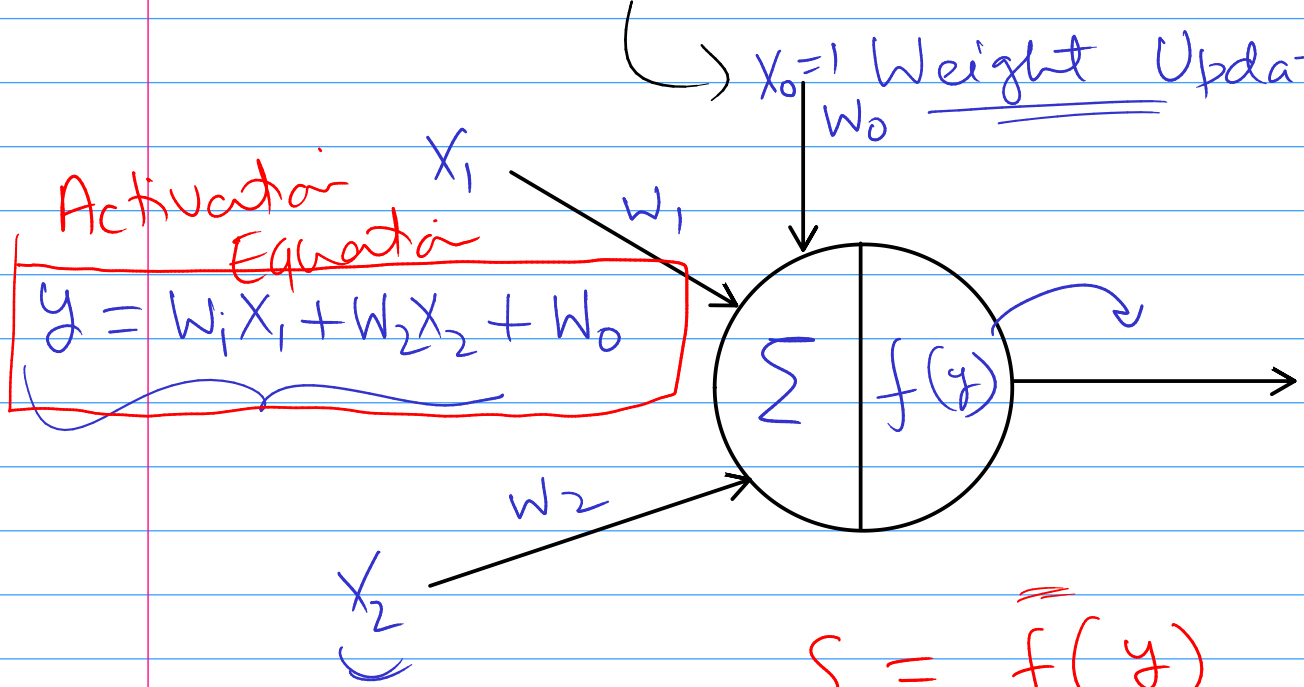
⇒ Tensor Operations:

activation | How it is calculated?

relu | how it is applied?

Gradient Decent Method.

↳ $x_0 = 1$ Weight Update ?



Activation Equation

$$y = w_1x_1 + w_2x_2 + w_0$$

$$S = f(y)$$

↳ activation function.

$$S = relu(x_1w_1 + x_2w_2 + w_0)$$

MNIST Handwritten Digits ✓

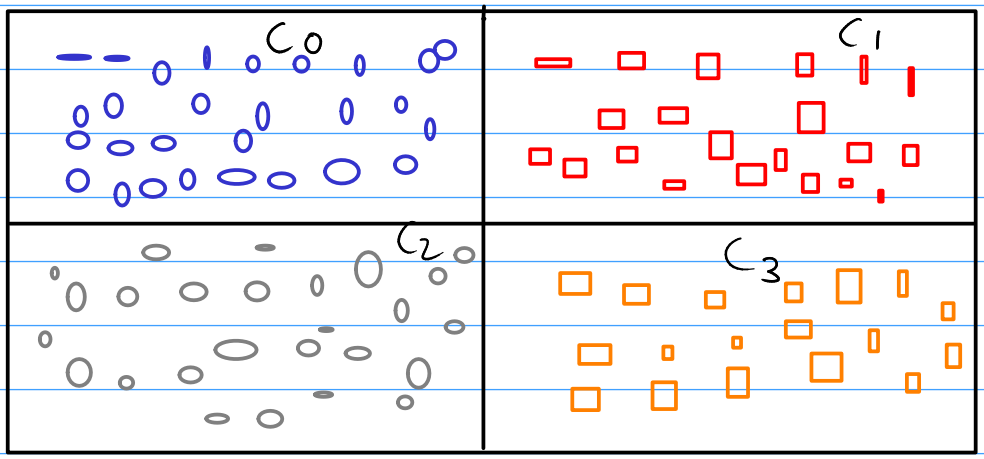
↳ modified National Institute of Standards & Technology

