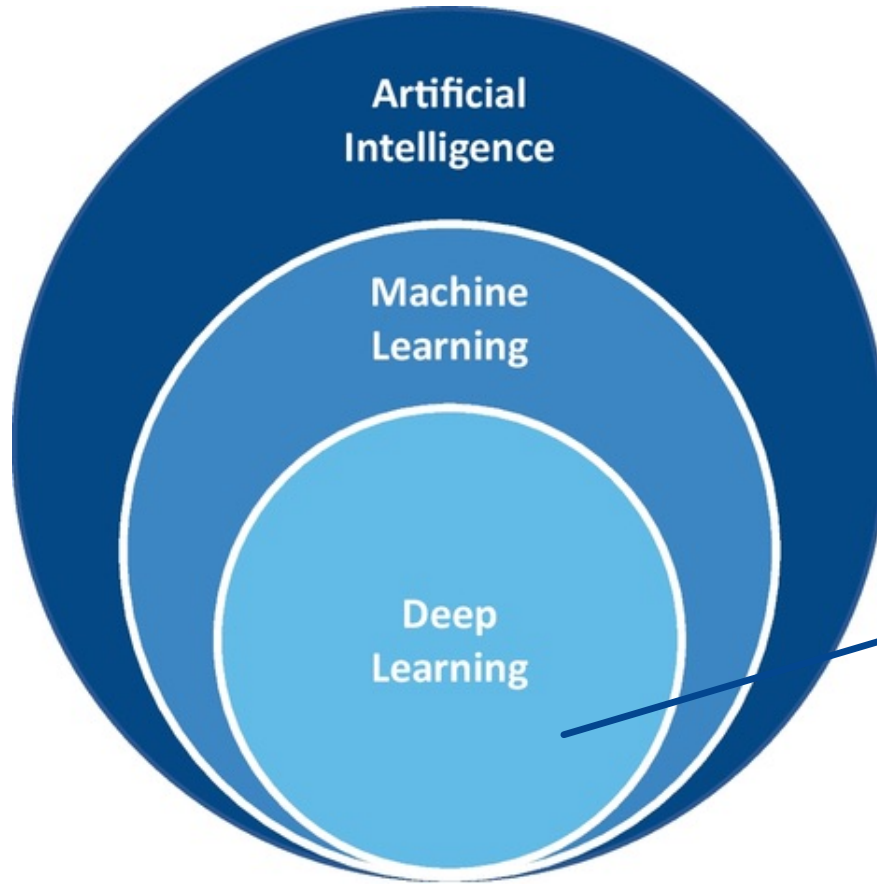


DEEP
LEARNING



done using Neural Networks

NEURAL NETS THROUGH THE YEARS

1942—First computational model for neural networks

1965—First functional networks of many layers

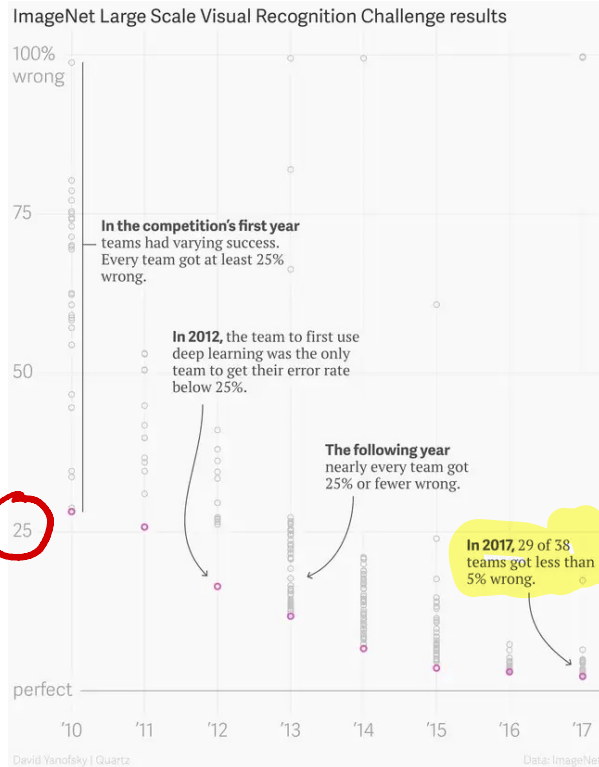
1975—Backpropagation algorithms for training multilayer networks

1990s—Datasets quite small; computers not that fast; other methods doing better

2005-2007—Unsupervised learning with deep nets; use of GPUs

2009—ImageNet: Image database of 14 million images for more than 21000 concepts

2012—AlexNet: Winner of ImageNet Large Scale Visual Recognition Challenge 2012



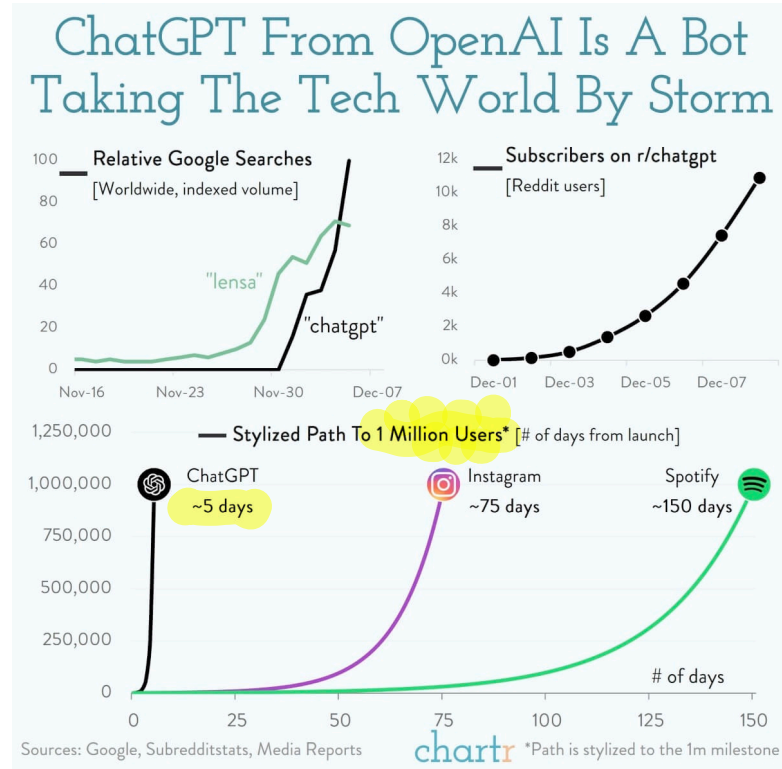
NEURAL NETS THROUGH THE YEARS

2017—A team of Google researchers proposes a new simple network architecture, the Transformer. Transformers enabled advancements in generative models compared to older long short-term memory models

