

Neural Networks Explained, really

Roger Lam
lamroger.com

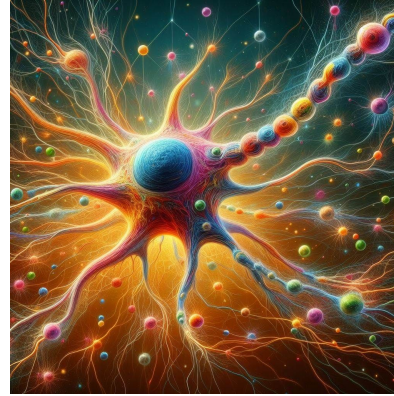
“Neural Network”?

What’s the first thing that comes to mind when you hear the word “neural”?

Maybe “neuron”?

And “network”?

Maybe “a connection of things”?

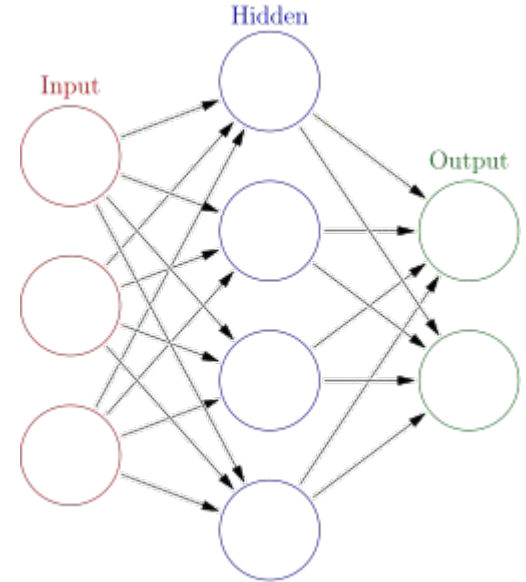
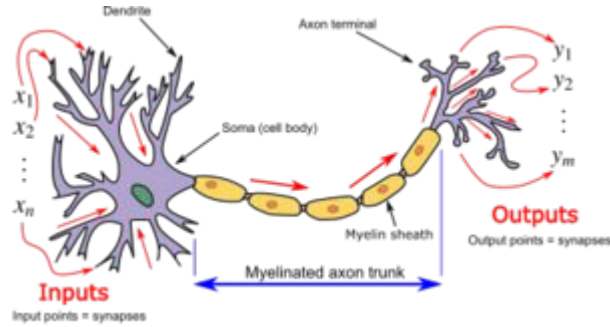


Neural Networks are inspired by neurons!

Neurons take input and give outputs.

Similarly, Neural networks are like multiple neurons.

They take input, pass it through to other nodes, and generates an output.

