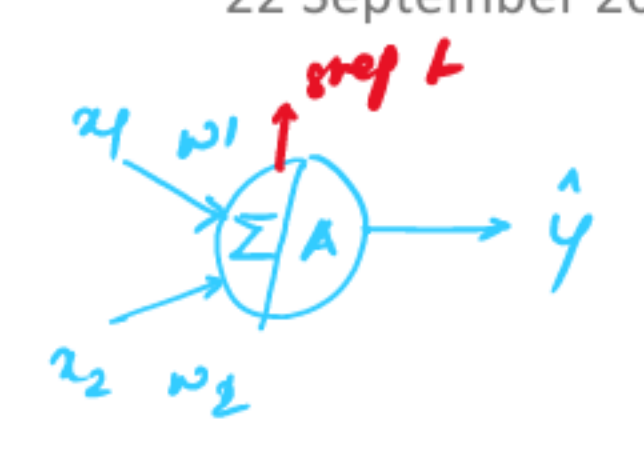
Activation functions

22 September 2023 07:09



step 1 → $x_1 w_1 + x_2 w_2 + bias$ → Result

Sigmoid

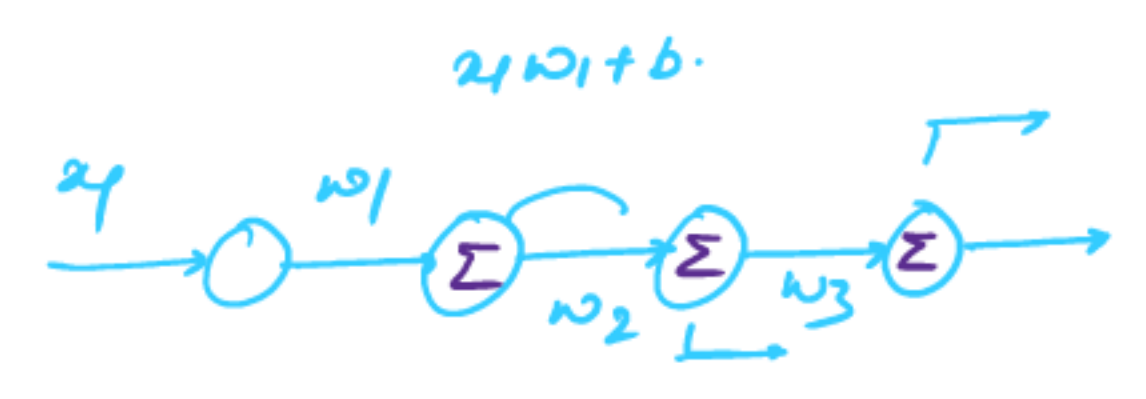
$\frac{1}{1+e^{-y}}$ ← Input

output 0 - 1

Major 0.8 > 1

Act (x)

