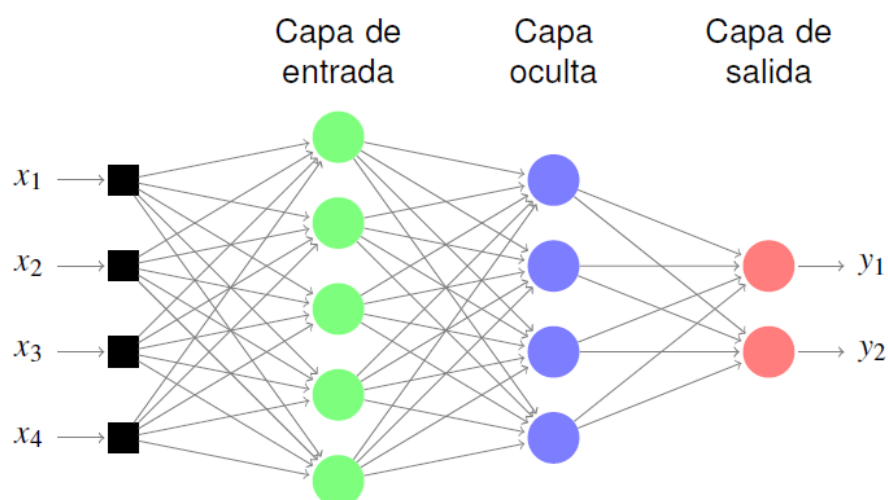


Este tipo de redes puede resolver problemas linealmente separables, donde la frontera de decisión define las regiones donde la salida tiene uno de los 2 valores posibles luego de la función de activación no lineal, así:



Redes multicapa

El Perceptrón Multicapa (MLP, por sus siglas en inglés) es una extensión del Perceptrón simple, diseñada para abordar problemas de clasificación más complejos, especialmente aquellos que no son linealmente separables. A diferencia del Perceptrón de una sola capa, el MLP está compuesto por múltiples capas de neuronas: una capa de entrada, una o más capas ocultas, y una capa de salida. Cada capa está completamente conectada a la siguiente, y cada neurona de una capa aplica una función de activación no lineal, como la sigmoide, ReLU, o tangente hiperbólica.



Esta estructura permite que el MLP capture relaciones no lineales en los datos, lo que lo hace adecuado para tareas complejas como la clasificación de imágenes, reconocimiento de voz y predicción de series temporales. A continuación se observan las regiones de decisión que pueden lograrse con este tipo de arquitecturas:

<i>Estructura</i>	<i>Tipos de regiones de decisión</i>	<i>Problema XOR</i>	<i>Separación en clases</i>	<i>Formas regiones más generales</i>
<i>Una capa</i> 	<i>hemiplano limitado por hiperplano</i>			
<i>Dos capas</i> 	<i>Regiones convexas abiertas o cerradas</i>			
<i>Tres capas</i> 	Arbitrarias (Complejidad limitada por N°. de Nodos)			

Entrenamiento del Perceptrón

El proceso de entrenamiento consiste en ajustar los pesos y el sesgo para que el Perceptrón pueda clasificar correctamente las entradas. Se utiliza un algoritmo de aprendizaje supervisado conocido como *Backpropagation*. El entrenamiento se realiza en varias iteraciones o épocas, donde el Perceptrón ajusta sus parámetros para minimizar el error en las predicciones.

Básicamente, el algoritmo utiliza el cálculo de una función de costo para realizar un cambio iterativo de pesos en función del gradiente descendente en la curva de error. Este método permite, luego de cierta cantidad de épocas, encontrar un mínimo local o -idealmente-

global en la superficie de error, cuyas variables de entrada representan los valores óptimos de los pesos buscados.

La combinación de múltiples capas y funciones de activación no lineales permite que el MLP aprenda representaciones internas de los datos, lo que resulta en un modelo mucho más potente que el Perceptrón simple. Sin embargo, su entrenamiento puede ser computacionalmente intensivo y requiere un ajuste cuidadoso de hiperparámetros, como la tasa de aprendizaje y el número de neuronas en cada capa.