Can you use Softmax?

↳ activation function.

classification → 2- class problem

→ Multiclass problem..



multiclass problem:

$$S = e^{y} \rightarrow e^{x_1 w_1 + x_2 w_2 + w_0}$$

$$S_i = \frac{e^{y_i}}{\sum e^{y_i}}$$

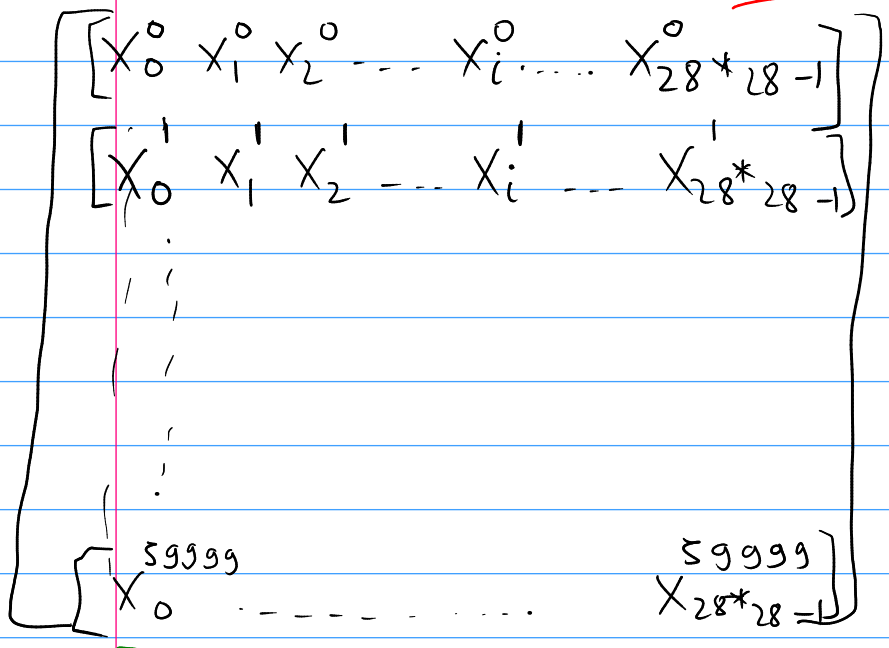
$$\sum S_i = 1$$

$$\sum_{i=0}^{3} S_i = \frac{e^{y_0} + e^{y_1} + e^{y_2} + e^{y_3}}{\sum e^{y_i}}$$

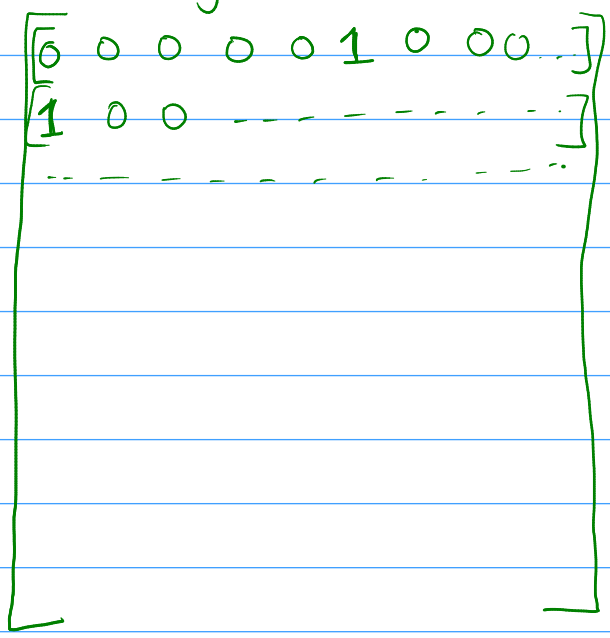
$$= 1$$

(60000, 28, 28) 3 channels to_categorical 28 * 28

↳ Image Data



Categorical



'cat'

Data → encoding

Labels.