$x_1 w_1 + b$

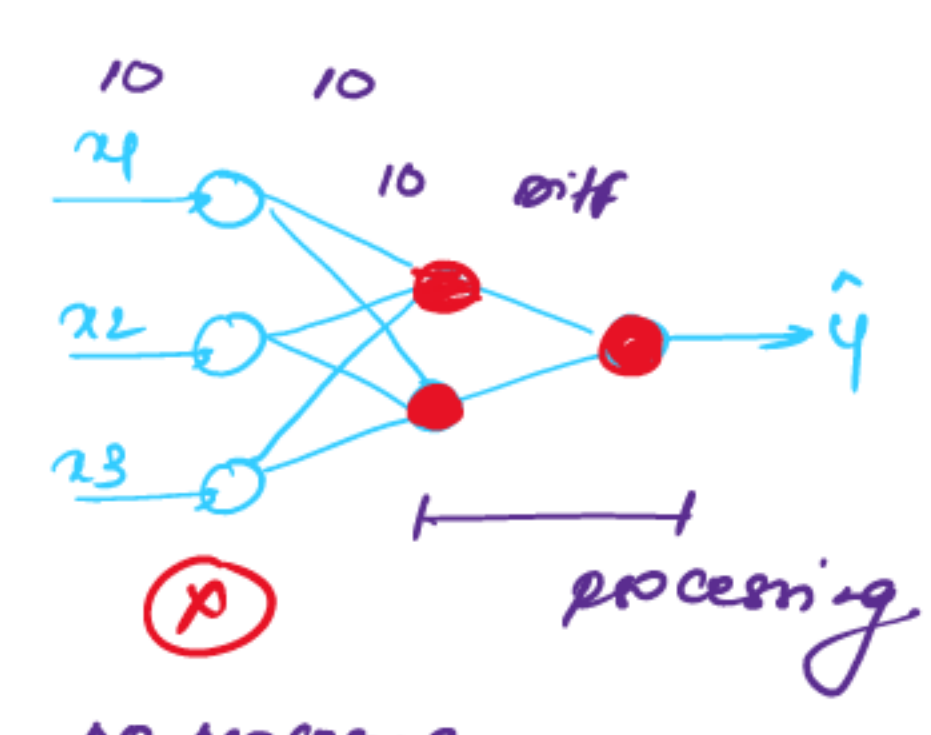


$(x_1 w + b)\, w_2 + b_2)$

Output 0-1

Behave like one single layer.

Input 50y output 1



Input linear

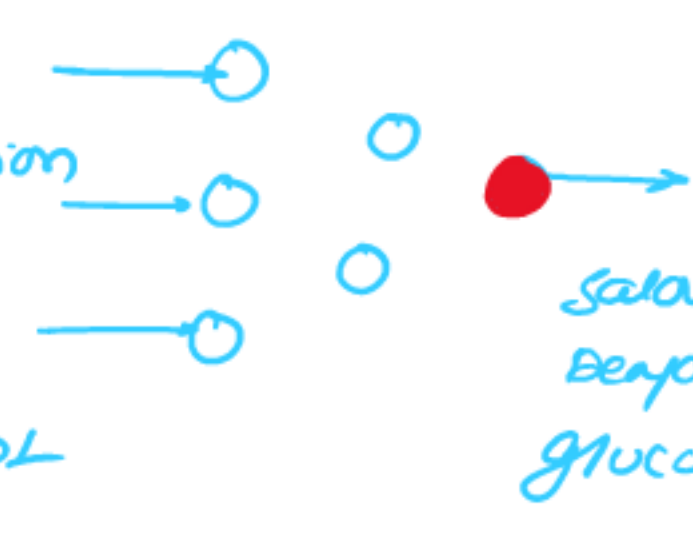
