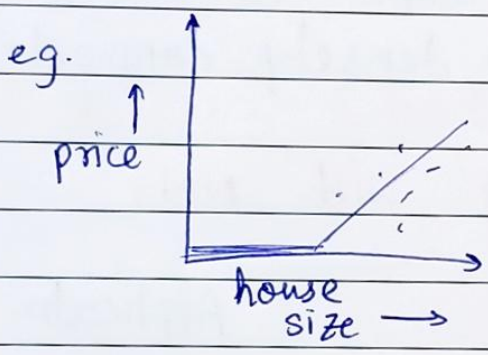
Date : 19/7/24

MON TUE WED THU FRI SAT SUN

☐ ☐ ☐ ☐ ☑ ☐ ☐

NEURAL NETWORKS & DEEP LEARNING (C-1)

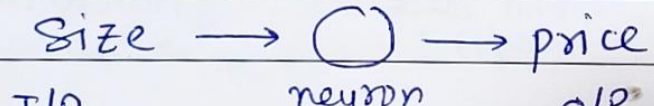
Neural Network



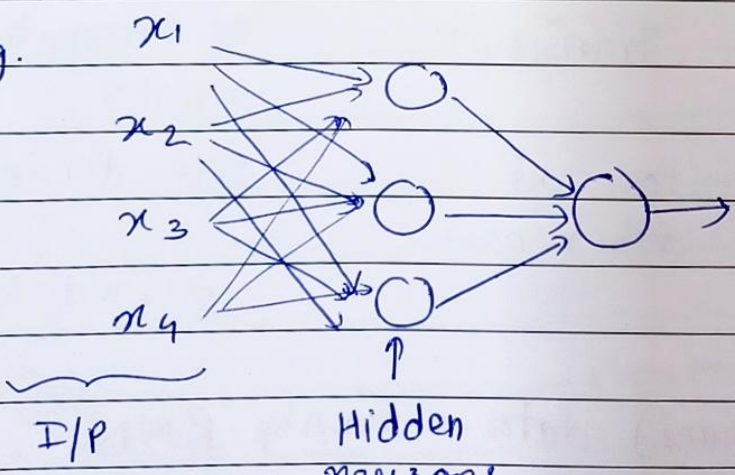
ReLU function.

Rectified Linear Unit

max (–, 0)



eg.



→ (Koala example)

Date : .....
MON TUE WED THU FRI SAT SUN
☐ ☐ ☐ ☐ ☐ ☐ ☐

N.N. has ability to decide by itself what I/P features each of the hidden neuron wants/requires to do its computation.

Every I/P feature conn. to every middle neurons ⇒ densely connected.

Supervised Learning with NN.

I/P	O/P	Applicatn.
Ad, User Info.	Will click (0/1)?	Advertising ↳ Standard NN
Image	object	Tagging ↳ CNN
Audio	Transcript	Sp. recognition ↳ RNN
Img, Radar	Pos. of other cars	Self-driving ↳ Custom/ Hybrid NN

Temporal (time) data → mainly RNN.

↑ Data & ↑ Scale (i.e. more layers of NN)

⇓

↑ Performance

Date : .....

MON TUE WED THU FRI SAT SUN

☐ ☐ ☐ ☐ ☐ ☐ ☐

Structured Data

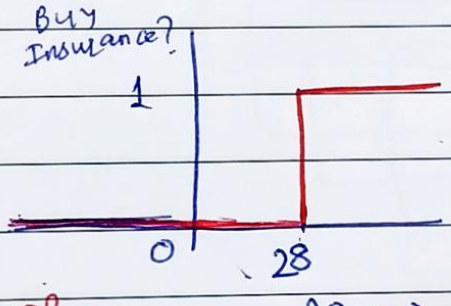
⇒ eg.

Size	#bedrooms	Price
↓	↓	↓

Unstructured Data

⇒ Audio, Image, Text

Basics – Activation functions.

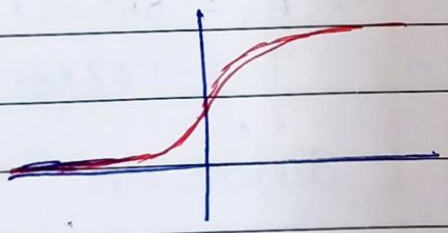


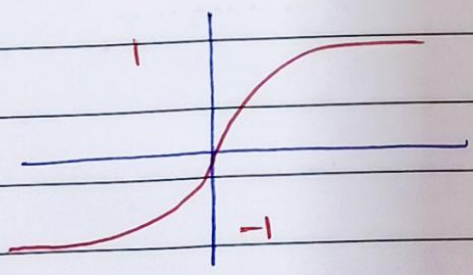
STEP Functn

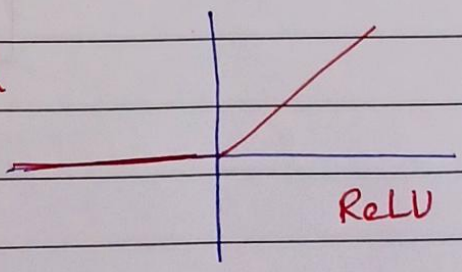
→ eg. People above 28y. will buy & below won't.