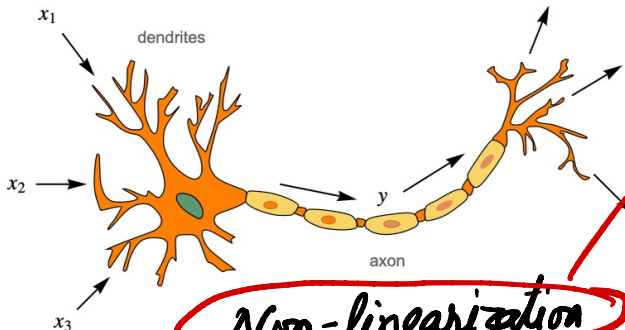
2018—OpenAI releases GPT (Generative Pre-trained Transformer), a language model that achieves state-of-the-art performance on various natural language processing tasks

2021—DALL.E, a transformer-based neural network-based system developed by OpenAI, generates images from textual descriptions

2022—ChatGPT releases GPT-3.5, an AI tool that reached one million users within five days. The tool can access data from the web from up to 2021.

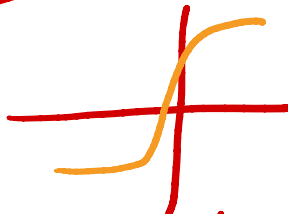
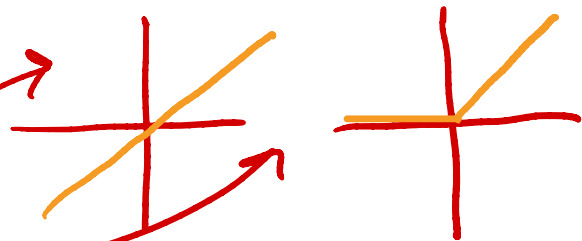