HL output layer }

Linear Activation function.

$f(x) = x$

it is called as no activation function.

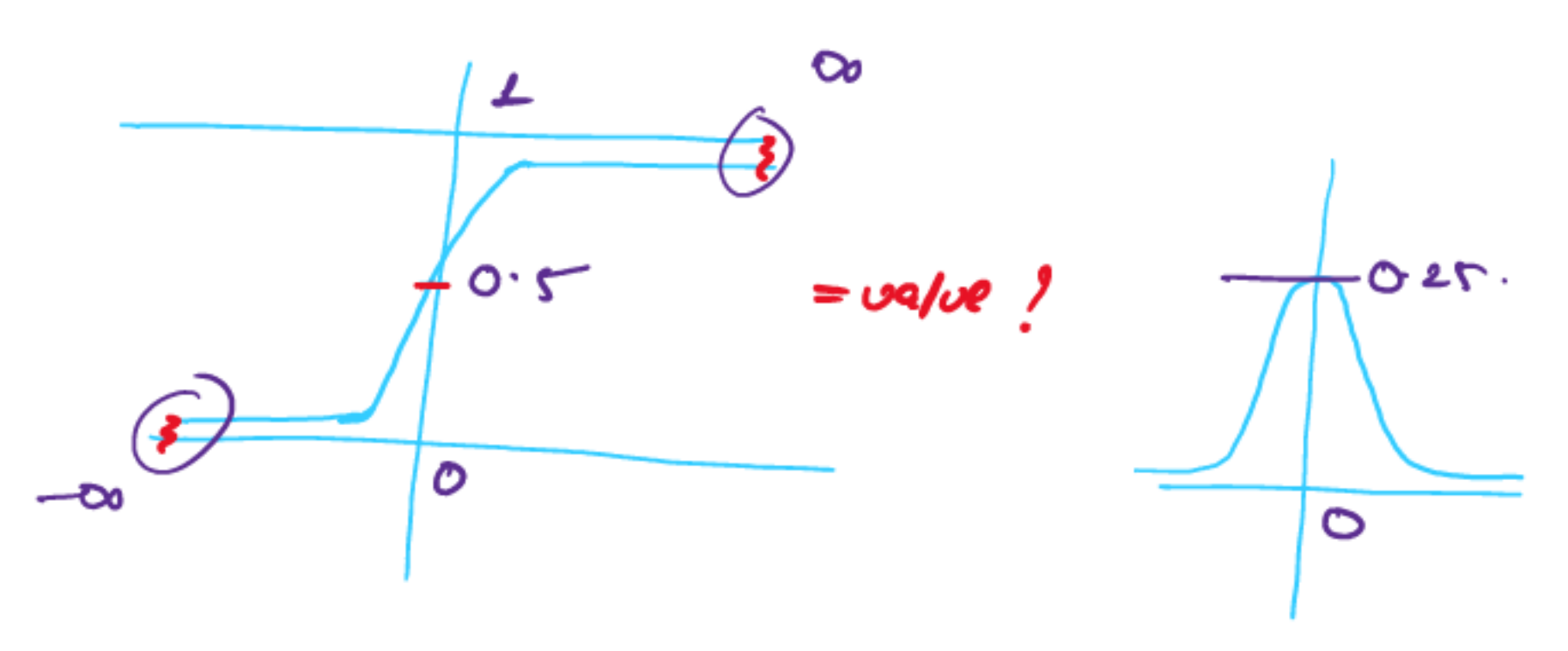
We never use this kind of Activation function in hidden layers.



