x_1

x_2

x_3

35 random w

$\hat{y}$

$$L(w) = (\hat{y} - y)^2$$

Find ∇L the gradient ?

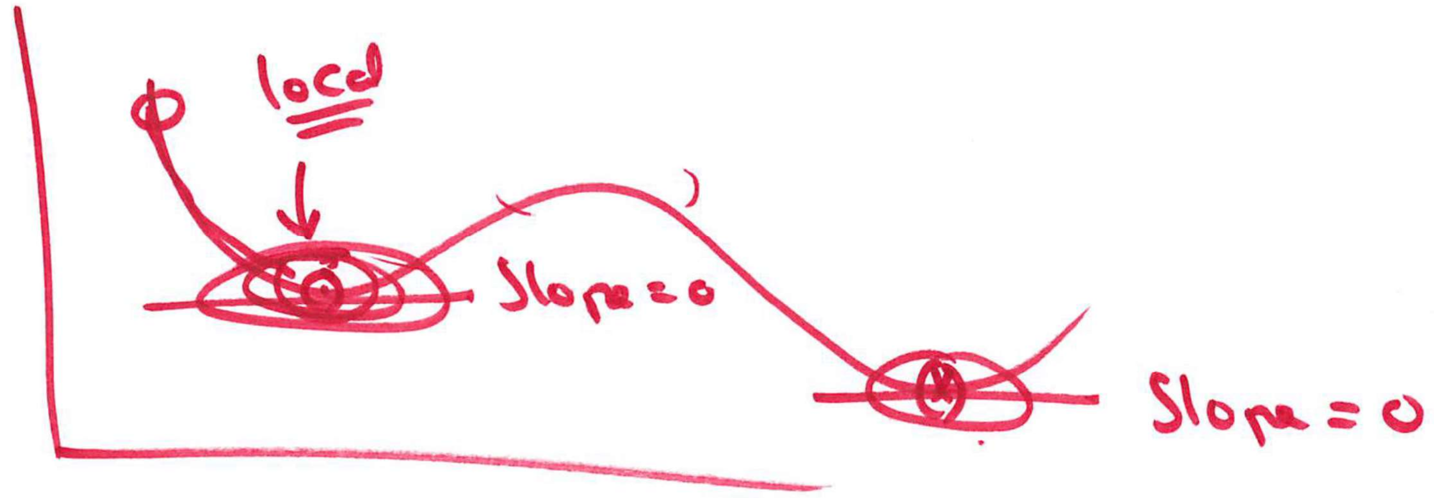
$$w^{new} = w^{old} - \eta \nabla L$$

η → learning rate

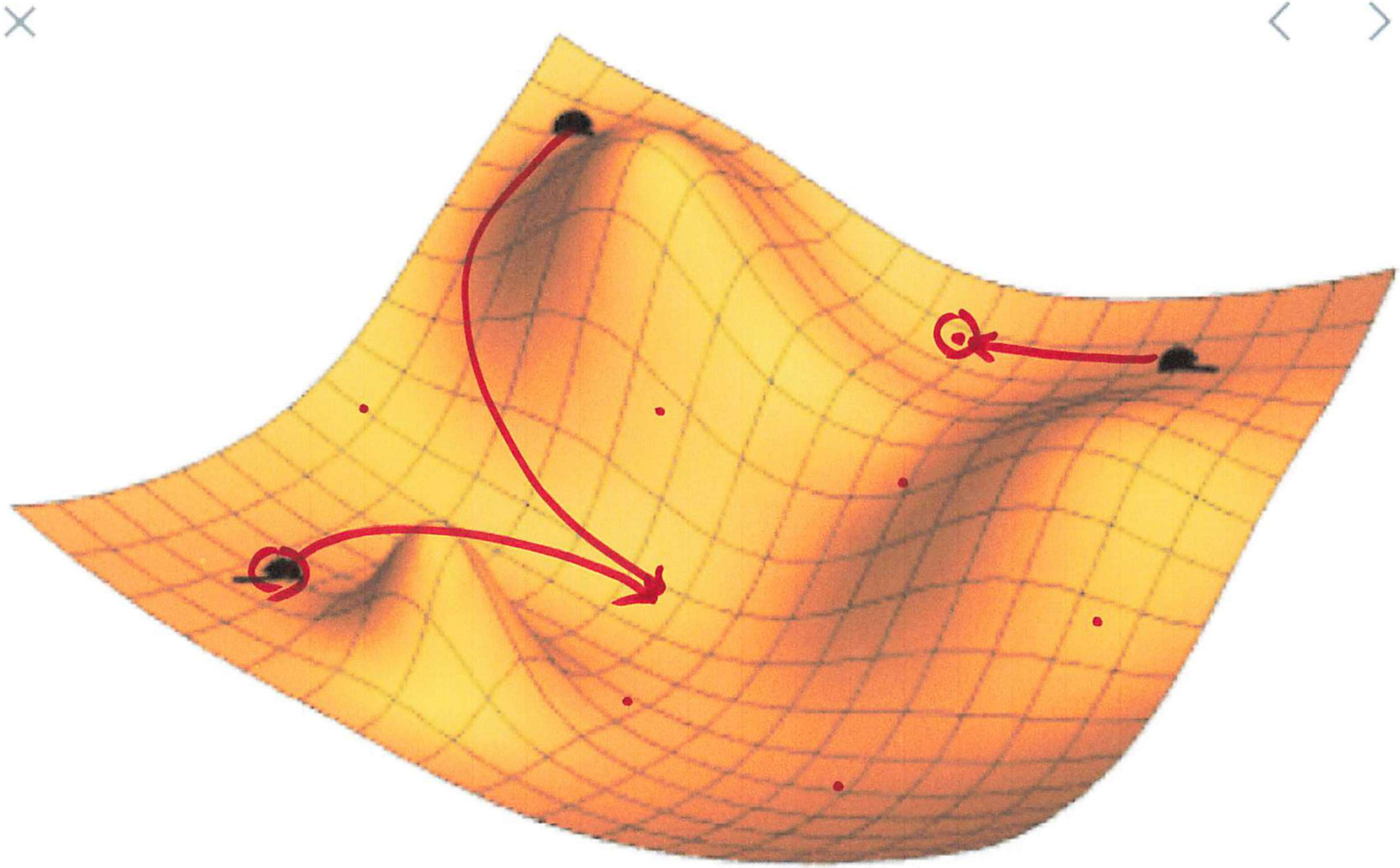
Backprop.

local v. global

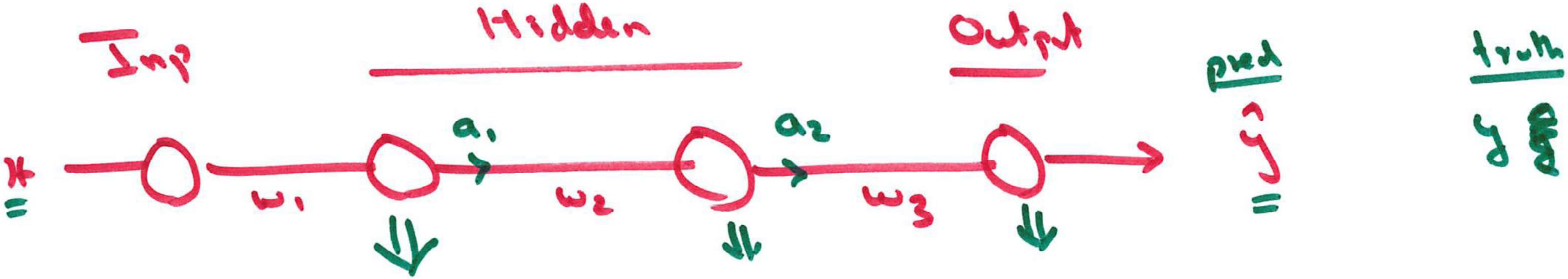
This file is meant for personal use by venkhatbalaji@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.



This file is meant for personal use by venkhatbalaji@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.



This file is meant for personal use by venkhatbalaji@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.



