General AI/ML

Unit 1: Intro to AI, ML and DL

1.2.3

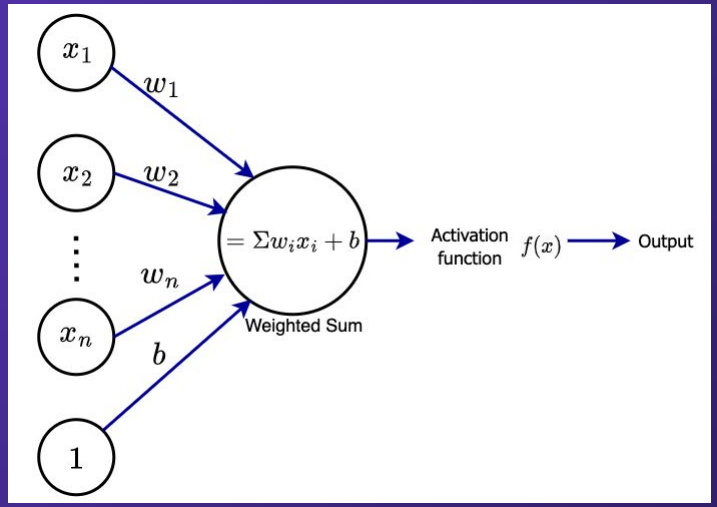
Introduction to Deep Learning

Activation Functions

TIL-AI
TODAY I LEARNED-AI

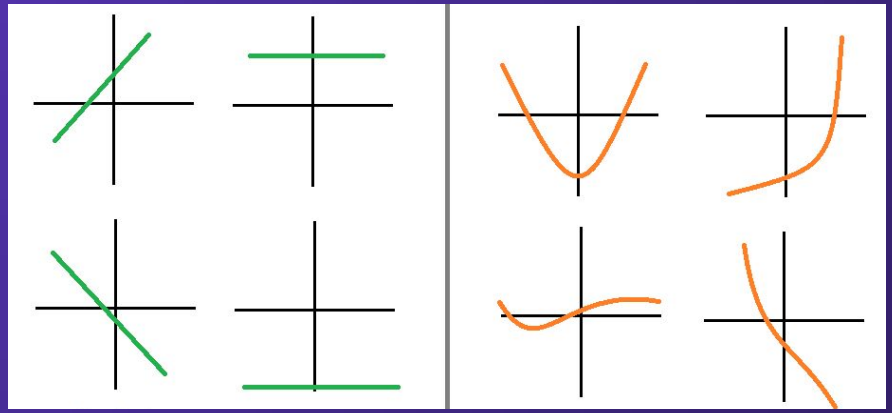
What are Activation Functions?

- Activation functions are mathematical functions applied to the output of a neuron (or node) in a neural network.
- They determine whether a neuron should 'fire' (activate), and to what degree, based on the weighted sum of its inputs.
- This non-linear behavior empowers neural networks to learn complex patterns that would be impossible with linear models.



Why Non-linearity Matters

- Most real-world data and relationships do not follow simple linear patterns.
- Linear models are restricted in their ability to learn complex patterns.
- Non-linear activation functions allow neural networks to approximate nearly any function, making them universal approximators.

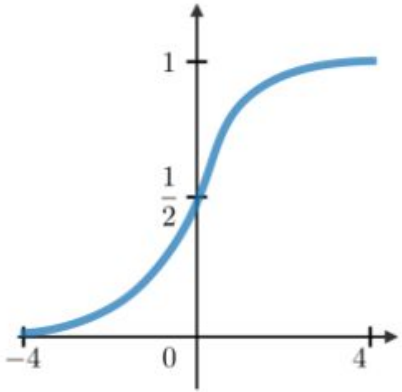
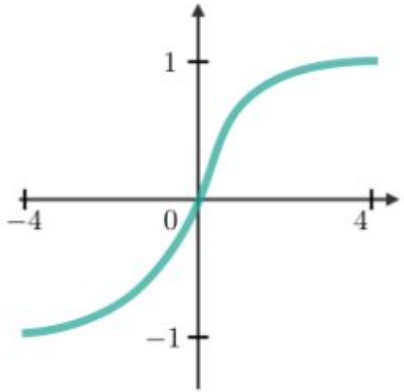
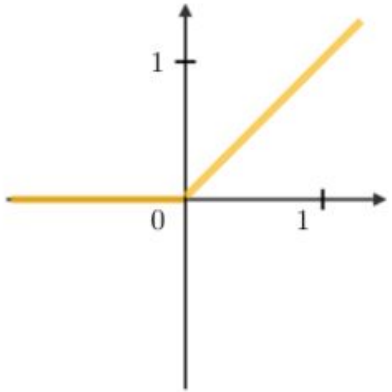
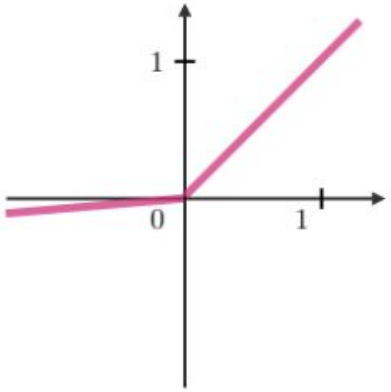


Linear Functions

Non-linear Functions

TIL-AI
TODAY I LEARNED-AI

Common Activation Functions

Sigmoid	Tanh	ReLU	Leaky ReLU
$g(z) = \frac{1}{1+e^{-z}}$	$g(z) = \frac{e^{z}-e^{-z}}{e^{z}+e^{-z}}$	$g(z) = \max(0, z)$	$g(z) = \max(\epsilon z, z)$ with $\epsilon \ll 1$
			

Choosing an Activation Function

- Task: The type of problem (classification, regression) influences the choice.
- Output Layer: Sigmoid for binary classification, softmax for multi-class, linear for regression
- Hidden Layers: ReLU is a good default, experiment with Leaky ReLU, explore others if needed.
- Experimentation & Research: The field evolves, and best practices change, so stay updated.

TIL-AI
TODAY I LEARNED-AI

Activation Functions