But X.

↓







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

sigmoid

Date :.....

MON TUE WED THU FRI SAT SUN

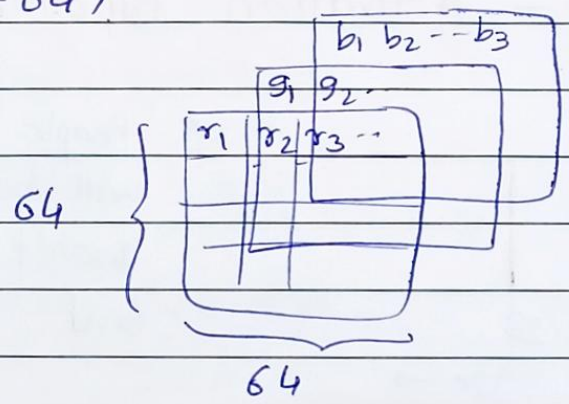
Hidden Layer mostly - ReLU or Leaky ReLU.

O/P → sigmoid mostly.

Binary classification

O/P → 0 or 1

eg. cat image (64 × 64) ⟶ cat OR non-cat



Feature vector

$$X = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ g_1 \\ g_2 \\ \vdots \\ b_1 \\ b_2 \\ \vdots \end{bmatrix} \Rightarrow n_x = 64 \times 64 \times 3 = 12288$$

↑ dimension.

Date :.....
MON TUE WED THU FRI SAT SUN
☐ ☐ ☐ ☐ ☐ ☐ ☐

(x, y) = Training eg.

$x \in \mathbb{R}^{n_x}$, $y = \{0, 1\}$

$m \rightarrow$ # training examples:

$\{(x^1, y^1), (x^2, y^2) \cdots (x^m, y^m)\}$

$$X = \begin{bmatrix} | & | & & | \\ x^1 & x^2 & \cdots & x^m \\ | & | & & | \end{bmatrix} \} n_x$$

$\underbrace{\qquad\qquad}_{m}$

compact notation to store m training examples.

For Y also, stack in cols.

$$Y = [y^1, y^2 \cdots y^m]$$

Date :
MON TUE WED THU FRI SAT SUN

Logistic Regression

→ Supervised ML Algo. used for binary class. tasks.

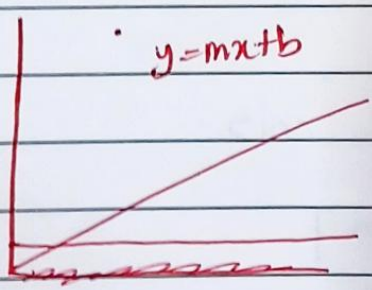
Given I/P x, $\hat{y} = P(y=1|x)$ probability

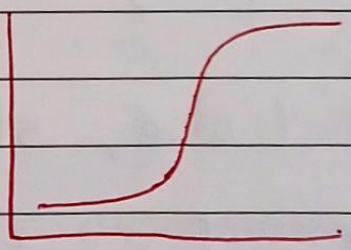
Params: $w \in \mathbb{R}^{n_x}$, $b \in \mathbb{R}$

O/P: $\hat{y} = w^T x + b$? is not bet^n [0,1]

$\therefore \hat{y} = \sigma(w^T x + b)$

z ↳ limits 0 to 1





$y = \frac{1}{1+e^{-(mx+b)}}$

no. of weights (w) = no. of I/P features.

Date :.....
MON TUE WED THU FRI SAT SUN
☐ ☐ ☐ ☐ ☐ ☐ ☐

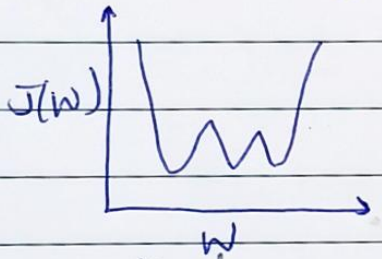
Loss Function:

→ Measures how far an estimated value is from its true value.

→ For general Linear Regr:

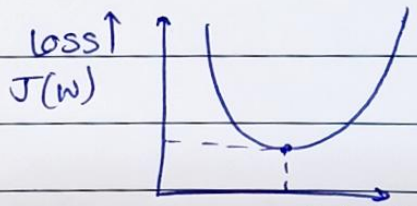
$$\text{Loss} = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

→ But for Logistic Regr. ↑ ✱ don't work.



Gradient Descent

Curve with multiple local minima



optimal Grad. Desc.

Loss fn applied to single training eg.
Cost fn: Cost of parameters.

$$L(\hat{y}, y) = -(y \log \hat{y} + (1-y) \log(1-\hat{y}))$$

$$J(w,b) = \frac{1}{m} \sum_{i=1}^{m} L(\hat{y}^i, y^i)$$