App.

Output layer ← Regression problem.

Sigmoid



$1 - sig(x)$

Input $-\infty$ to ∞

output 0 to 1

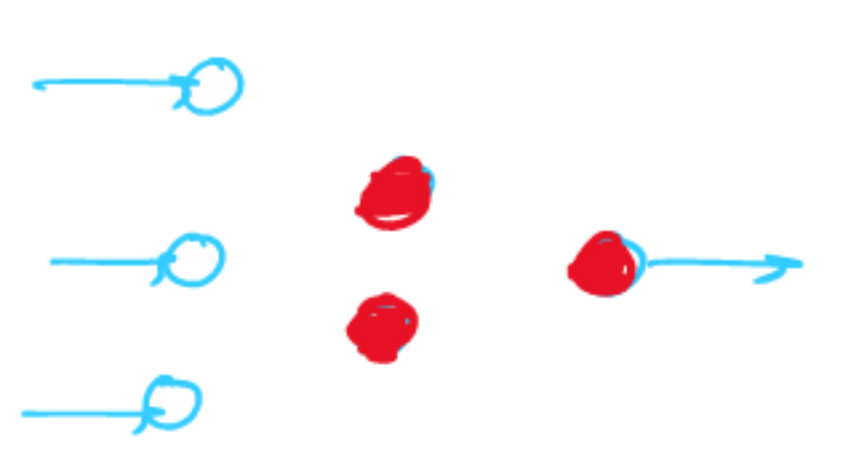
Derivative 0 to 0.25

$\frac{1}{1+e^{-y}}$ ← (-y) Input ← $x_1 w_1 + x_2 w_2 + b$ → 0 to 1

exponent → Calculation are expensive

used

* hidden layer
* output layer → Binary



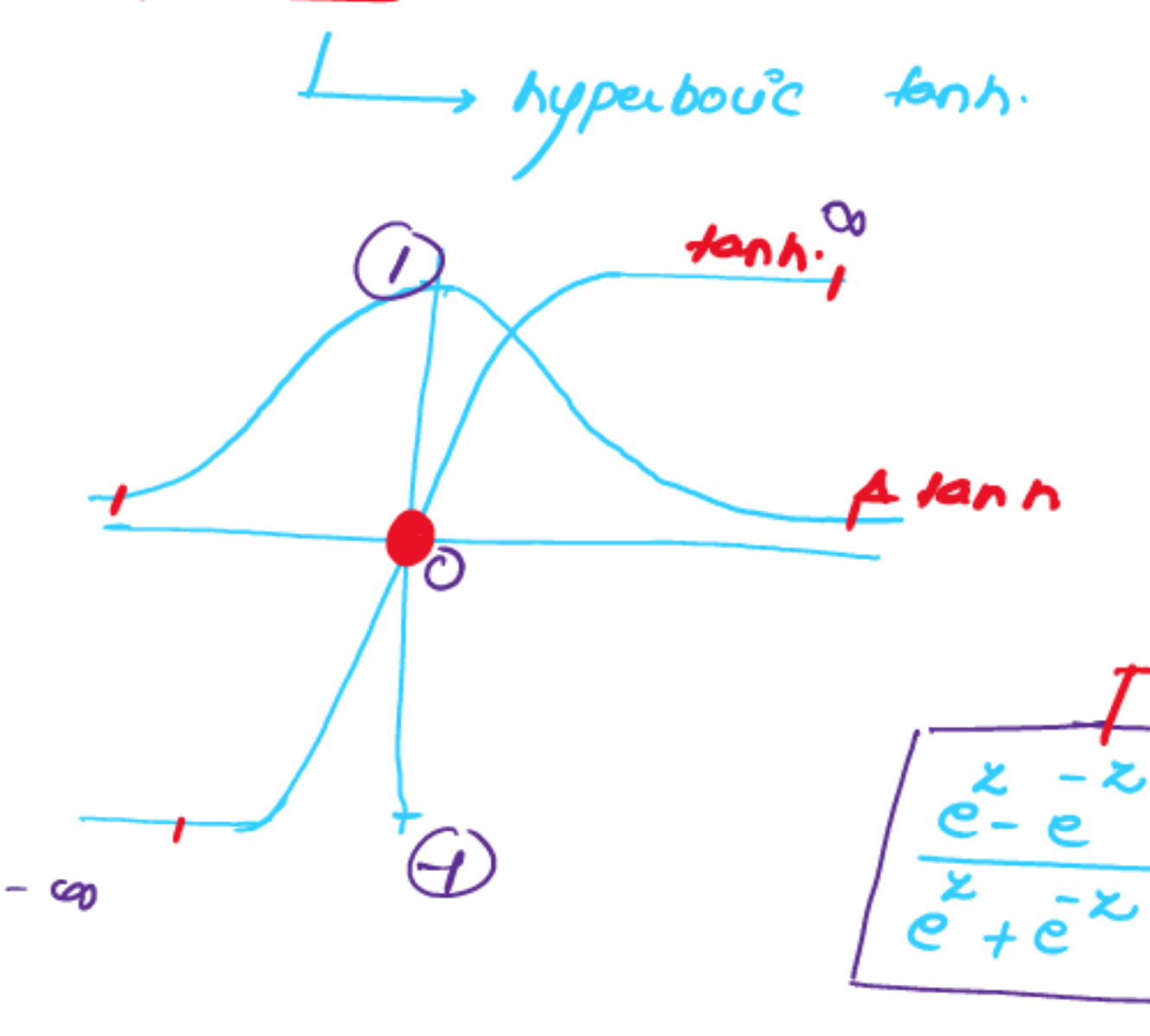
1. sigmoid is having smooth curvature
2. sigmoid in output layer used to Binary class classification.

Problem → vanishing gradient problem. }

Input	HL1	HL2	HL3	HL4	output
	sig	sig	sig	sig	

Tanh

→ hyperbolic tanh.

