Artificial neural networks: (ANN).

Deep Learning

$y = f(x) + \varepsilon$

- Reg
- Class
- Parametric
  - functional form
    - Deep learning
- Non-Parametric
  - Trees

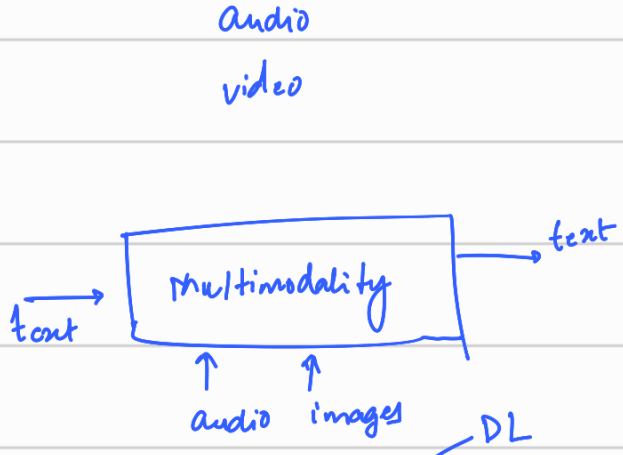
Deep learning

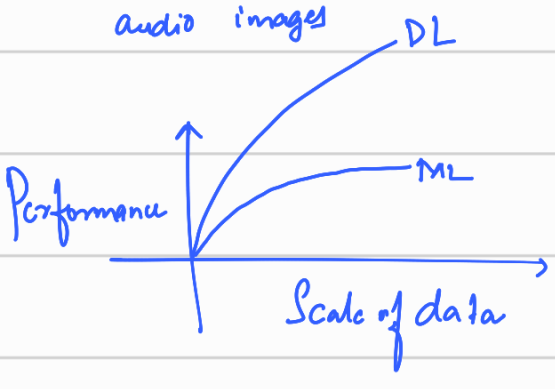
images . Graphs

text

audio

video





Foundation of NN

Deep Learning

ANN → Tables

CNN → Images

RNN / LSTM / Transformers → text / seq.

GNN → Graphs

Neuron?

Neural network

NN, ANN, tensorflow / Keras

how to improve NN ✓

Explainability

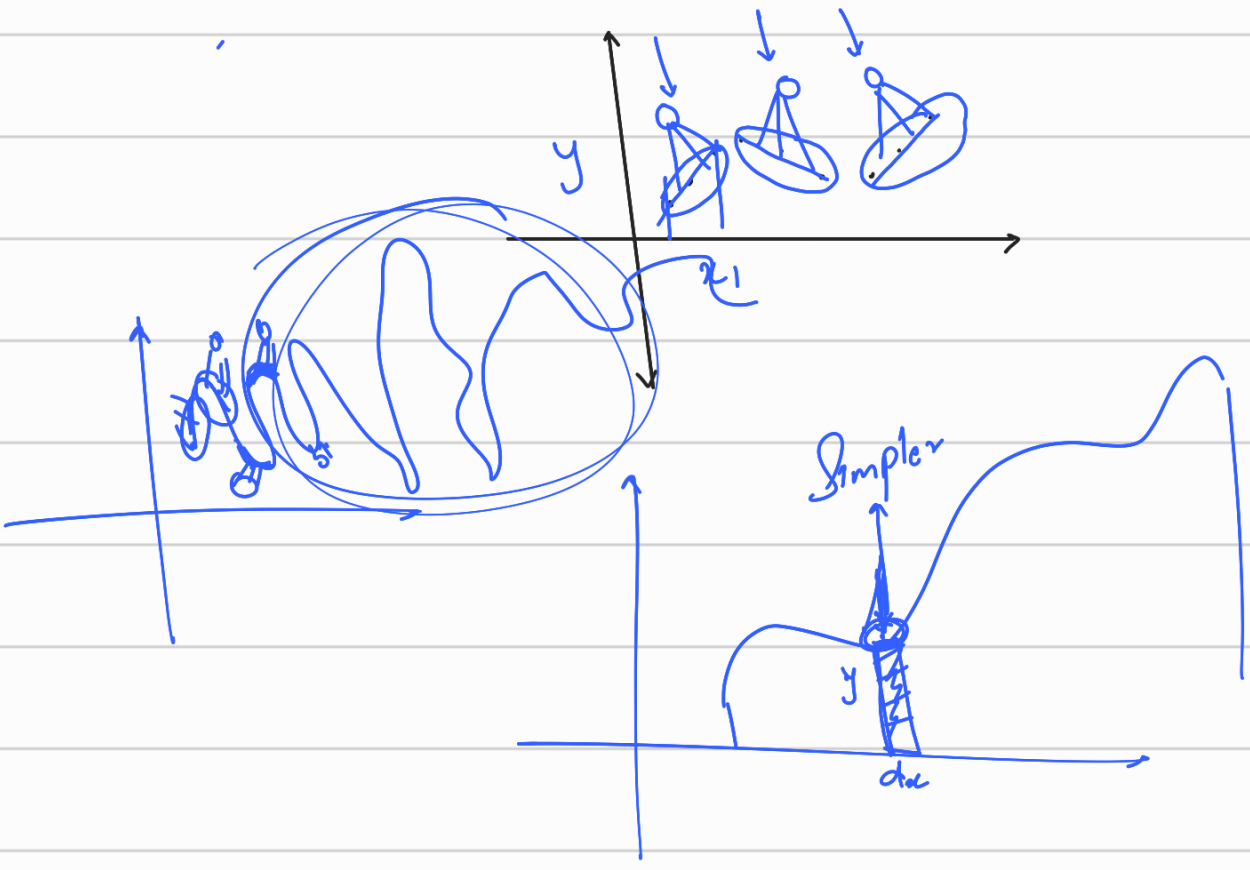
$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

