

→ The neural networks are further developed in various architectures, and now these

Artificial neural networks (ANNs) are the very core of deep learning.

They are very powerful, making them ideal to tackle large and highly complex machine learning tasks such as classifying billions of images (e.g. Google Images), recommending the best videos to watch, etc.

### Simplest Artificial Neural Network Architecture

#### The Perceptron

→ The Perceptron is one of the simplest ANN Architecture, it is based on an artificial neuron called — Threshold Logic Unit (TLU) —.

→ The inputs and outputs are now numbers (instead of binary on/off values) and each input connection is associated with a weight.

→ Before seeing this architecture, it would be much clear if we can understand the Perceptron Arch.

## perceptron Algorithm

Input:-  $\{(x_1, y_1), \dots, (x_n, y_n)\}$   $x_i \in \mathbb{R}^d$

Iteration:-  $w^0 = 0 \in \mathbb{R}^d$   $[0, \dots, 0]$

until convergence

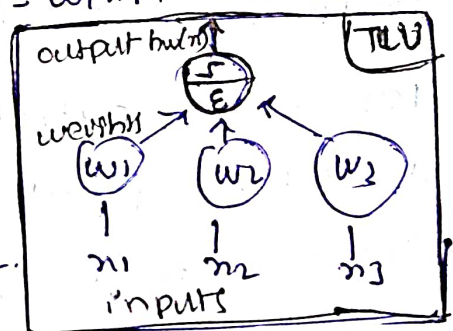
- pick  $(x_i, y_i)$  pair from the dataset
- If  $\text{sign}(w^T x_i) = y_i$  (do nothing)
- Else,  $w^{t+1} = w^t + x_i y_i$  update Rule.  
 $\downarrow \quad \downarrow$   
 $\mathbb{R}^d \quad \{\pm 1\}$
- End.

Assumption:- the dataset is linearly separable.

→ The TLV computes a weighted sum of its inputs, then applies a step function to that sum and outputs the result.

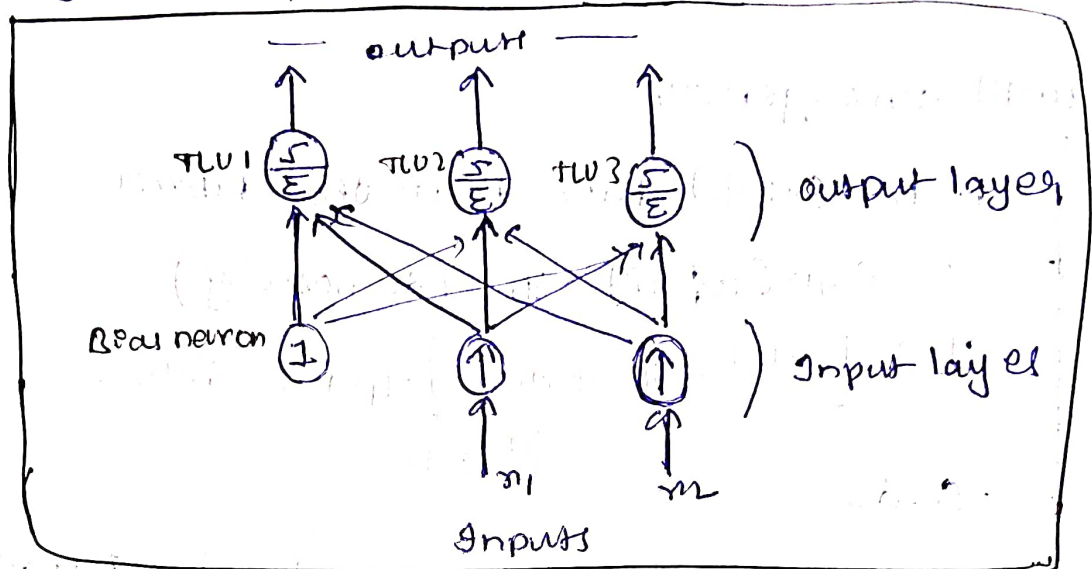
$$h_w(x) = \text{step}(z), \text{ where } z = x^T w = w_1 x_1 + \dots + w_n x_n$$

→ Training a TLV in this case means that finding the right values for  $w_0, w_1$  &  $w_2$ .



→ Here, a perceptron is simply composed of a single layer TLVs with each TLV connected to all the inputs. An extra bias feature is generally added ( $x_0 = 1$ ), which just outputs 1 all the time, called a bias neuron.

→ Here is a perceptron with two inputs and 3 outputs, this perceptron can classify instances simultaneously into 3 different binary classes, making it a multi-output classifier.



→ So how this perceptron is trained?

Just like the update rule we have seen before, → Perceptron is fed one training instance at a time, and for each instance it makes its predictions.

For every output neuron that produced a wrong prediction, it reinforces the connection weights from the inputs that would have contributed to correct prediction.

→ And if the training instances are linearly separable it is said that Perceptron would converge to a solution.

How? →