Perceptron



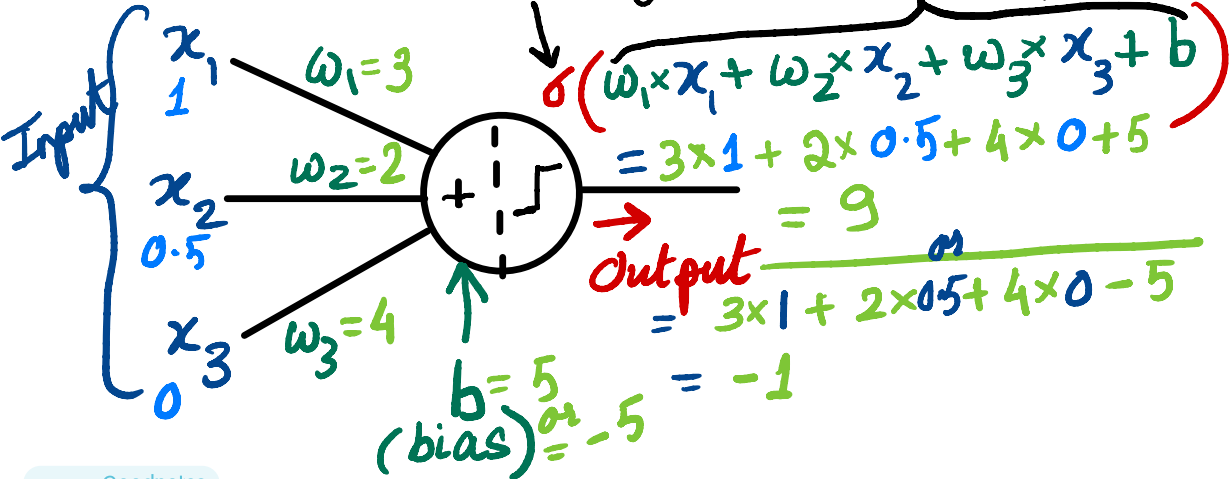
Activation function

linear
step



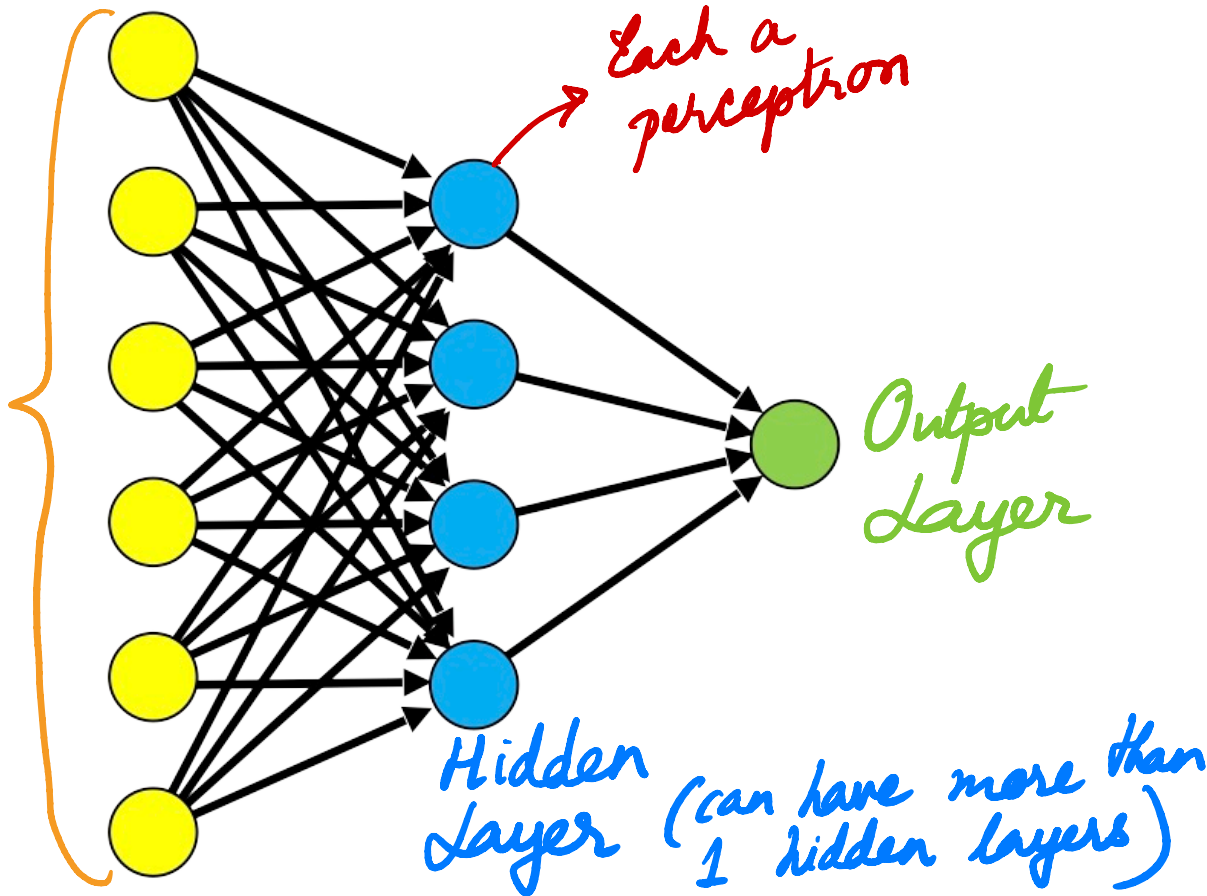
also squishes output to be between -1 and 1