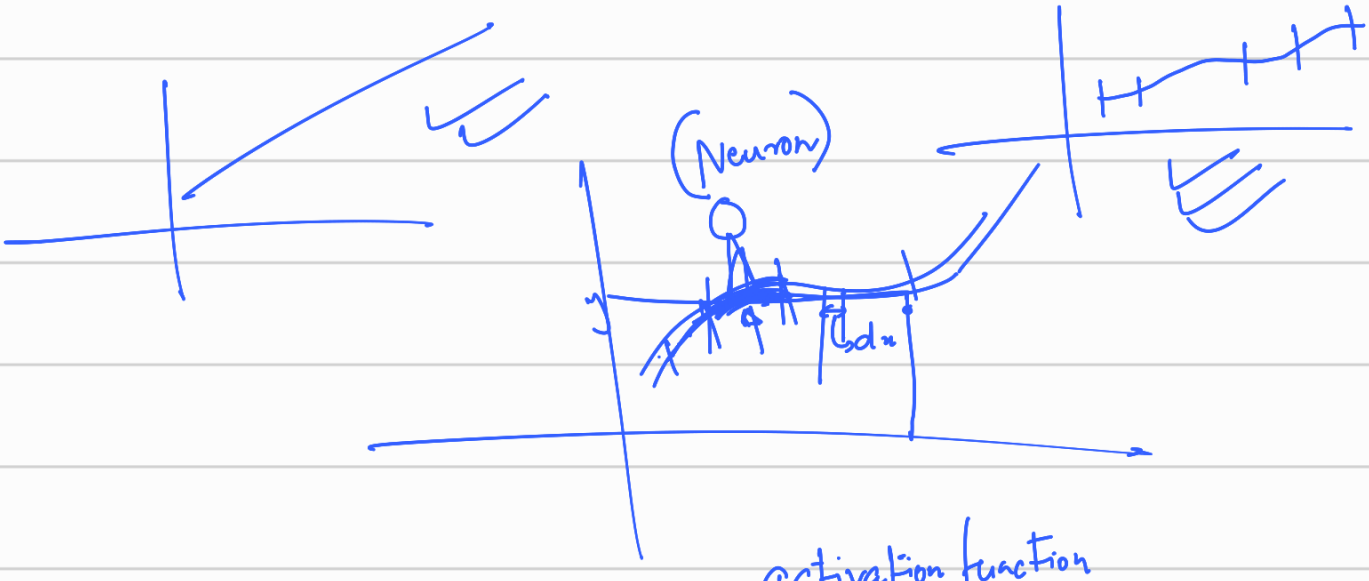
$y = \beta_0 + \beta_1 x_1^2 + \beta_2 x_2^2$