	Cat	Dog	Goat	
Data →	1	0	0	= 1
Data2	0	1	0	

'cat' x 5

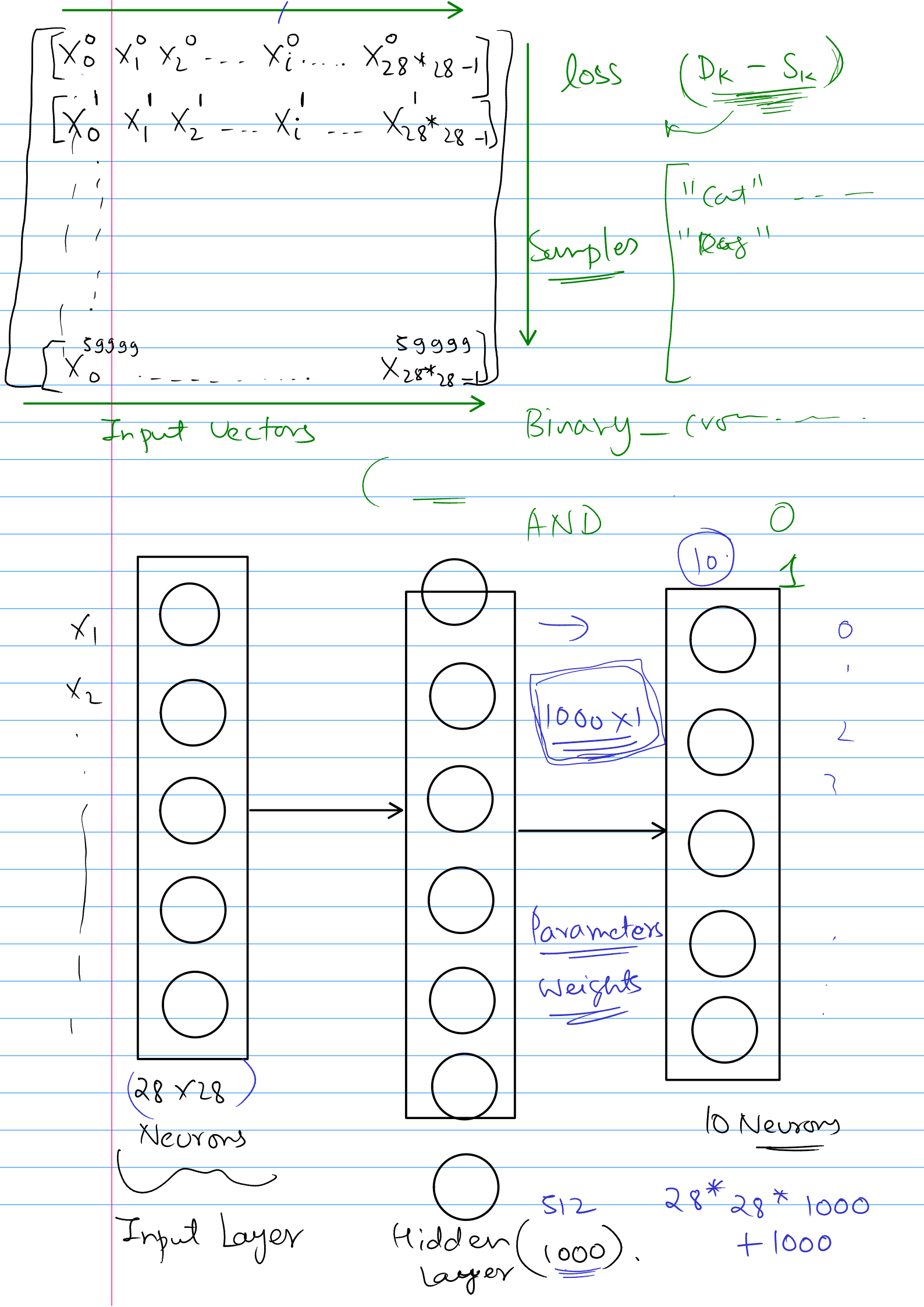
cat cat ... 5x

Categorical Data

loss function for ↳ categorical Data

"Categorical_crossentropy" } For Classification Problem.