Non-linearization

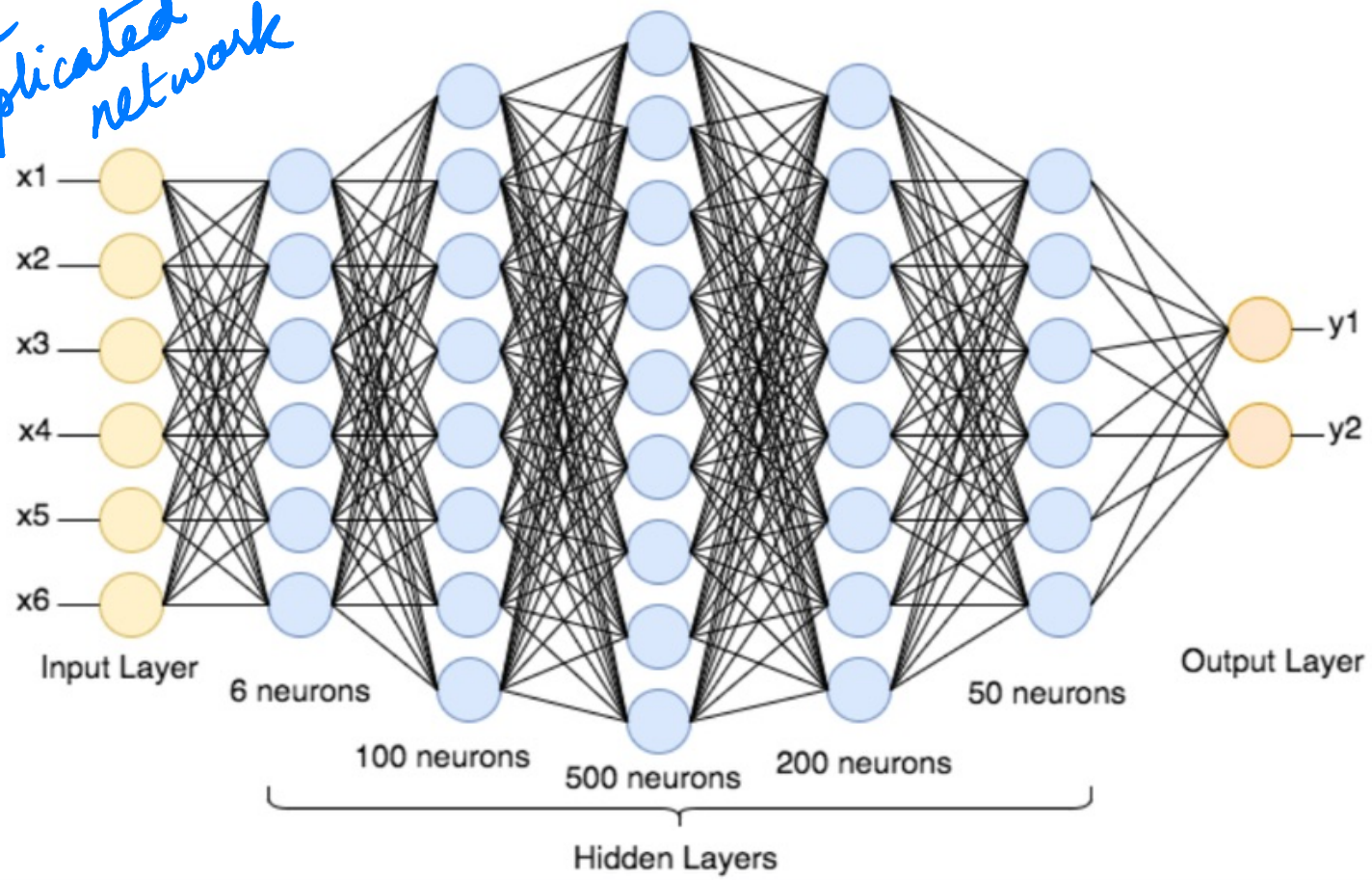


Neural Networks
Consist of Many
Artificial Neurons

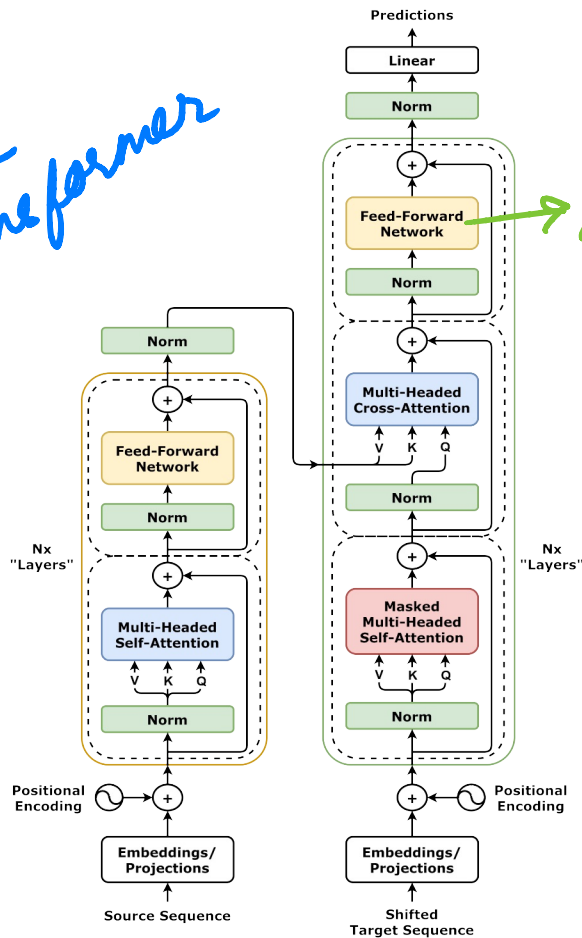
Input
Layer



More complicated network

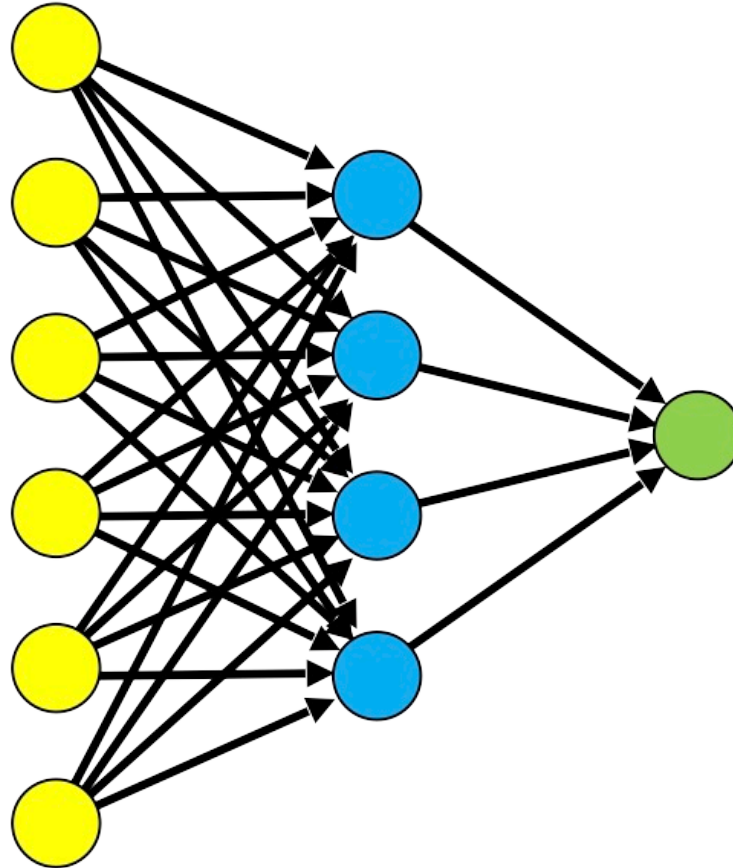


Transformer
architecture
(as in Generative
Pre-trained Transformer
of GPT)



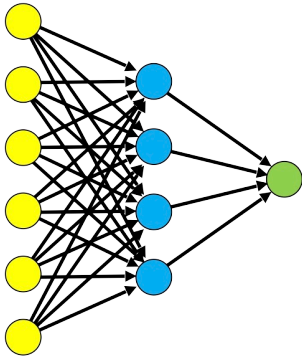
Each a full-fledged network

Neural Networks Consist of Many Artificial Neurons



In theory, a small number of fully connected hidden layers can learn anything.

Neural Networks Consist of Many Artificial Neurons



In theory, a small number of fully connected hidden layers can learn anything.

In practice, three things happened before neural networks enjoyed major success.

By the late 1990s, researchers had built

- **One: new architectures** that **leveraged relationships between the inputs**, and
- **Two: deeper networks** to **capture more complex functions more quickly**.

- **Three:** The last change came in early 2007 with the release of the first easily programmable graphics processing unit (GPU), NVIDIA's GTX80

GPUs had developed

- to meet the entertainment market demand for high-resolution graphics,

- and by 2007 offered much **more raw computational power than processors**.

More recent designs have stopped using human-modeled features, and instead allow training of the neural network to derive the features of importance from the data.

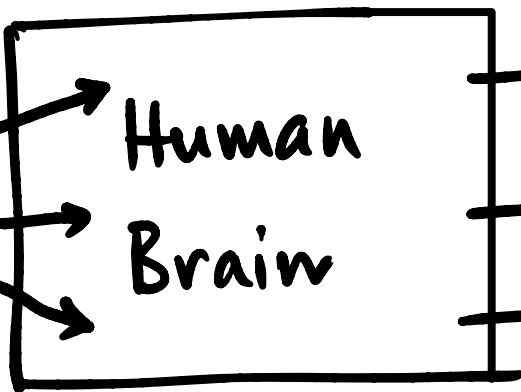
This approach is called deep learning.

