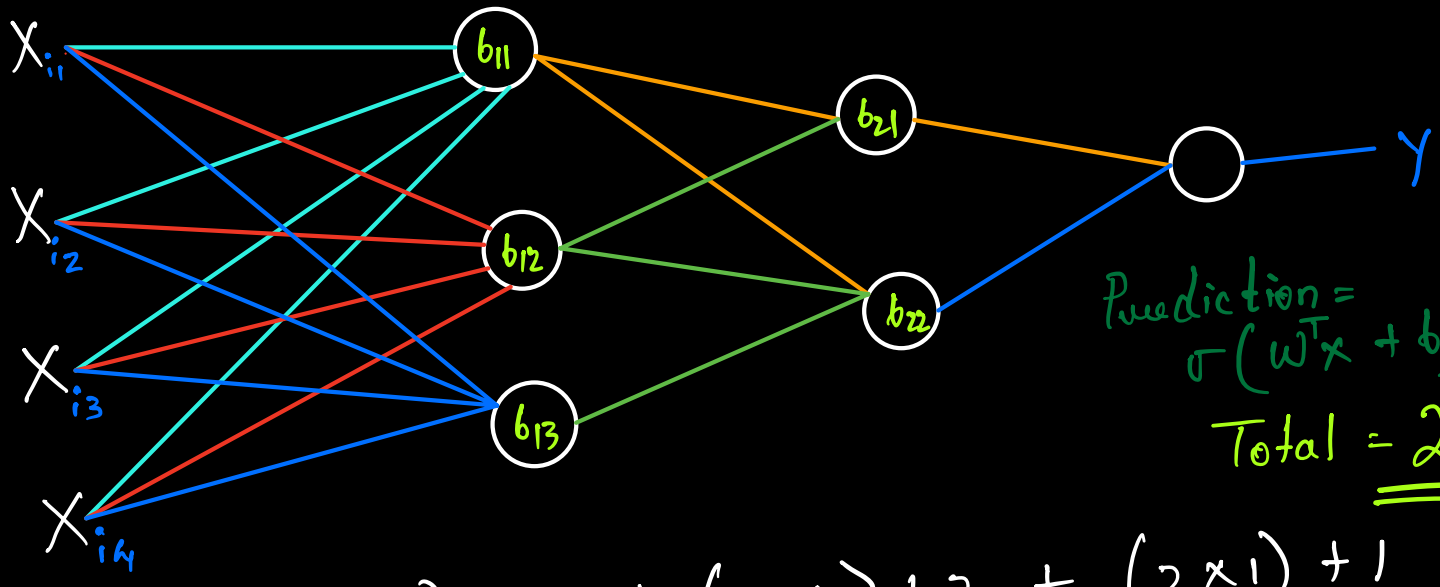
Today's Agenda

1) Recap of OOPs Project
2) MLP → Internal Maths (Linear Algebra)
   → Practical (MNIST) → Classification

Multi Layer Perceptron

Non - Linear

Parameters = Weights + Bias



Prediction = $\sigma(W^T x + b)$

Total = 26

$= (4 \times 3) + 3 + (3 \times 2) + 2 + (2 \times 1) + 1$

$= 15 + 8 + 3$

$= 26$

Layer 1

$$\begin{bmatrix} w'_{11} & w'_{12} & w'_{13} \\ w'_{21} & w'_{22} & w'_{23} \\ w'_{31} & w'_{32} & w'_{33} \\ w'_{41} & w'_{42} & w'_{43} \end{bmatrix} \begin{bmatrix} x_{i1} \\ x_{i2} \\ x_{i3} \\ x_{i4} \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \end{bmatrix}$$

$4\times3 \xrightarrow{T} 3\times4 \qquad 4\times1 \qquad 3\times1$

$= (m\times n)\ (n\times k) \quad = 3\times1 \ + \ 3\times1$

$= (m\times k) \quad = 3\times1$

$$= \begin{bmatrix} w'_{11}\cdot x_{i1} + w'_{21}\cdot x_{i2} + w'_{31}\cdot x_{i3} + w'_{41}\cdot x_{i4} \\ w'_{12}\cdot x_{i1} + \\ w'_{13}\cdot x_{i1} \end{bmatrix}$$

$$= \begin{bmatrix} \quad \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \end{bmatrix}$$

$\Downarrow$

$a^{[1]}$

$$= \begin{bmatrix} w_{11}\cdot x_{i1} + w_{21}\cdot x_{i2} + w_{31}\cdot x_{i3} + w_{41}\cdot x_{i4} + b_{11} \\ + b_{22} \\ + b_{33} \end{bmatrix}$$

$$= \sigma\left(\begin{bmatrix} o_{11} \\ o_{12} \\ o_{13} \end{bmatrix}\right)$$

Layer 2

$$\begin{bmatrix} w^2_{11} & w^2_{12} \\ w^2_{21} & w^2_{22} \\ w^2_{31} & w^2_{32} \end{bmatrix} \cdot \begin{bmatrix} o_{11} \\ o_{12} \\ o_{13} \end{bmatrix} + \begin{bmatrix} b_{21} \\ b_{22} \end{bmatrix}$$

$3\times 2 \quad 3\times 1 \quad 2\times 1$

$\longrightarrow 2\times 3$

$$= \begin{bmatrix} w^2_{11} \cdot o_{11} + w^2_{21} \cdot o_{12} + w^2_{31} \cdot o_{13} + b_{21} \\ \end{bmatrix}$$

$$= \sigma\left(\begin{bmatrix} o_{21} \\ o_{22} \end{bmatrix}\right) = a^{[2]}$$