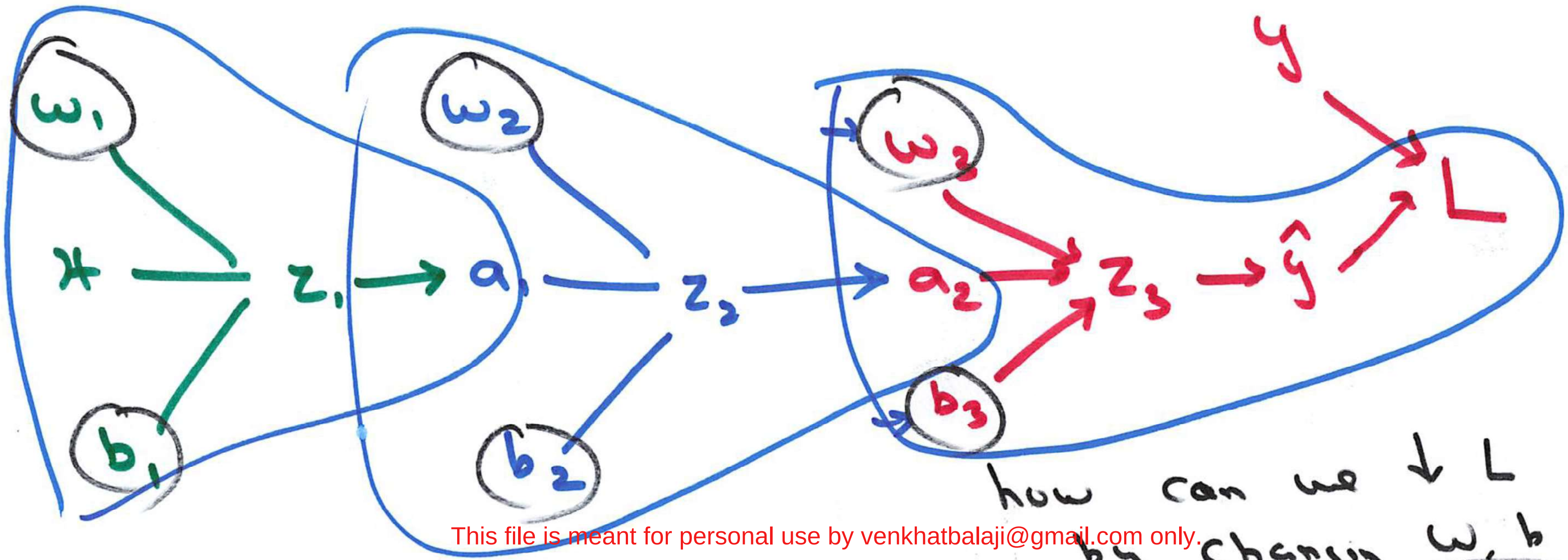
$$\begin{cases} z_1 = x w_1 + b_1 \\ a_1 = \sigma(z_1) \end{cases}$$

$$z_2 = a_1 w_2 + b_2$$
$$a_2 = \sigma(z_2)$$

$$z_3 = a_2 w_3 + b_3$$
$$\hat{y} = \sigma(z_3)$$

$$\boxed{L(w) = \sum(\hat{y} - y)^2}$$

$\cancel{\nabla w}$ ∇L ?



how can we ↓ L by changin w, b

This file is meant for personal use by venkhatbalaji@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.

Slope of L with respect to ω_3 $\Bigg\} = \dfrac{\partial L}{\partial \omega_3}$

$$\frac{\partial L}{\partial \omega_3} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_3} \cdot \frac{\partial z_3}{\partial \omega_3}$$

$$\frac{\partial L}{\partial \hat{y}} \rightarrow 2(\hat{y} - y) \qquad \frac{\partial \hat{y}}{\partial z_3} \rightarrow \sigma'(z_3) \qquad \frac{\partial z_3}{\partial \omega_3} \rightarrow a_2$$

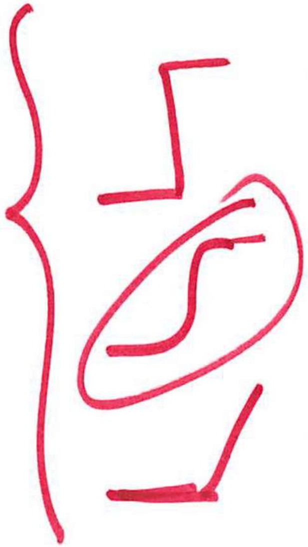
$$\frac{\partial a_2}{\partial \omega_2}$$

This file is meant for personal use by venkhatbalaji@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.