It's the number of node layers, or depth, of neural networks that distinguishes a single neural network from a deep learning algorithm.

Deep learning is possible due to the sheer volume of data now available in many problems.

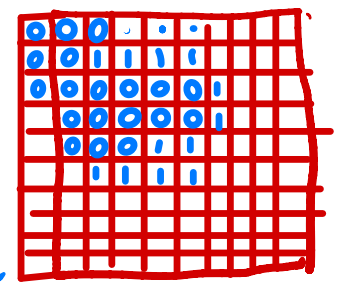
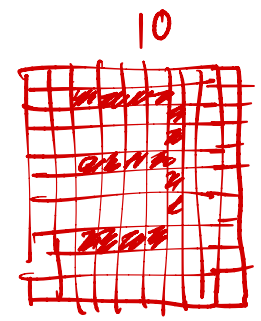
HANDWRITTEN DIGITS RECOGNITION

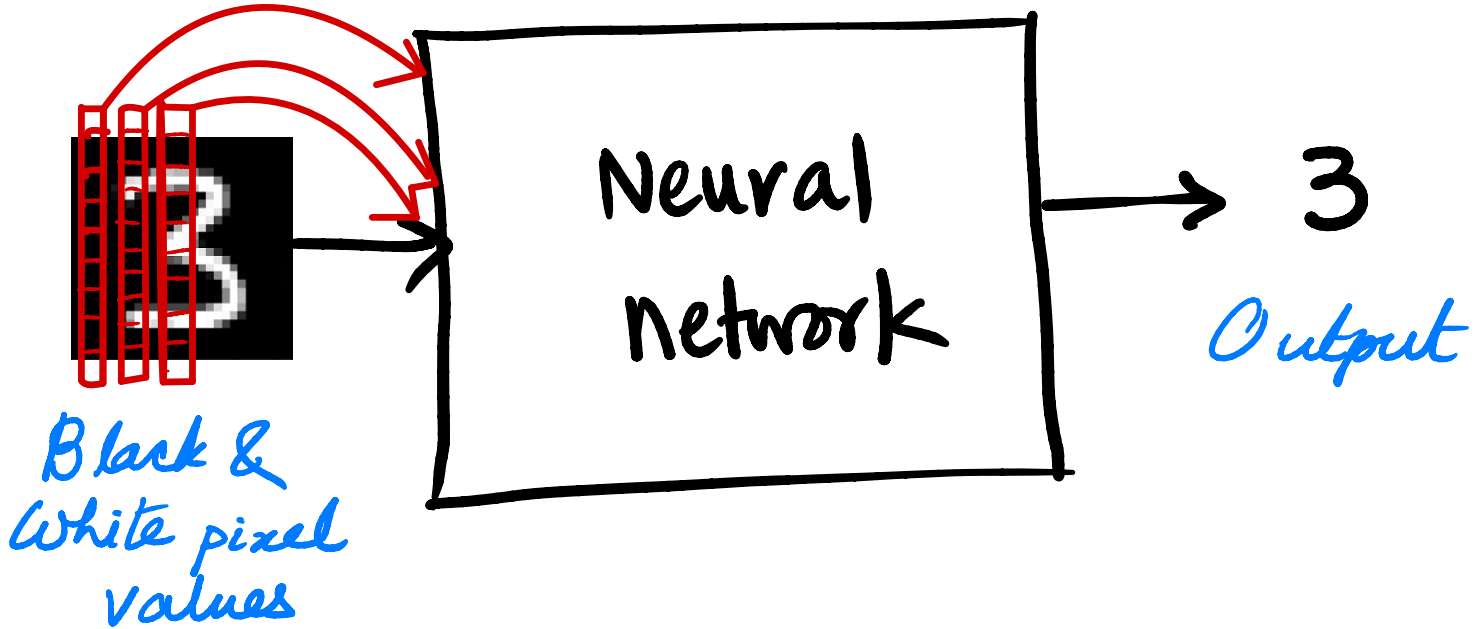
10 pixel by 10 pixel image

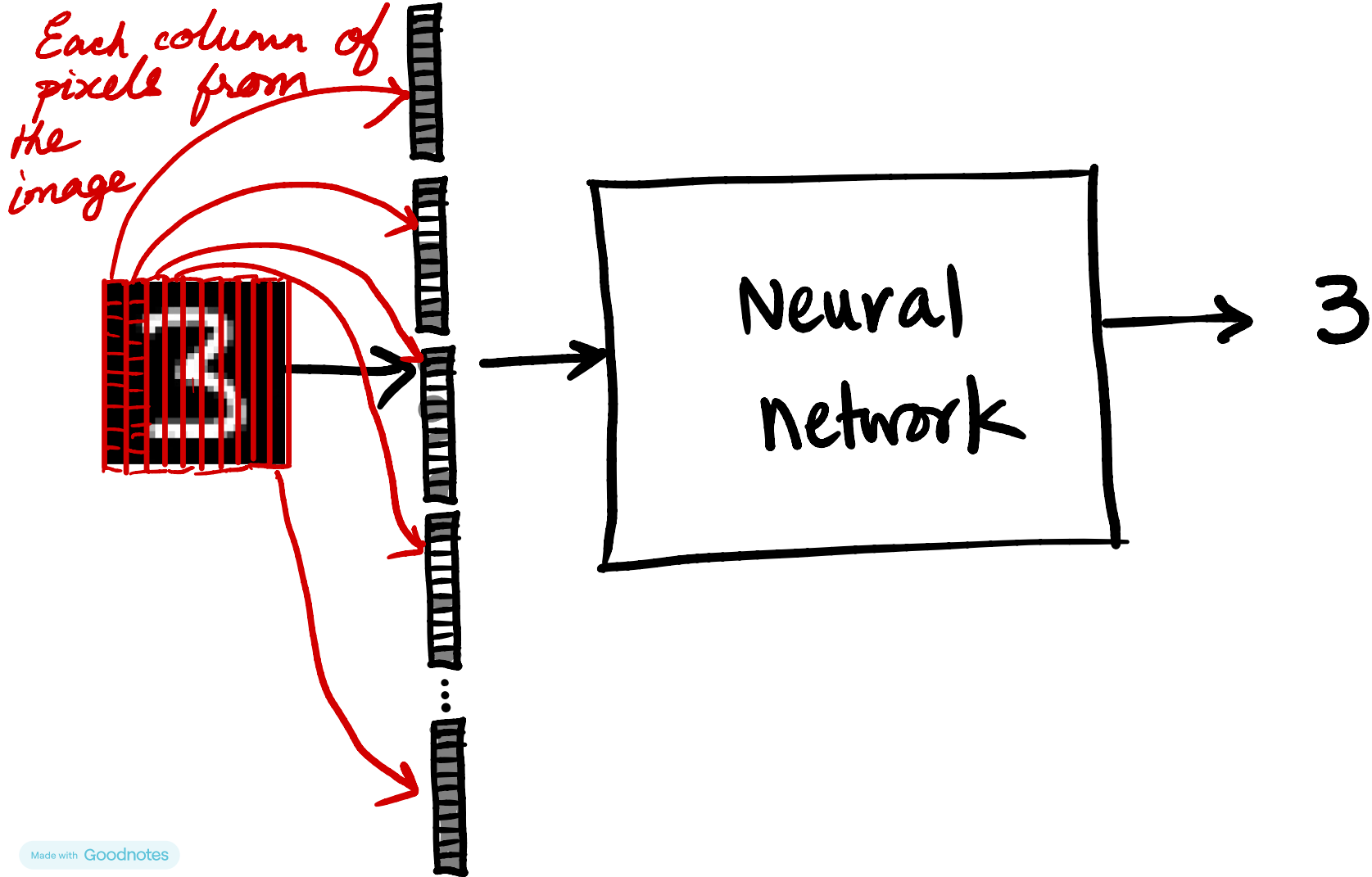


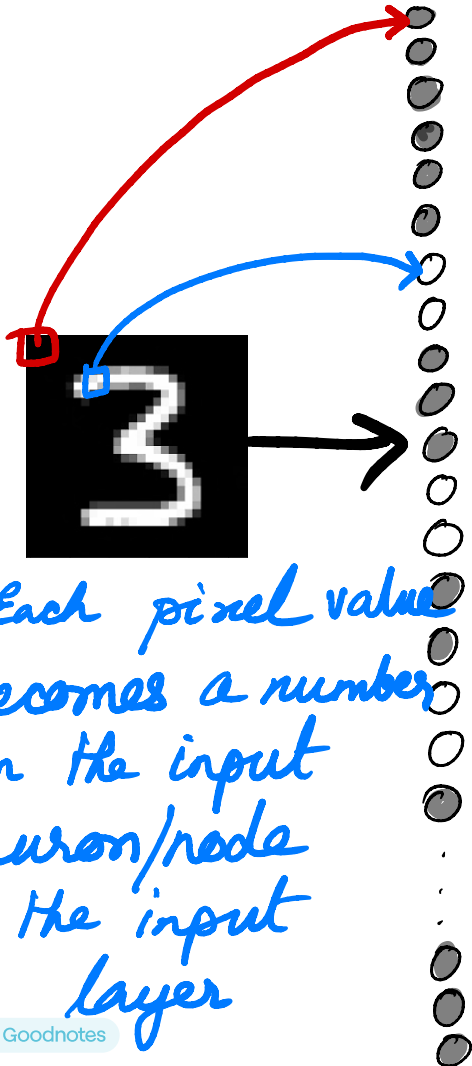
3
7
6