Layer 3

$$\begin{bmatrix} w_{11} \\ w_{21} \end{bmatrix}^T \begin{bmatrix} o_{21} \\ o_{22} \end{bmatrix} + \begin{bmatrix} b_{31} \end{bmatrix}$$

$$= \begin{bmatrix} w_{11}\, o_{21} + w_{21}\, o_{22} + b_{31} \end{bmatrix}$$

$$= \sigma\left(\begin{bmatrix} o_{31} \end{bmatrix}\right) \Rightarrow a^{[3]}$$

$$a^{[1]} = \sigma\left(a^{[0]} \cdot w^{[1]} + b^{[1]}\right)$$

$$a^{[2]} = \sigma\left(a^{[1]} \cdot w^{[2]} + b^{[2]}\right)$$

$$a^{[3]} = \sigma\left(a^{[2]} \cdot w^{[3]} + b^{[3]}\right)$$

$$= \sigma\Big(\sigma\Big(\big(\sigma\big(a^{[0]} \cdot w^{[1]} + b^{[1]}\big)\big) \cdot w^{[2]} + b^{[2]}\Big) w^{[3]} + b^{[3]}\Big)$$

$$= a^{[3]}$$

Forward Propagation

1 pass = F.P + B.P

1 F.P

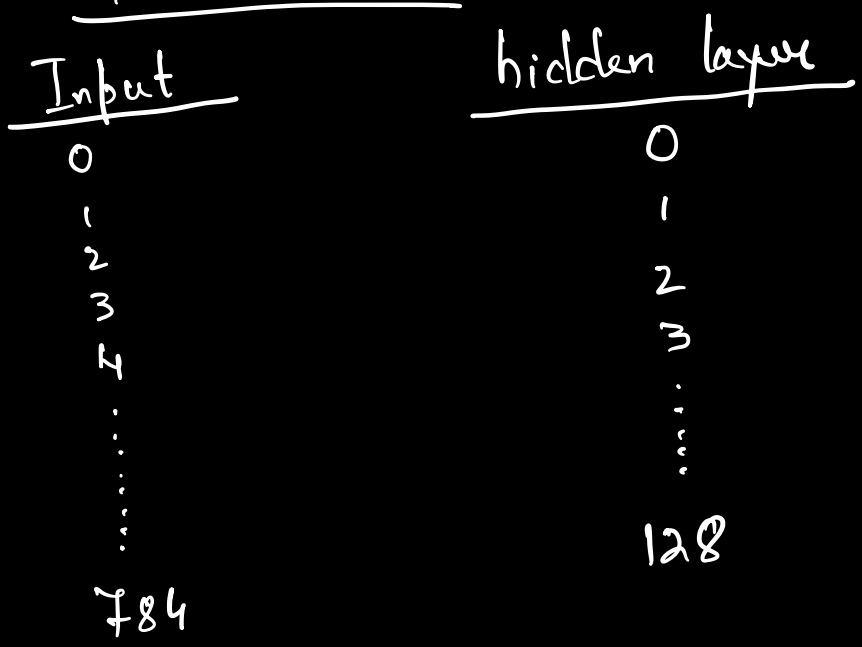
Input → [B.B] → Output

Layer Understanding

1) Dense → Linear, Hidden, Fully Connected
2) Flatten layer
   ↳ Reshape to 1D

Input → image (2D) → 28 X 28 → 784 → (784, 1) (1D)

Network



Node Selection Number

Hyperparameter

Thumb Rule :- 2^n

Software + Hardware (GPU)

2^n

SIMD

GPU Memory allocation optimal

Nodes, Batches, Kernels

2^n

91

128

64

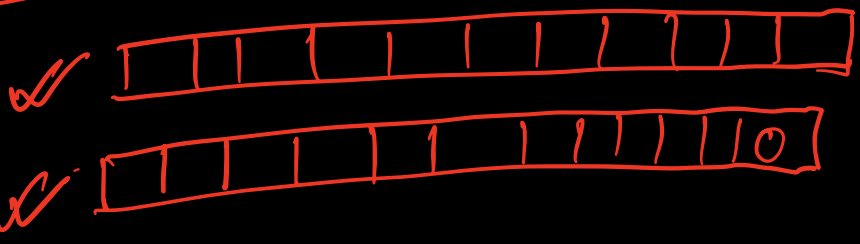
32

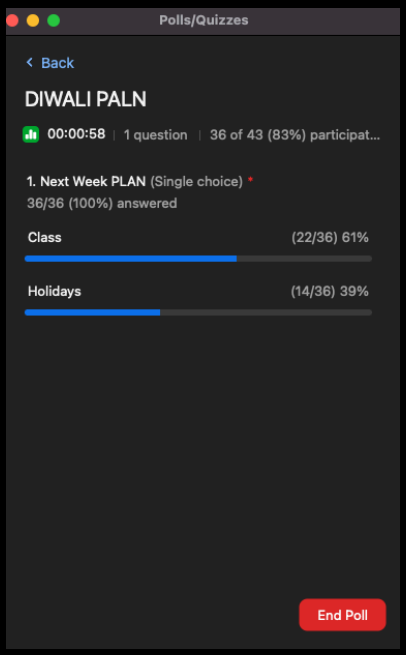
256

127

130

128




Polls/Quizzes
Back
DIWALI PALN
00:00:58 | 1 question | 36 of 43 (83%) participat...
1. Next Week PLAN (Single choice) *
36/36 (100%) answered
Class
(22/36) 61%
Holidays
(14/36) 39%
End Poll