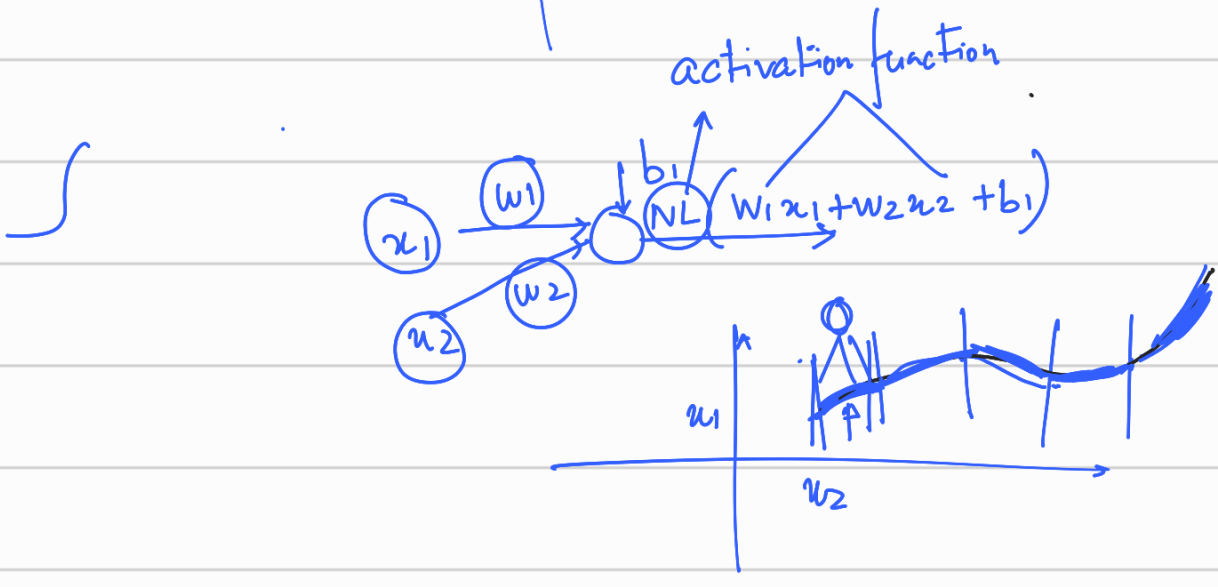
Deep Learning → parametric approach → $y \rightarrow f(x) + \varepsilon$

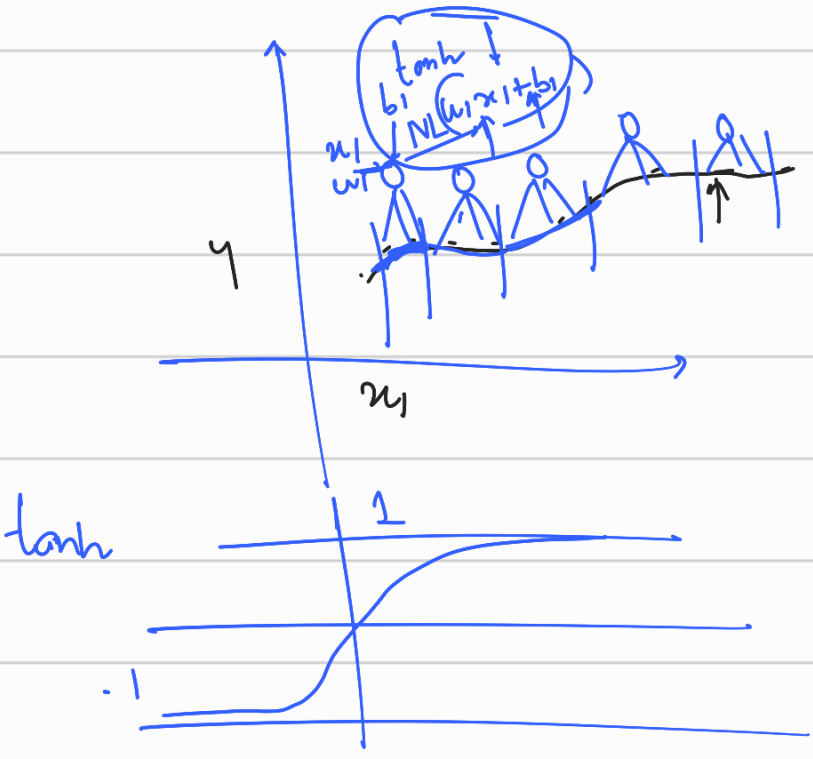
↓

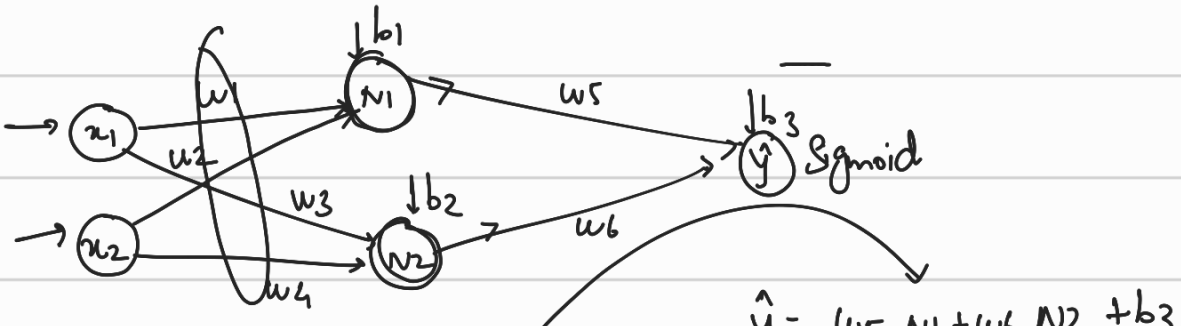
functional form.