"Binary_crossentropy".


loss
$(D_k - S_{ik})$
Samples
"Cat"
"Dog"
Input Vectors
Binary_
AND
0
1
10
X_1
X_2
1000 x1
0
1
2
3
Parameters
Weights
28 x 28
Neurons
Input Layer
Hidden Layer
512
(1000)
10 Neurons
$28*28*1000$
$+1000$

$W =$

$$\begin{bmatrix} W_0^0 & W_1 & \cdots & W_i & W_{28*28} \\ \vdots & & & & \\ W_0^{1000} & W_1^{1000} & \cdots & & W_{28*28} \end{bmatrix}$$

$X^T W$

$\text{Relu}(X^T W)$

28×28 $\quad$ 28×28

$n \times m$ $\quad$ $n \times 1000$

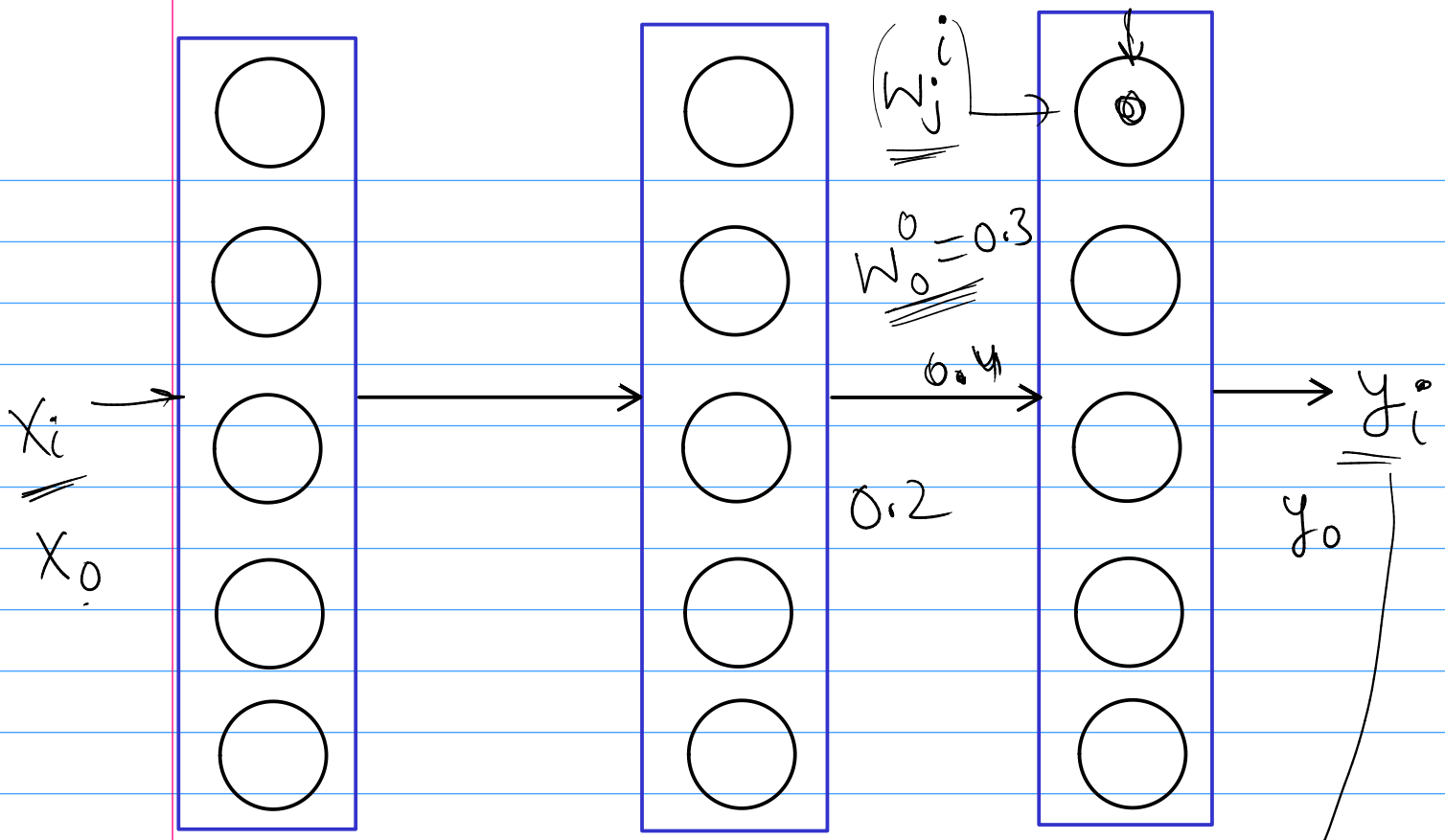
$1000 \times m$

10×1

28×28

Softmax