$\square = f(\sum \underline{\underline{w_i}}\, x_i + b)$

non linear func

$\square = f(\sum \underline{\underline{w_i}}\, \square + b)$

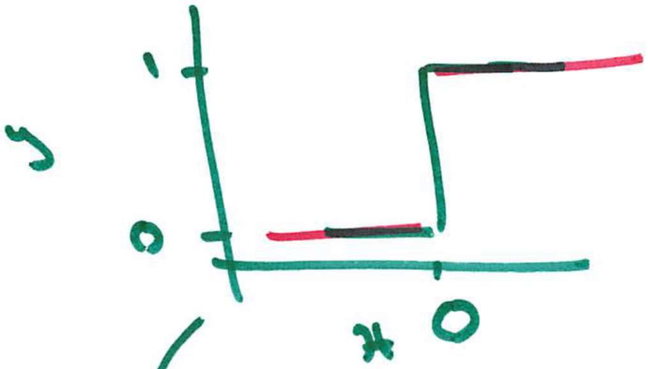


Step

Sigmod

ReLu

This file is meant for personal use by venkhatbalaji@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.



$$y = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

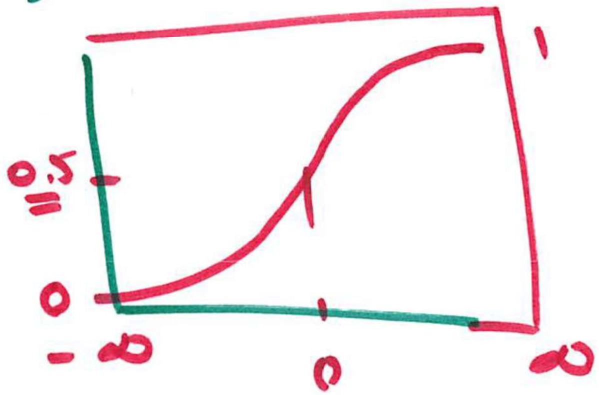
~~$y = 0$~~ $\quad y = f(x)$

$$w^n = w^o - \eta\left(\frac{\text{slope of } L}{w}\right)$$

$\sigma'(x)$

Sigmoid $\rightarrow$ 0 to 1

$$y = \sigma(x) = \frac{1}{1+e^{-x}} = \frac{e^x}{e^x+1}$$



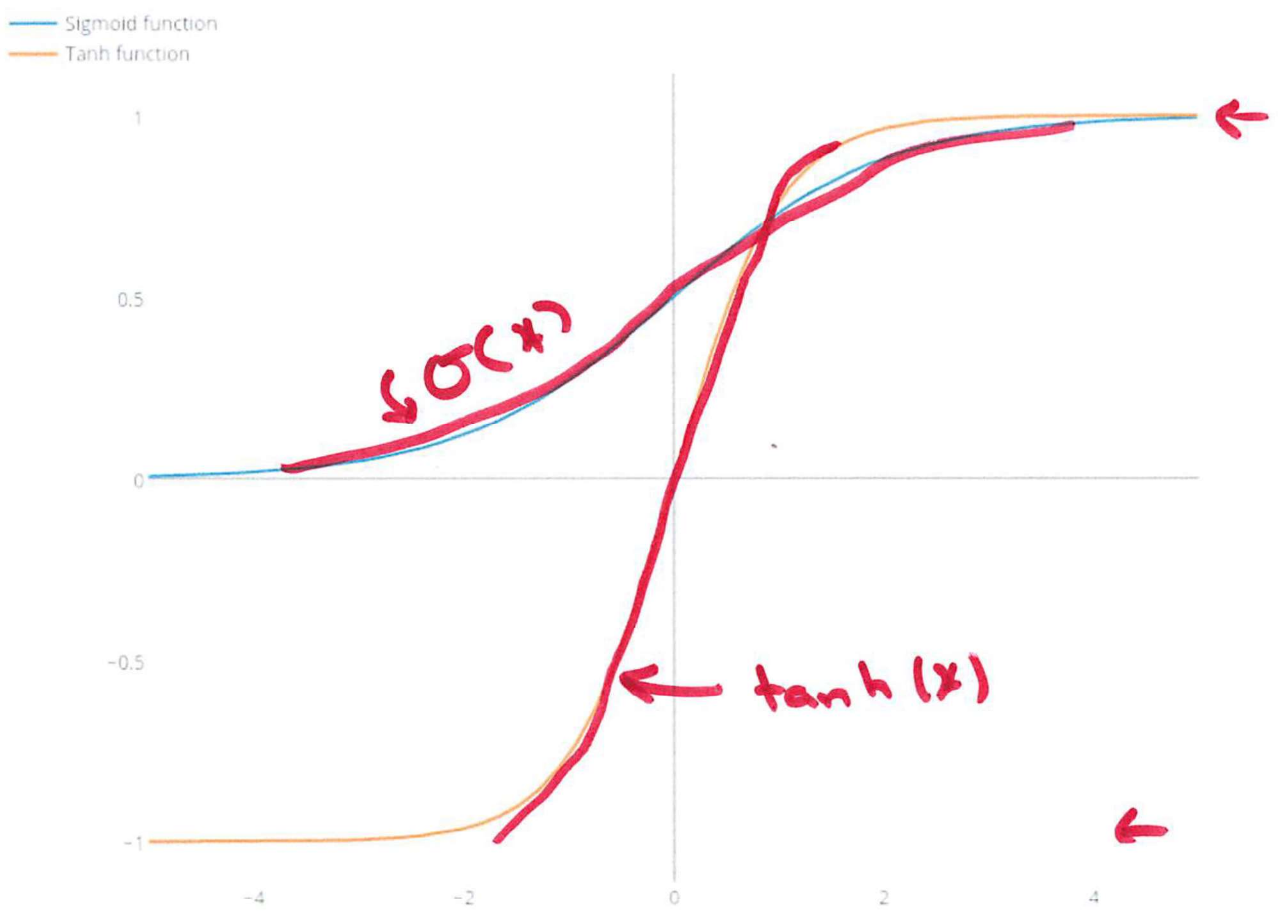
$\boxed{2\sigma(x) - 1} \rightarrow -1$ to 1

