

$$\text{activation } a = f(x) = \frac{1}{1 + e^{-(wx+b)}} \quad \text{output}$$

The term a stands for activation, and it's actually a term from neuroscience, and it refers to how much a neuron is sending a high output to other neurons downstream from it.

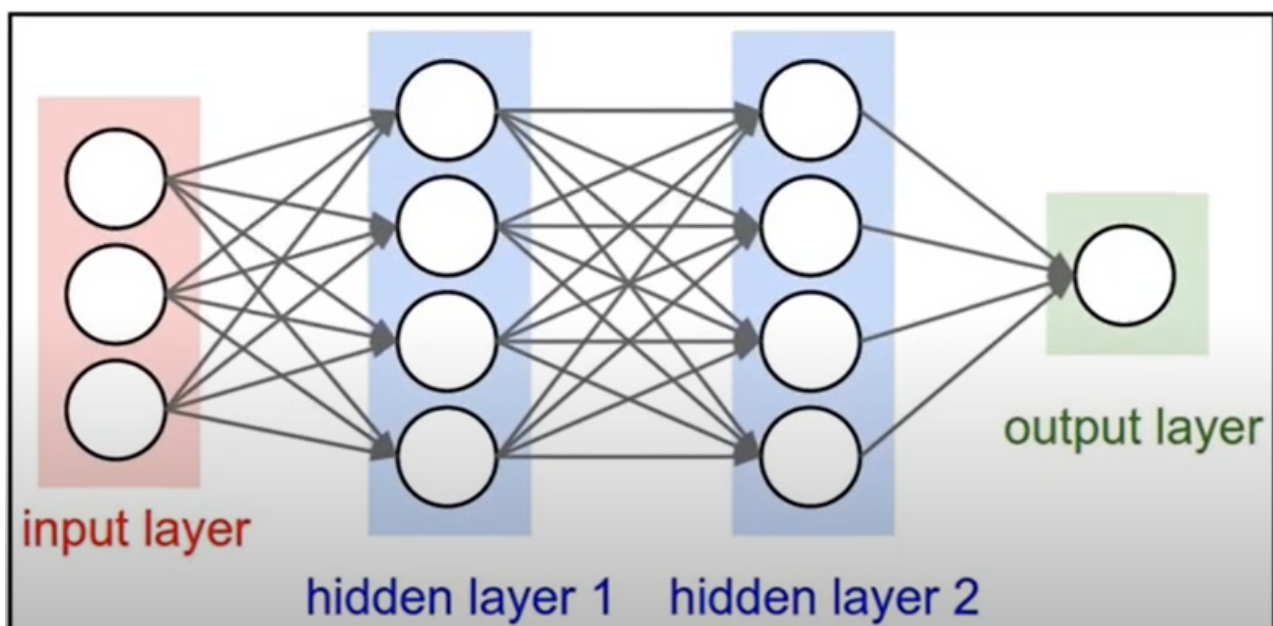
آلية تدريب الآلة :

- يقوم المبرمج بإعطاء مئات الصور لسيارات بانواع و اشكال و الوان مختلفة للخوارزم , , وتحديد رقم 1 في الـ y ليعرف الخوارزم ان هذه سيارة
- ثم يقوم بإعطاء مئات من صور أخرى ليست سيارات , وتحديد رقم 0 للـ y ليعرف الخوارزم ان هذه ليست سيارة
- يقوم الخوارزم , بتحويل كل صور السيارات لمصفوفات بأحجام كبيرة , , وتحديد الثبتات الملائمة لضبط هذه الارقام والتي كلها صور سيارات ($y = 1$)
- يقوم أيضا بتحويل الصور الأخرى (ليست سيارات) لمصفوفات , وإيجاد الثبتات الملائمة لها ($y=0$)
- الان قمنا بعمل تدريب الآلة (او الخوارزم) , وصار لديه القدرة علي تحديد اي صورة هي سيارة او ليست سيارة
- كل ما سبق يتم بما يسمى عينة التدريب Training Sample