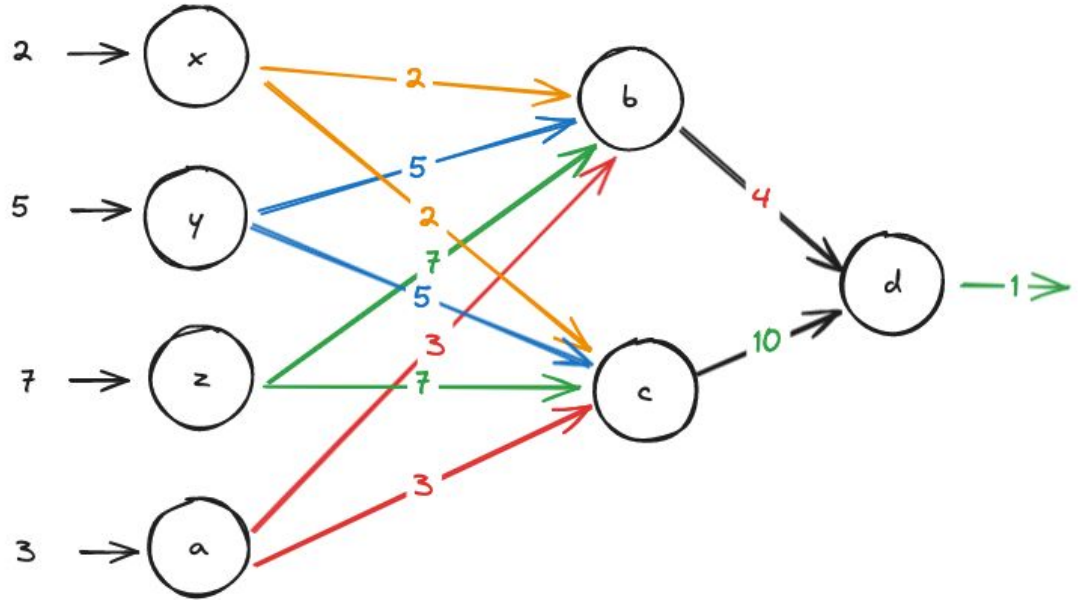


Ok, but what does it look like?

In this simplified example, we have $[2, 5, 7, 3]$ as our input.

Each node is an equation that generates a result.

We keep passing results through and end up with the value 1.

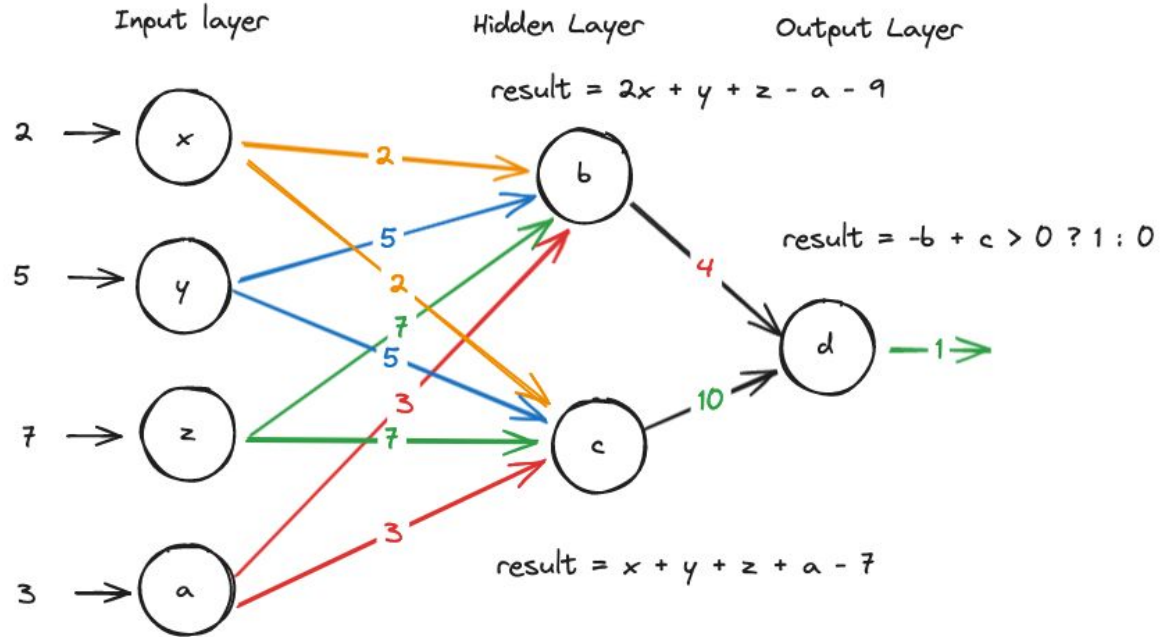


Each node is an equation

They do actual calculations.

The last node does binary classification.

If $-b + c$ is > 0 , return true, else return false.