$$y = \tanh(x) = 2\sigma(2x) - 1 \rightarrow -1 \text{ to } 1$$

$$\hookrightarrow \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x} - 1}{e^{2x} + 1}$$

This file is meant for personal use by venkhatbalaji@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.

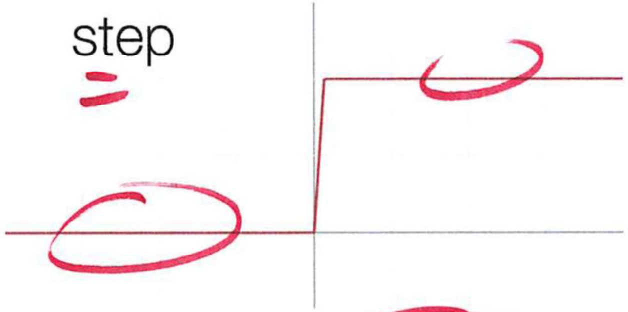
$$\tanh(x) = 2\sigma(2x) - 1$$



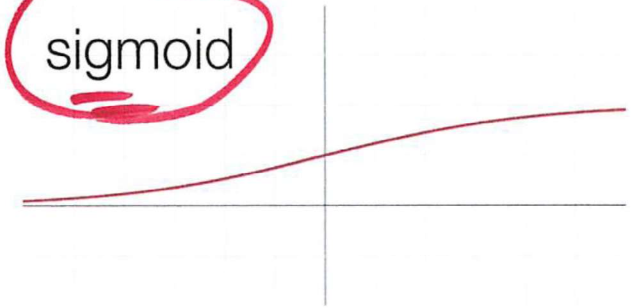
This file is meant for personal use by venkhatbalaji@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.

