Relatn b/w i/p & o/p → y Assume ⟹ Linear relation

x

y_1 = make - Pty

y_2 = ---- Char

⋮

y_5 =

x_1 = no. of gfs

x_2 = no. of hrs he studi

x_3 = parents ...

Linear Regress

$y = f(x)$ → Linear

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$y_1' = w_{11}x_1 + w_{21}x_2 + w_{31}x_3 + b_0 + b_1$$

$$y_2' = w_{12} \quad w_{22} \quad w_{32} \quad b_2$$

y_3'

$$Y' = w \cdot x + b$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} w_{11} & w_{21} & w_{31} \\ w_{12} & w_{22} & w_{32} \\ w_{13} & w_{23} & w_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$(m \times 1)$ $\quad (m \times n)$ $\quad (n \times 1)$

$$y = g(f(h \dots (x)))$$

$$y = w_1[w_2[w_3 \dots x]]$$

$w_1\ w_2 \dots$ matrix → linear

$$y = \sigma_2\{w_2[\sigma_1(w_1 x)]\}$$

activation function

Bias

$$\underbrace{\begin{bmatrix} w_{11} & w_{21} & w_{31} & b_1 \\ - & - & - & b_2 \\ - & - & - & b_3 \\ b_1 & b_2 & b_3 & \end{bmatrix}}_{W} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix}}_{X}$$

Imple

m training egs → $(X_1, X_2 \dots X_m)$

(x_1, x_2, x_3, x_n)

$Y_1 \quad Y_2 \quad Y_m$

$X_1 \quad x_2 \quad x_3 \dots x_m$

$$\begin{bmatrix} y & & \\ & & \\ & & \end{bmatrix} = \begin{bmatrix} & & \\ & W & \\ & & \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & x_1 \\ x_{21} & x_{22} & x_2 \\ x_{31} & x_{32} & \end{bmatrix}$$

1 $X \to 3\times1$ $\quad Y = 5\times1$

$W \to 6\times3$ $\quad 5\times4$ $\quad W_3$

$6\times1 - W_2 - 4\times6 - 4\times1$

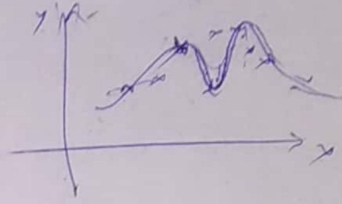
$Y = [\;W_3\;]\,\sigma\,[\;W_2\;]\,\sigma\,[\;W_1\;][X]$ $\quad 6\times5 \quad 5\times3 \quad 3\times1$

100 → 60 + 30 + 10 . results you are using.

↓ error

train (no of layers = 5) test

↓ error ↓↓



<u>Loss</u>

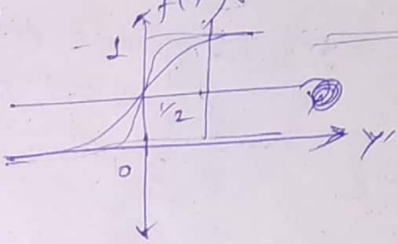
Correct - Y $\quad$ o/p from net → y'

Error $\quad$ Regression → mean sq. error $\frac{1}{n}(Y-Y')^T(Y-Y')$

Class → logistic function :- $f(y')$

$$y = \frac{1}{1-e^{-y/n}}$$

Normalization



<u>Optimisation</u> → Loss minimise by changing the parameters (Ws)

$$W_{i+1} = W_i - \mu\frac{\partial}{\partial W_i}(\text{Loss})$$

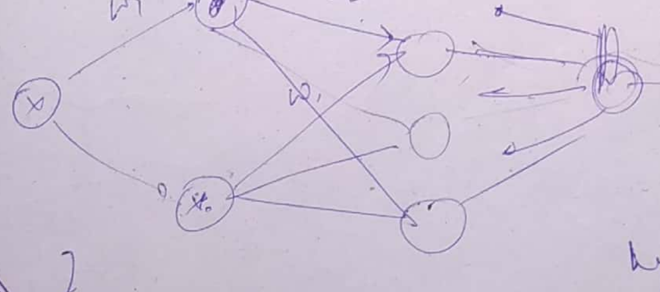
learning rate

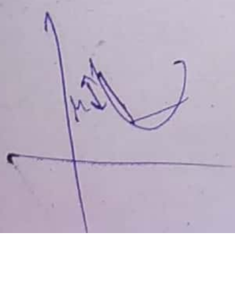
Supp. $\quad y = W_3[\sigma_2(W_2(\sigma_1(W_1(x))))]$ $\quad z_2 \quad z_1$

$$\frac{\partial}{\partial W_1} = \frac{\partial(\text{err})}{\partial W_3}\cdot\frac{\partial W_3}{\partial z_2}\cdot\frac{\partial z_2}{\partial W_2}\cdot\frac{\partial W_2}{\partial z_1}\cdot\frac{\partial z_1}{\partial W_1}$$

