

Práctica 6 - Jose Andrés Mejías Rojas

Activation Functions

Los activation functions definen cómo el weighth sum de los inputs son transformados de nodo a nodo.

Las redes neuronales tienen 3 capas: input, hidden, y output. Generalmente, las capas hidden generalmente tienen el mismo activation function. Por otro lado, la capa output tiene un activation function completamente distinto.

Activation functions para capas hidden

- Rectified Linear Activation (ReLU): $\max(0.0, x)$
- Logistic (Sigmoid): $1.0 / (1.0 + e^{-x})$
- Hyperbolic Tangent (Tanh): $(e^x - e^{-x}) / (e^x + e^{-x})$

Activation functions para capas output

- Linear: x (mismo resultado)
- Logistic (Sigmoid): $1.0 / (1.0 + e^{-x})$
- Softmax: $e^x / \sum(e^x)$ (x es un vector)

Comentario

Lo único relevante que quiero agregar de este artículo es que por fin entendí la dinámica de los activation functions. Honestamente no me quedaba tan claro cuál era su propósito. Sin embargo, como el autor dio ejemplos, fue más claro. Algo tan sencillo como explicar el $\max(0.0, x)$ me facilita entender los otros más complejos.

Una duda que tengo es que no entiendo muy bien la diferencia entre Sigmoid y Tanh (más allá de la fórmula obvio).