Activation Functions

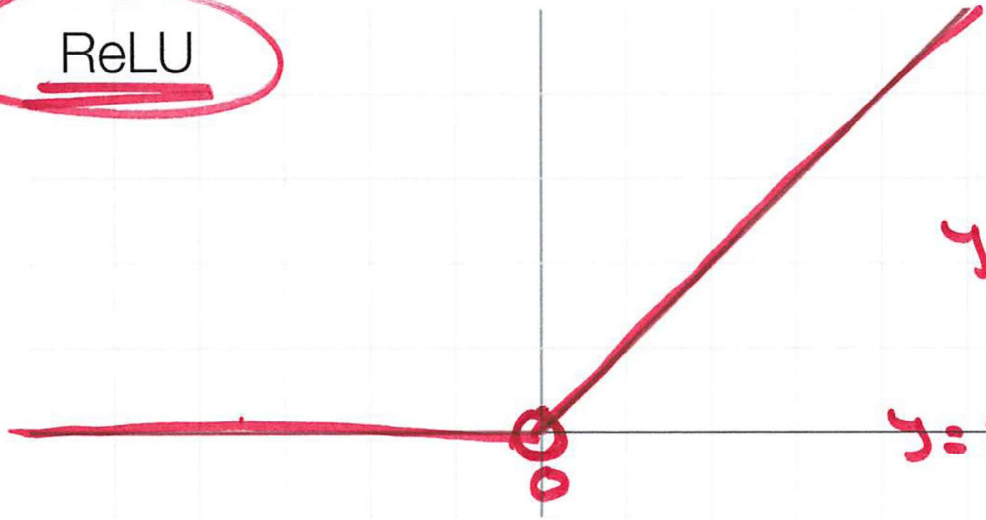
step



sigmoid



ReLU

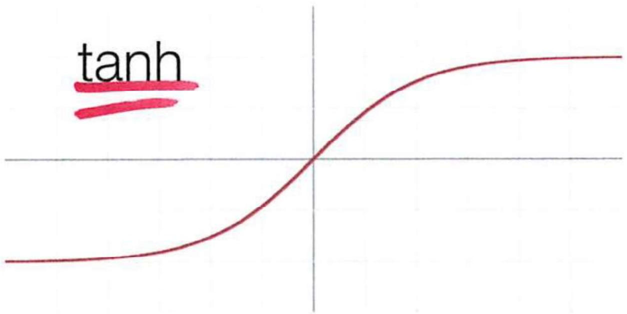


$$y = f(x)$$

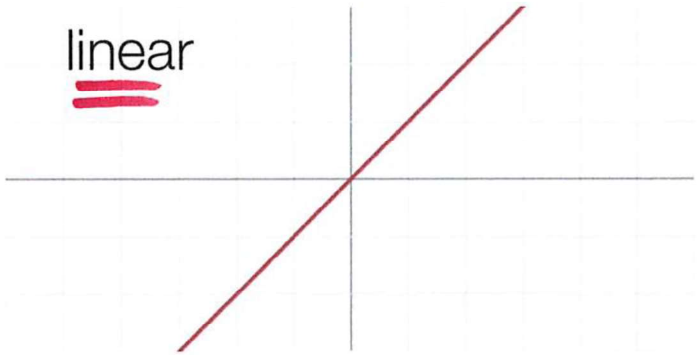
$$y = \begin{cases} 0 & \text{if } x \le 0 \\ x & \text{if } x > 0 \end{cases}$$