1 white pixel
0 Black pixels

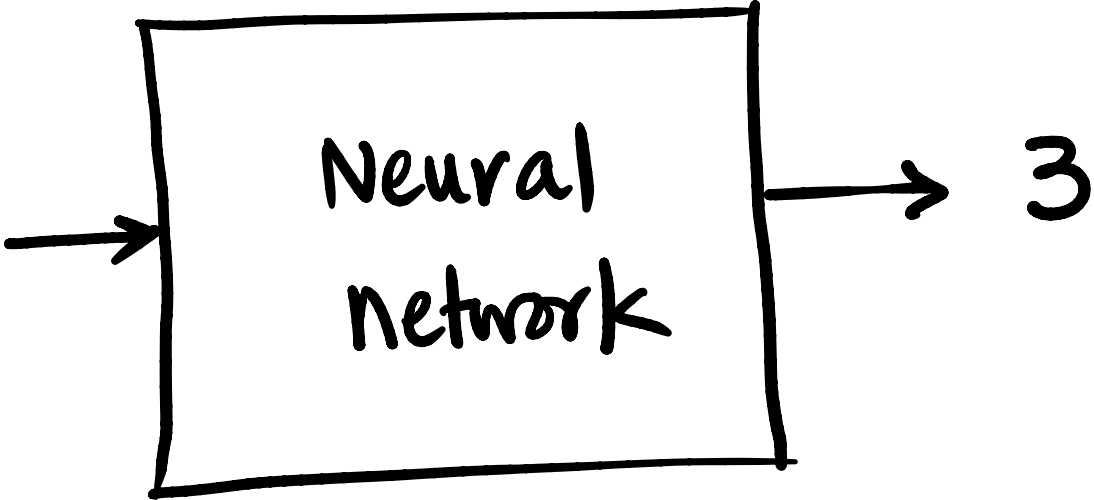


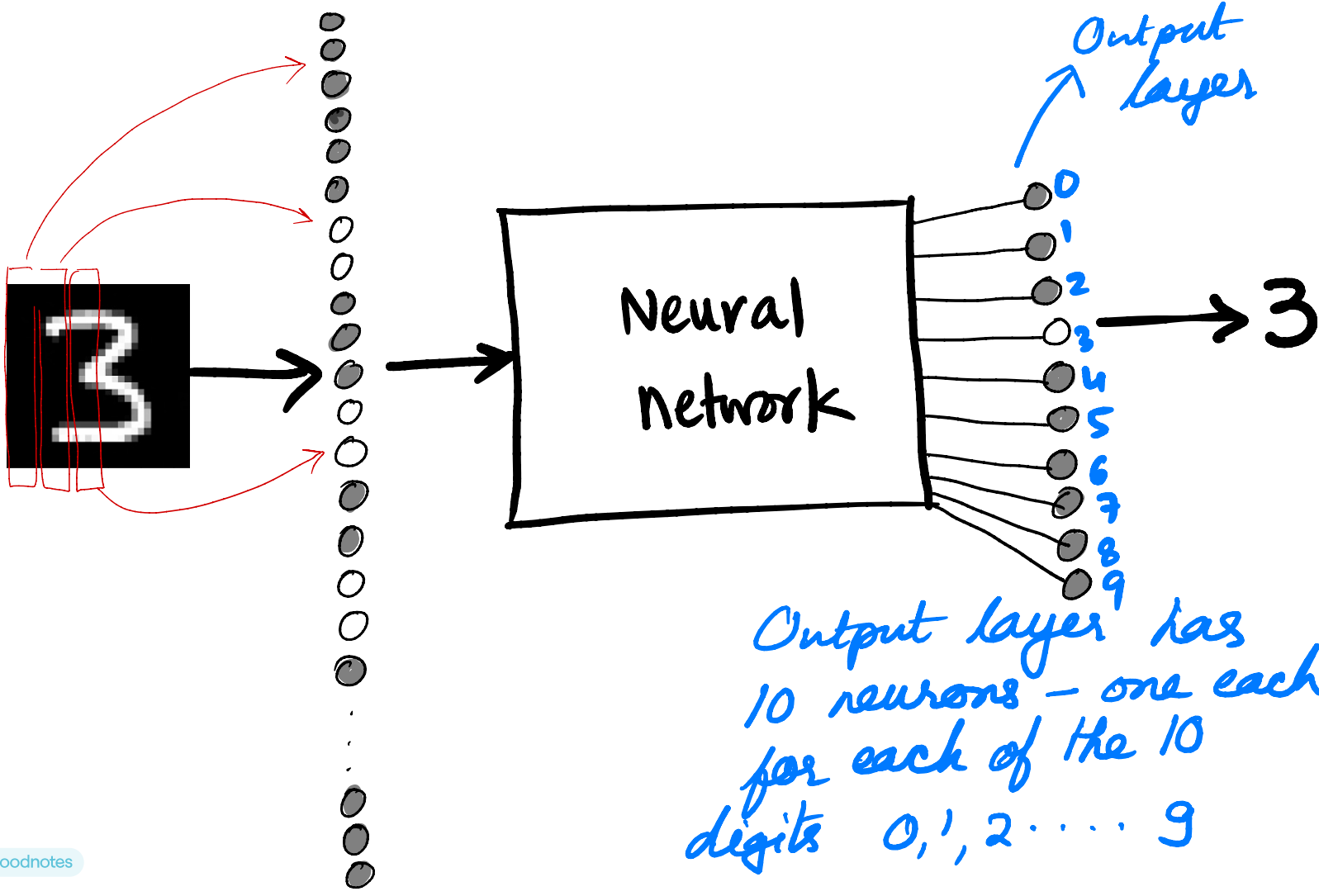


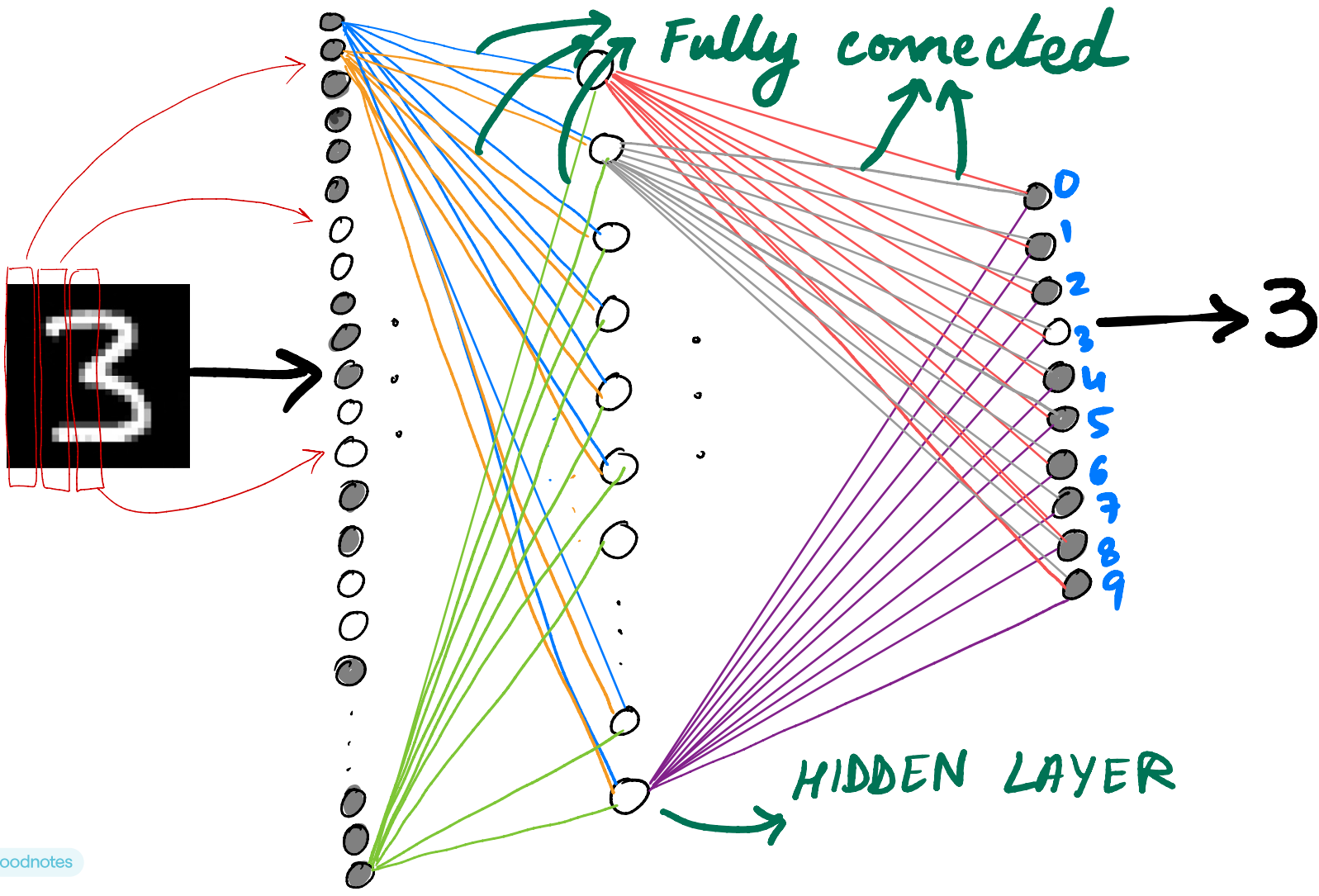


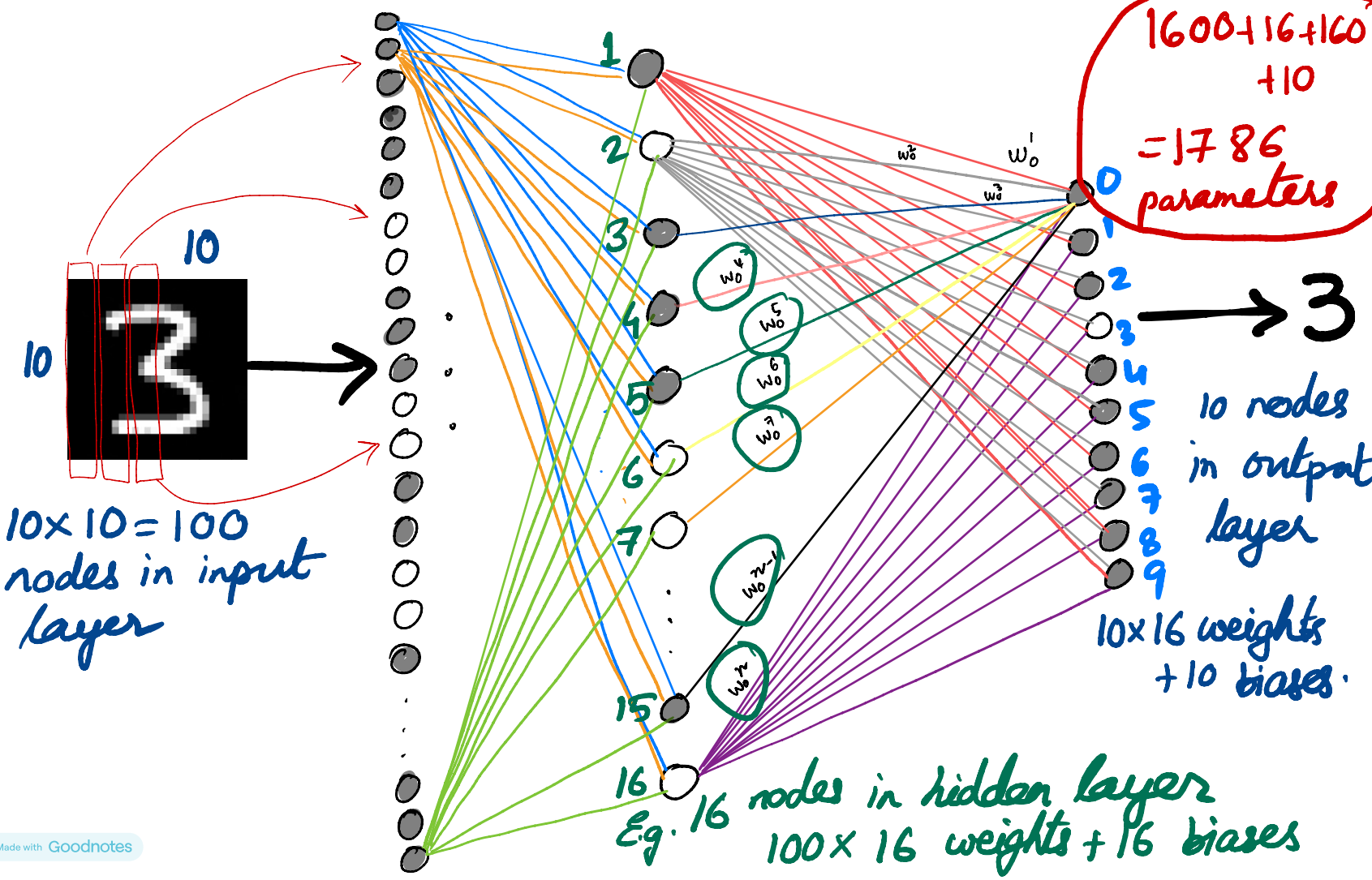


Each pixel value becomes a number in the input neuron/node in the input layer









Terminology You Should Know from These Slides

- 💡 Deep Neural Networks (DNNs)
- 💡 perceptron
- 💡 fully connected layer
- 💡 Input layer, hidden layer, output layer
- 💡 Graphics Processing Unit (GPU)
- 💡 Deep learning