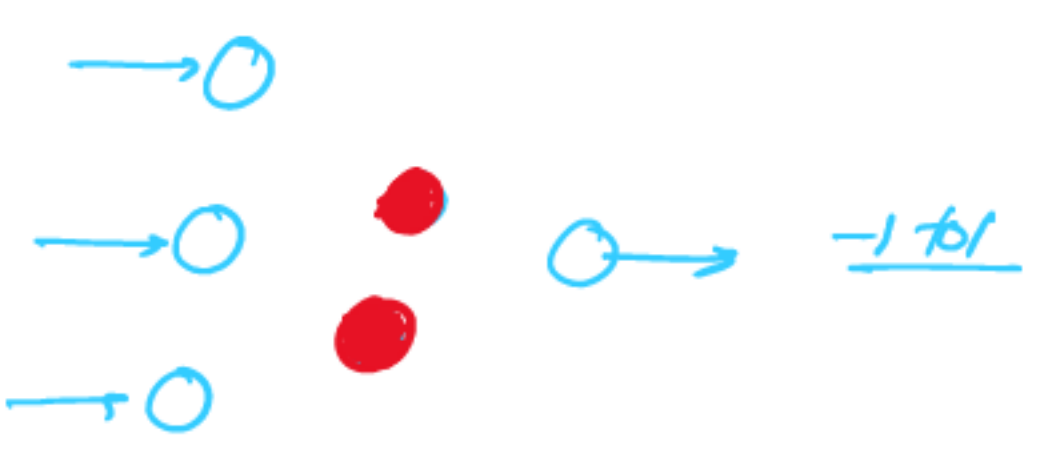


$\frac{e^{z} - e^{-z}}{e^{z} + e^{-z}}$ ← Input

Input $-\infty$ to ∞

Output -1 to 1

Derivat. 0 to 1 }

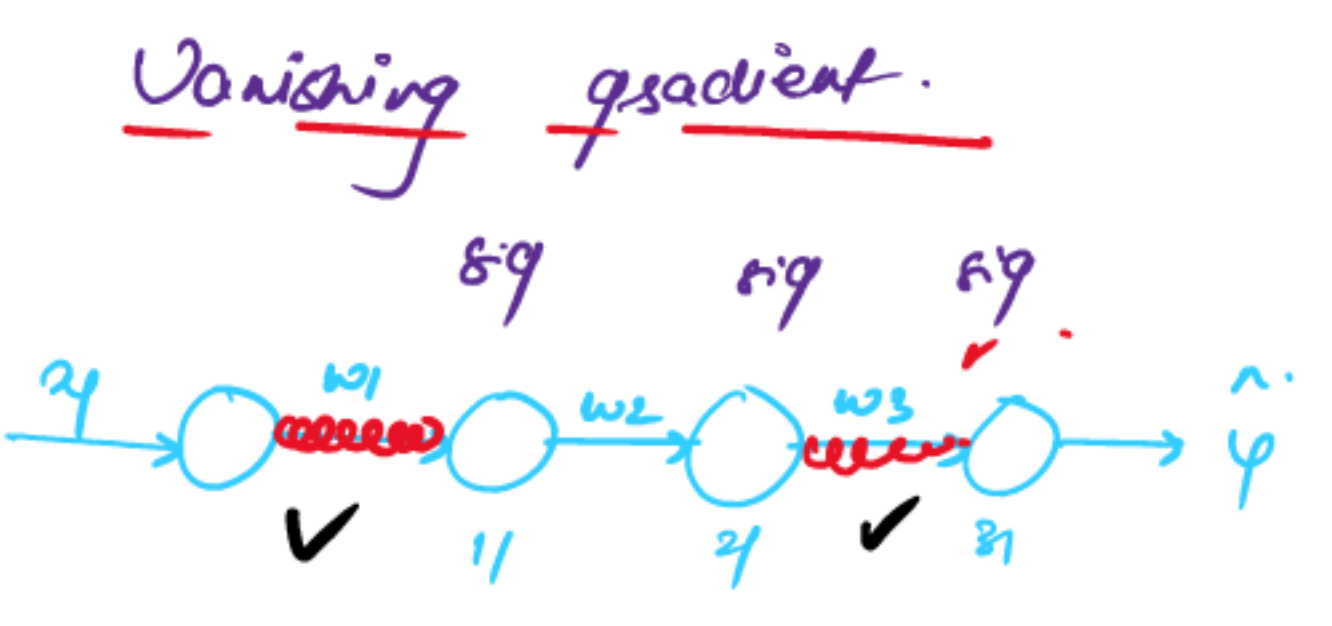


1. Calculations are expensive
2. zero centric function.

Problem. → vanishing gradient problem.

Input	HL1	HL2	HL3	HL4	Out
	tanh	tan	tanh	tan	

Vanishing gradient.



$w_{11} = w_{old} - \alpha \frac{\partial L}{\partial w_{old}}$

$$\frac{\partial L}{\partial w_{old}} = \frac{\partial L}{\partial O_{31}} \times \frac{\partial O_{31}}{\partial O_{21}} \times \frac{\partial O_{21}}{\partial O_{11}} \times \frac{\partial O_{11}}{\partial w_{11}}$$

$= 0.25 \times 0.1 \times 0.008 \times 0.004.$

= Result miniscule. small.

$w_{11}\,new = w_{11}\,old - \alpha \frac{\partial L}{\partial w_{11}}$

$= w_{old} - \alpha \times 0.00001$

$= w_{old} - small.$

$w_{new} = w_{old}$

10 — 9.999.

100 1000

10 - 9.I

10.81

1. parameters of higher layer changes significantly
2. changes in lower layers are miniscule.
3. stagnation in learning.
4. epoch ↑ perf stagnation