$$S_i = \frac{e^{y_i}}{\sum e^{y_i}}$$



1.1

0.9 ↓ (minimize) → 0.7

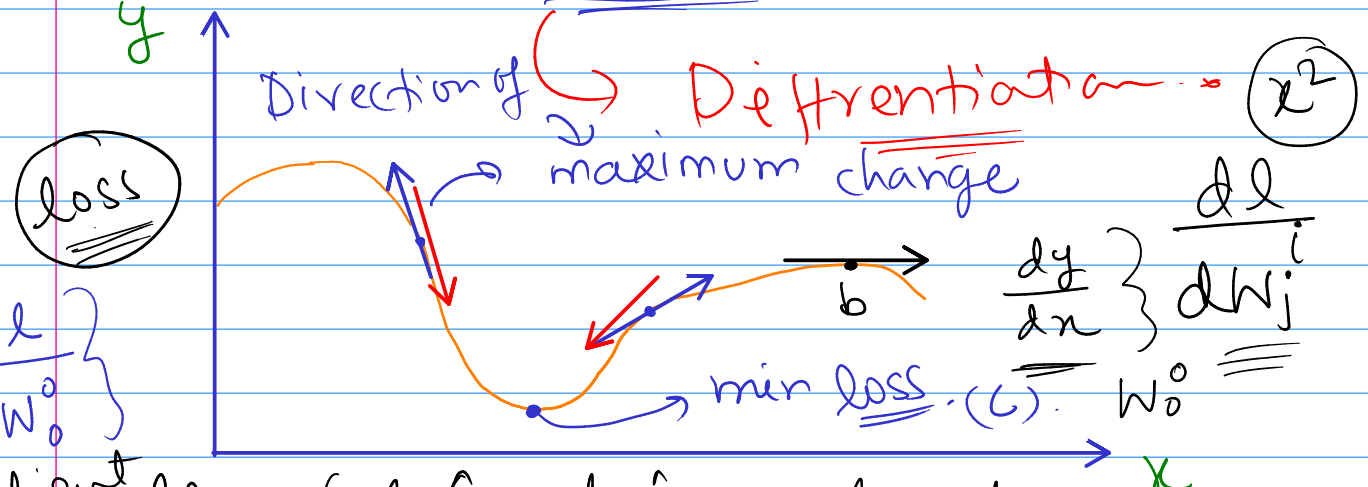
$loss = (y_i - y_{pred})$ → Calculated in forward pass.

$loss = f(y_i - y_{pred})$

100000

⇒ Weight Updates :-

⇒ Stochastic Gradient Decent Method :



Gradient $\frac{dl}{dW_j^i} = -\left(\frac{d}{dW_0^0}\hat{W}_{0t}^0 \; \frac{d}{dW_1^0}\hat{W}_{1t}^0 \cdots \frac{d}{dW_n^0} + \frac{d}{dW_0^1} + \cdots\right) l$

$$y = x^2 + 2x + 1$$

$$\frac{dy}{dx} = 2x = 0$$

$x = 0$

$$\frac{dy}{dx} = \tan\theta = m \simeq \infty$$

$\theta = 90^\circ$