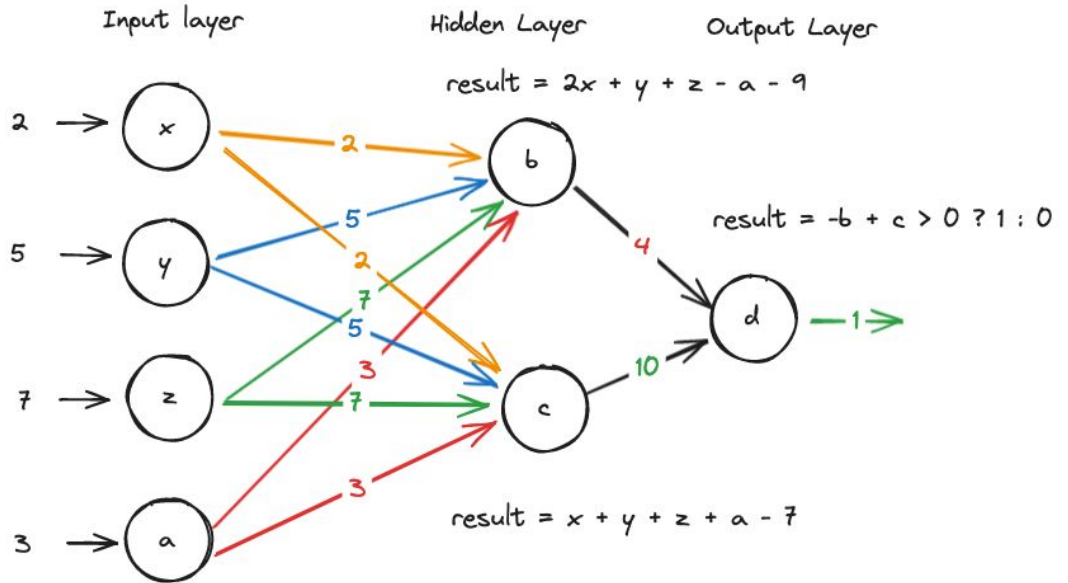


Lots of equations which can be represented by matrices

By using matrix multiplication, we can speed up calculations.

GPUs, or graphics cards, are used to do efficient matrix operations to render graphics on computers.

The same technology is used to do large scale AI / ML calculations.



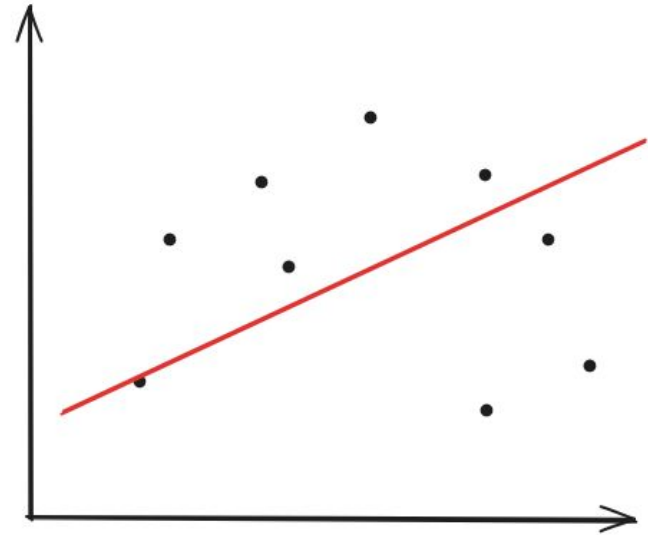
$$[2, 5, 7, 3] \times \begin{bmatrix} 2 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} + [-9, -7] = [4, 10]$$

Lots of Linear Equations but they're kinda boring.

You might have noticed that we only used linear equations. That's an oversimplification.