y
x_1
Simpler
y
dx


(Neuron)
y
dx


activation function
w_1
b_1
x_1
NL
$w_1x_1+w_2x_2+b_1$
w_2
x_2
x_1
x_2





$$\hat{y} = w_5 N_1 + w_6 N_2 + b_3$$

$$N_1 = w_1 x_1 + w_2 x_2 + b_1$$

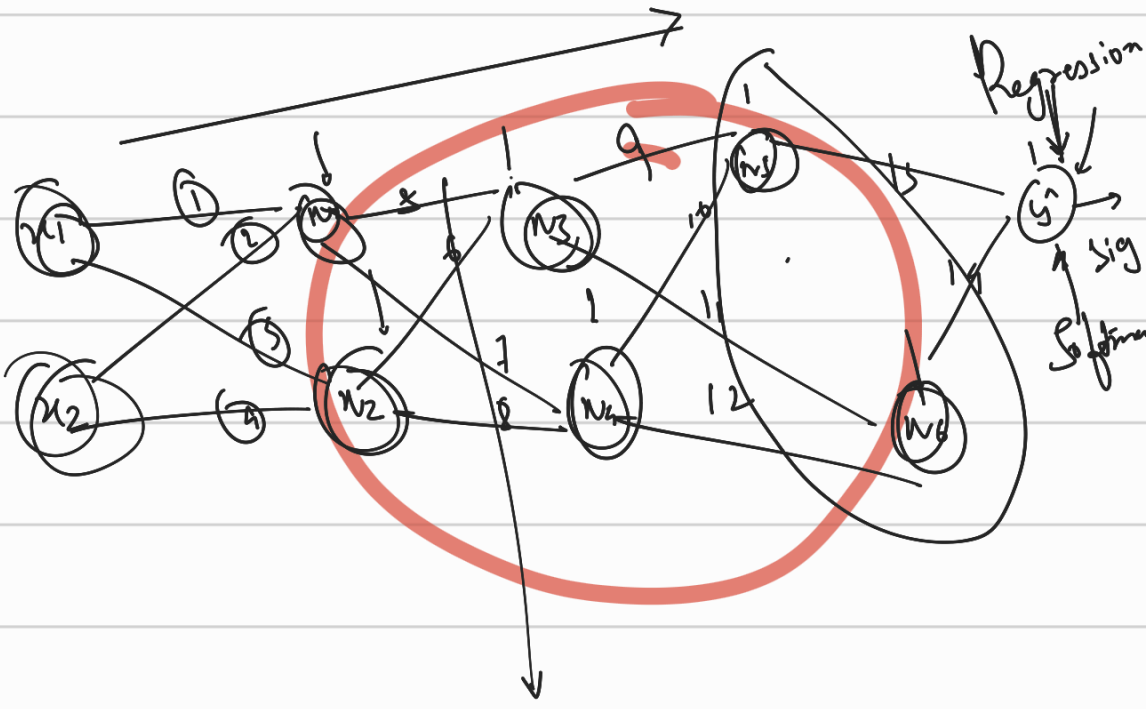
$$N_2 = w_3 x_1 + w_4 x_2 + b_2$$

$$\hat{y} = w_5(w_1 x_1 + w_2 x_2 + b_1) + w_6(w_3 x_1 + w_4 x_2 + b_2) + b_3$$

$$\hat{y} = w_5 w_1 x_1 + w_5 w_2 x_2 + w_5 b_1 + w_6 w_3 x_1 + w_6 w_4 x_2 + w_6 b_2 + b_3$$

$$\Rightarrow \hat{y} = x_1(w_5 w_1 + w_6 w_3) + x_2(w_5 w_2 + w_6 w_4) + w_5 b_1 + w_6 b_2 + b_3$$

$$\Rightarrow \hat{y} = Ax_1 + Bx_2 + C$$



Regression → o/p (No activation) ⟷ MSE

Binary class → o/p (Sigmoid) → Log loss.

MC → o/p (Softmax).