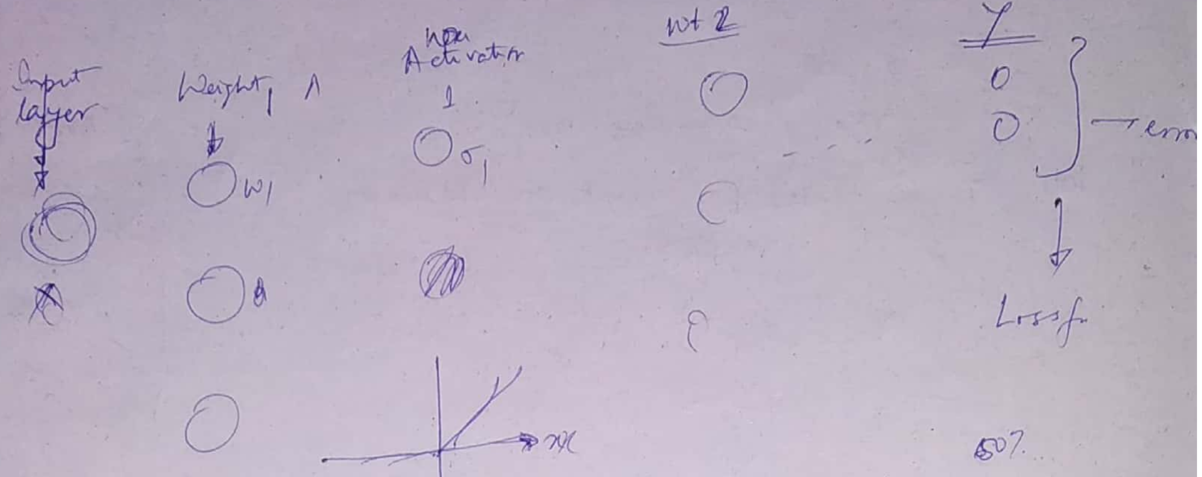
Backpropagation

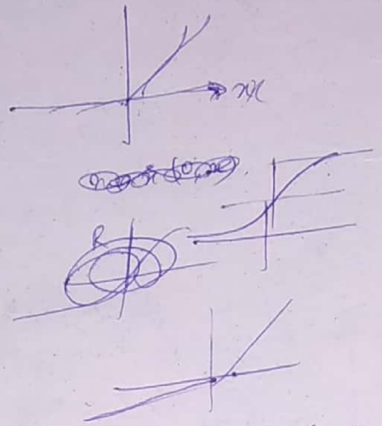
Gradient descent



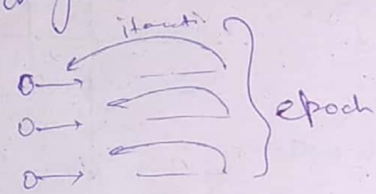


hyperpara ← no. of layers
" " neurons in a layer
μ



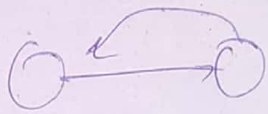


1. Batch gradient descent

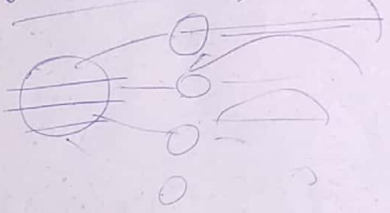


2. Stochastic descent

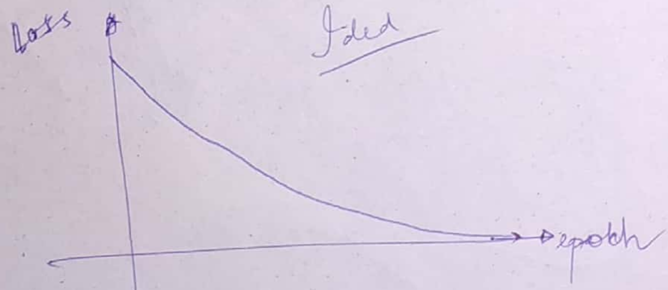
epoch = Iteration



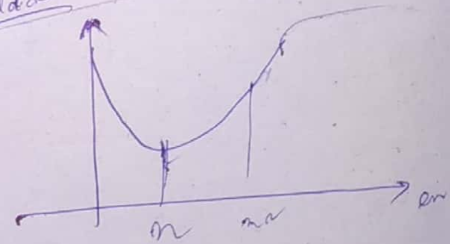
3 Stochastic batch

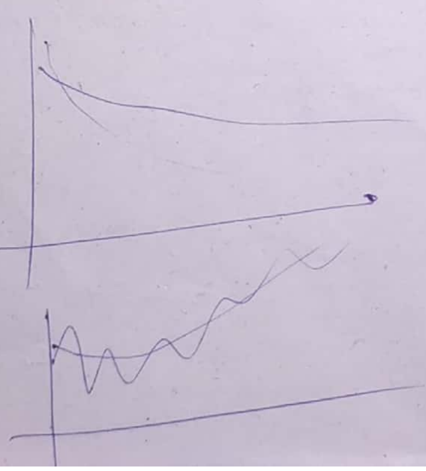


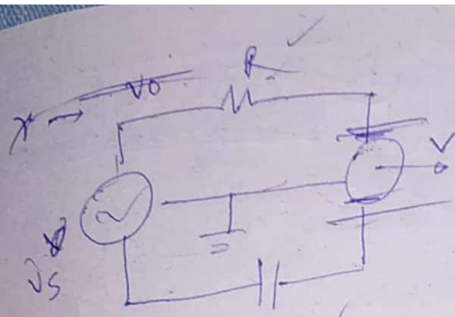
Ideal



Validation error







$\frac{1}{j\omega C}$ (ref)

V_s $\underset{x_2}{R}$, $\underset{x_3}{C}$, $\underset{x_4}{f}$

Measured |V| Mea ∠V

Real |V| Real ∠V

Error |V| — y_1 Error ∠V — y_2