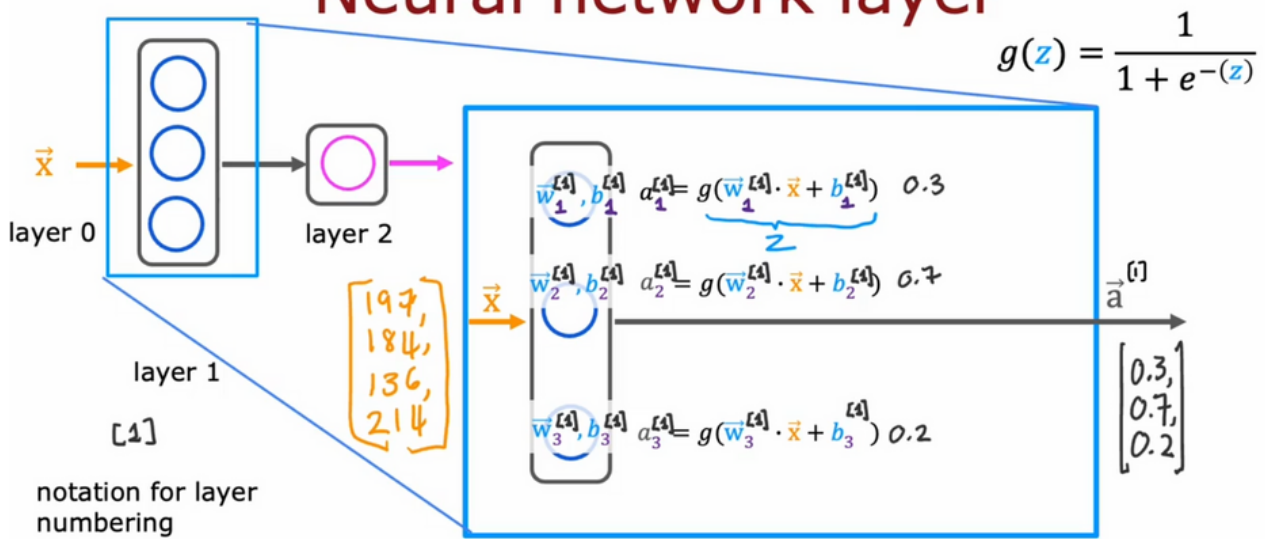
Neural Network Terminologies

1. **x :** Typically represents the input data or feature vector. In the context of a neural network, it's the data that you're passing into the network for processing.
2. **θ (theta):** Often used to represent the parameters (weights and biases) of the neural network. These parameters are learned during the training process to make accurate predictions.
3. **w :** Represents the weights associated with connections between neurons in a neural network. These weights are adjusted during training to influence the network's behavior.

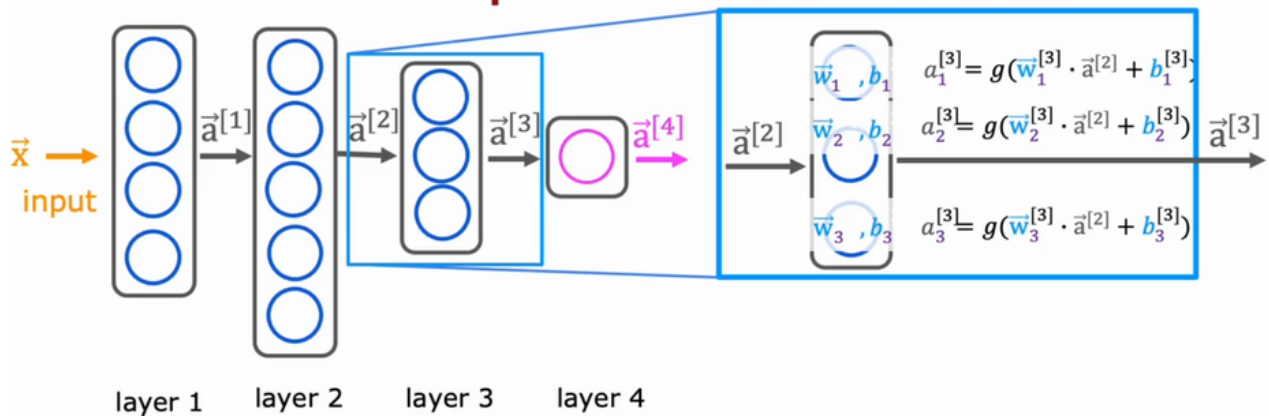
4. **z:** Represents the weighted sum of inputs to a neuron, before applying an activation function. It's calculated as the dot product of the input data (x) and the weights (w), plus a bias term.
5. **g:** Stands for the activation function that introduces non-linearity to the neuron's output. Common examples are ReLU, sigmoid, and tanh.
6. **Layer (L):** Refers to a collection of neurons that process input data together. The layers are stacked to form the neural network's architecture.
7. **\hat{y} (y hat):** Represents the predicted output of the neural network. After passing the input data through the network, \hat{y} is the value produced by the output layer.
8. **y:** Stands for the actual target output or ground truth corresponding to the input data. During training, the neural network aims to minimize the difference between \hat{y} and y .
9. **δ (delta):** Often used in the context of backpropagation, δ represents the error signal associated with a neuron. It's calculated by taking the derivative of the loss function with respect to the neuron's output.
10. **σ (sigma):** Typically used to represent the sigmoid function, which is an activation function that maps the weighted sum of inputs to a value between 0 and 1.
11. **Σ (sigma):** Represents summation. In neural networks, it's used to indicate the sum of weighted inputs.
12. **D:** Refers to the dimensionality or the number of features in the input data (x). It corresponds to the number of input neurons in the network's input layer.
13. **cost:** Also known as the loss or error, the cost quantifies the discrepancy between the predicted output (\hat{y}) and the actual target output (y). The goal during training is to minimize this cost.



Neural network layer



More complex neural network



Activation value of layer l , unit(neuron) j

$$a_j^{[l]} = g(\vec{w}_j^{[l]} \cdot \vec{a}^{[l-1]} + b_j^{[l]})$$

sigmoid "activation function"

output of layer $l-1$ (previous layer)

Parameters w & b of layer l , unit j