

¿Qué es una Red Neuronal? – DotCSV

Parte 1: La Neurona

Una neurona es la capacidad básica de procesamiento dentro de una red neuronal. Tiene conexiones de entrada a través de las que recibe estímulos externos (valores de entrada), con los cuales la neurona realizará un cálculo interno (similar a una regresión lineal) y generará un valor de salida. Internamente, utiliza todos los valores de entrada para realizar una suma ponderada de ellos de según el peso (parámetro) dado a cada conexión de entrada, el cual indica con qué intensidad dicha variable afectará a la neurona.

Cada neurona funciona para puertas AND y OR, para XOR, se necesitan dos neuronas, debido a que con una sola neurona no se puede separar a un conjunto de puntos linealmente distribuido de dicha manera.

Parte 2: La Red

Podemos juntar neuronas en la misma capa, teniendo así capas de entrada, capas ocultas y capas de salida. Cuando se colocan las neuronas de manera secuencial, una de ellas recibe la información procesada por la anterior lo que permite que la red pueda aprender conocimiento jerarquizado, aprendiendo conocimiento más básico en cada capa y haciéndolo cada vez más complejo según la cantidad de capas (Deep Learning).

La suma de regresiones lineales nos da como resultado una línea recta, haciendo que el sistema colapse y que todas las operaciones se resuman a una única recta. Para evitar que el sistema colapse es necesario manipular cada una de las regresiones lineales resultantes a un modelo no lineal, lo que hace necesaria a la función de activación. Al evaluar el resultado de la regresión lineal en la función de activación se distorsiona el valor de salida, añadiéndole deformaciones no lineales, logrando así encadenar de forma efectiva la computación de varias neuronas.

Algunas funciones de activación:

- Escalonada: No es muy útil para el aprendizaje de las neuronas.
- Sigmoide: Ideal para probabilidades, los valores muy grandes se saturan en 1 y los muy pequeños, en 0.
- tanh: Su rango varía de -1 a 1.
- Unidad rectificada lineal (relu): Se comporta como una función lineal cuando el valor de entrada es positivo y constante a cero cuando el valor de entrada es negativo.

El efecto de la función de activación es el de distorsionar el plano generado por la neurona. Al juntar neuronas, se juntan las distorsiones del plano, dando como resultado un nuevo plano distorsionado, así, las redes neuronales son capaces de desarrollar soluciones muy complejas.

Parte 3: Backpropagation

Algoritmo de backpropagation: Permite obtener los errores presentes en cada neurona del modelo con su respectiva ponderación (responsabilidad) con un único pase hacia atrás. Dichos errores ayudan a calcular las derivadas parciales de cada parámetro de la red, conformando un vector gradiente que se utiliza en el algoritmo de descenso del gradiente para minimizar el error y entrenar a la red.

Las redes neuronales se asemejan a una cadena de responsabilidades, donde cada nodo es una neurona especializada en una tarea determinada. El análisis de cuánta responsabilidad tiene cada neurona en el resultado obtenido se realiza hacia atrás, desde la señal de error hasta las primeras capas, considerando que el error de la capa final depende de los errores de las capas previas (como una epidemia). Si el error no depende de una capa, tampoco dependerá de las anteriores a ella, en caso contrario, resulta necesario analizar las capas anteriores para identificar el origen del error (retropropagación de errores).