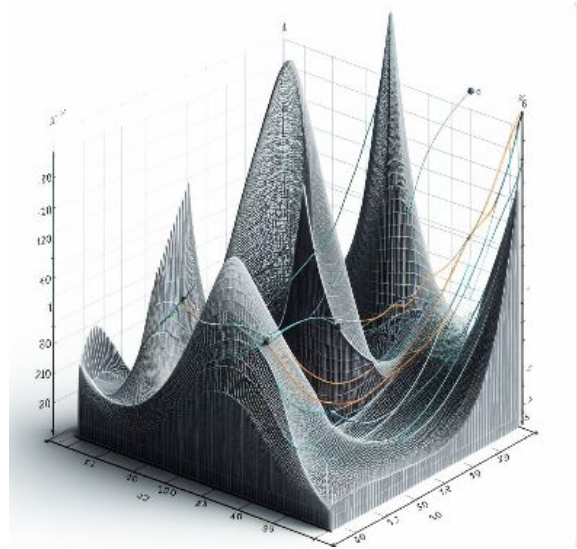
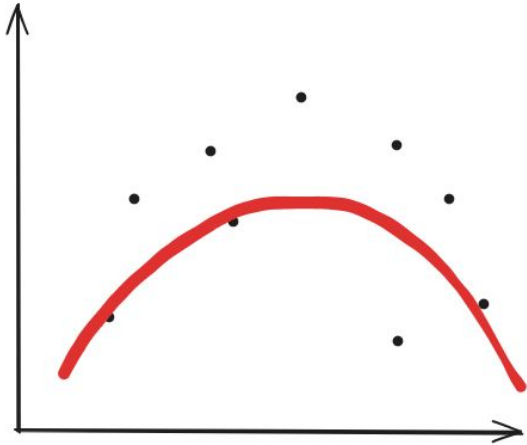
Linear equations aren't great at representing anything other than data in straight lines.

We need more skillz to model our complex world.



We need curves

That's better!

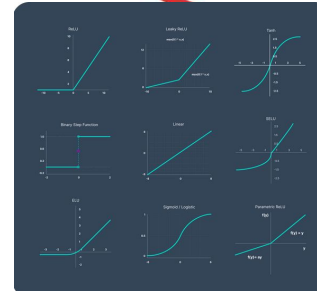
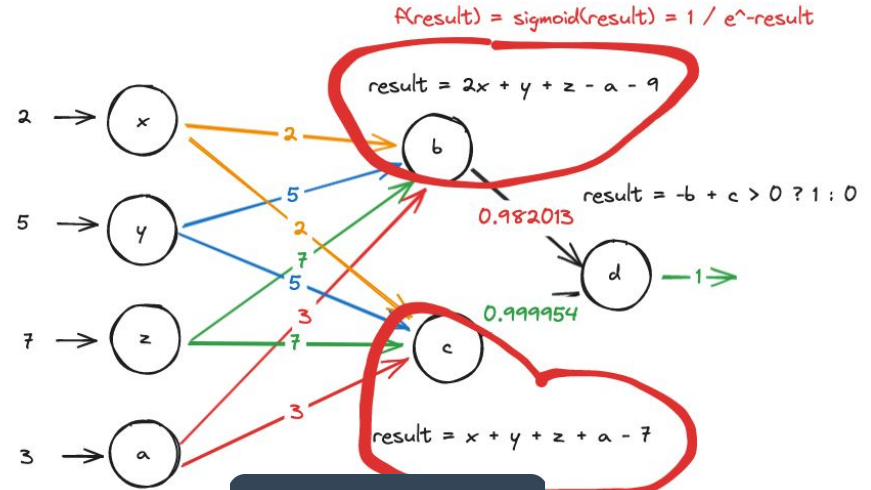


We need to introduce non-linearity

We actually skipped a step in normal neural networks.

Each node usually has a non-linear “activation function”.

We apply it after our linear calculation for each node.



What else does it help with?

Activation functions help with

- stability - squashing values into a contained set
- domain performance - some work better in different types of problems
- backpropagation - I won't go into backpropagation here but we need curves to determine how to "train" or change the equations in our neural network

Now imagine BILLIONS of nodes!

GPT-3 has 175 billion parameters!

All those parameters are used to model the nuances and patterns in human language.



Thanks!

Hope this demystified neural networks a little bit.

If you want to learn more, I highly recommend the [3Blue1Brown Neural Network YouTube course](#). It's 1 hour total split up into 4 chapters.

Roger Lam - lamroger.com