$$y = \max(0, x)$$

tanh

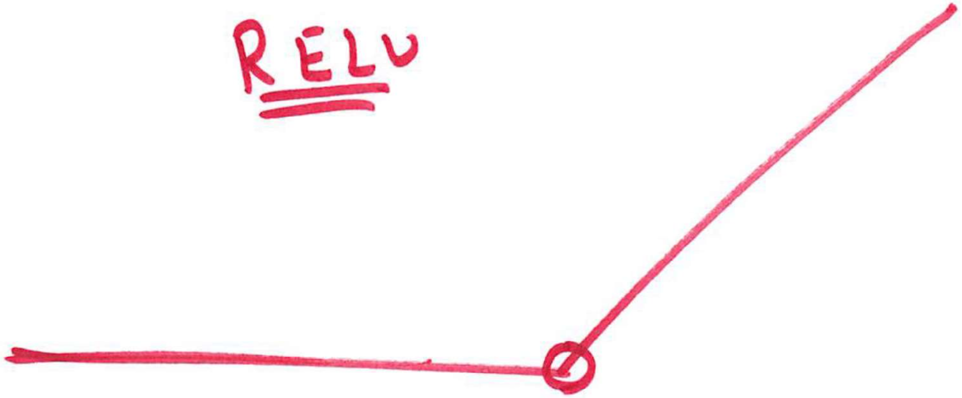


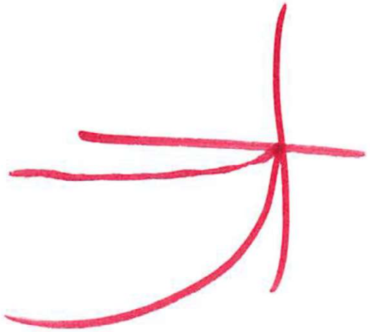
linear



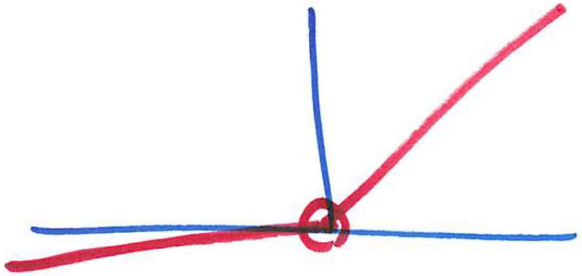
This file is meant for personal use by venkhatbalaji@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.

RELU



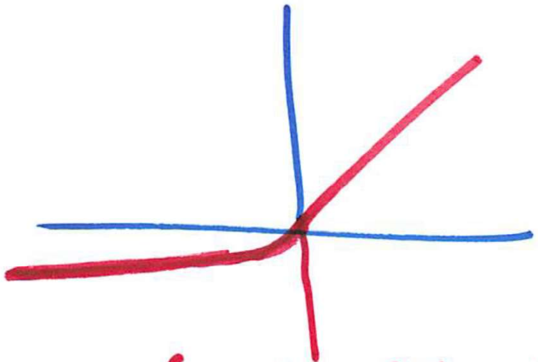


leaky RELU



$$y = \begin{cases} x & \text{if } x \geq 0 \\ 0.001x & \text{if } x < 0 \end{cases}$$

EReLU



$$y = \begin{cases} x & \text{if } x \geq 0 \\ \alpha(e^{x} - 1) & \text{if } x < 0 \end{cases}$$

This file is meant for personal use by venkhatbalaji@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.

Hidden → ReLU, Leaky ReLU, ERelU, Sigmoid, tanh, ~~linear~~

Output layer → Regression → linear

Classification

binary

Sigmoid / tanh

multiclass

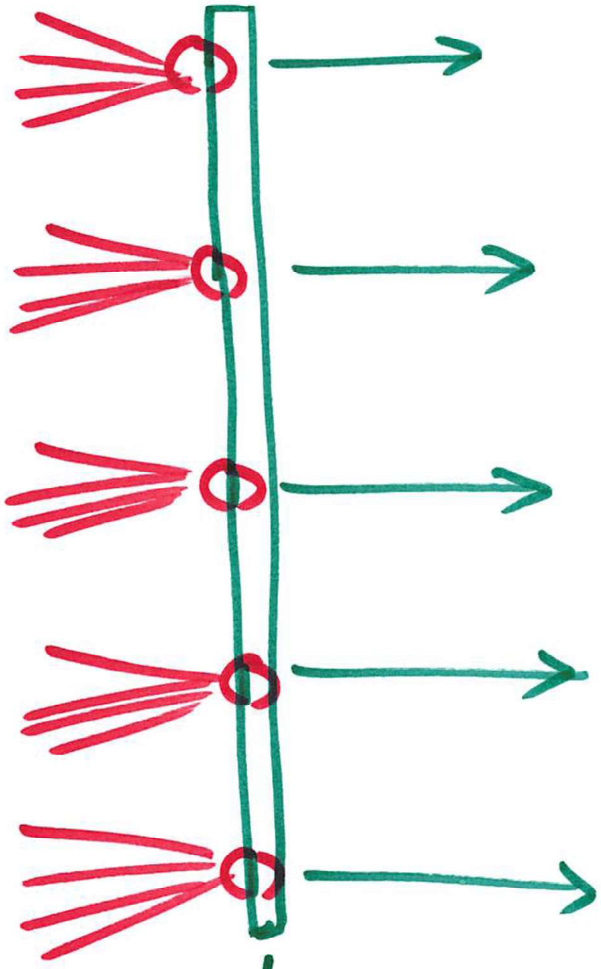
A B C D E

Softmax

1

This file is meant for personal use by venkhatbalaji@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.

$f(\dots) \rightarrow$



Softmax activation

$$z_1 = (\Sigma w x + b)$$

$$\cancel{a_1 = f(z_1)}$$

This file is meant for personal use by venkhatbalaji@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.