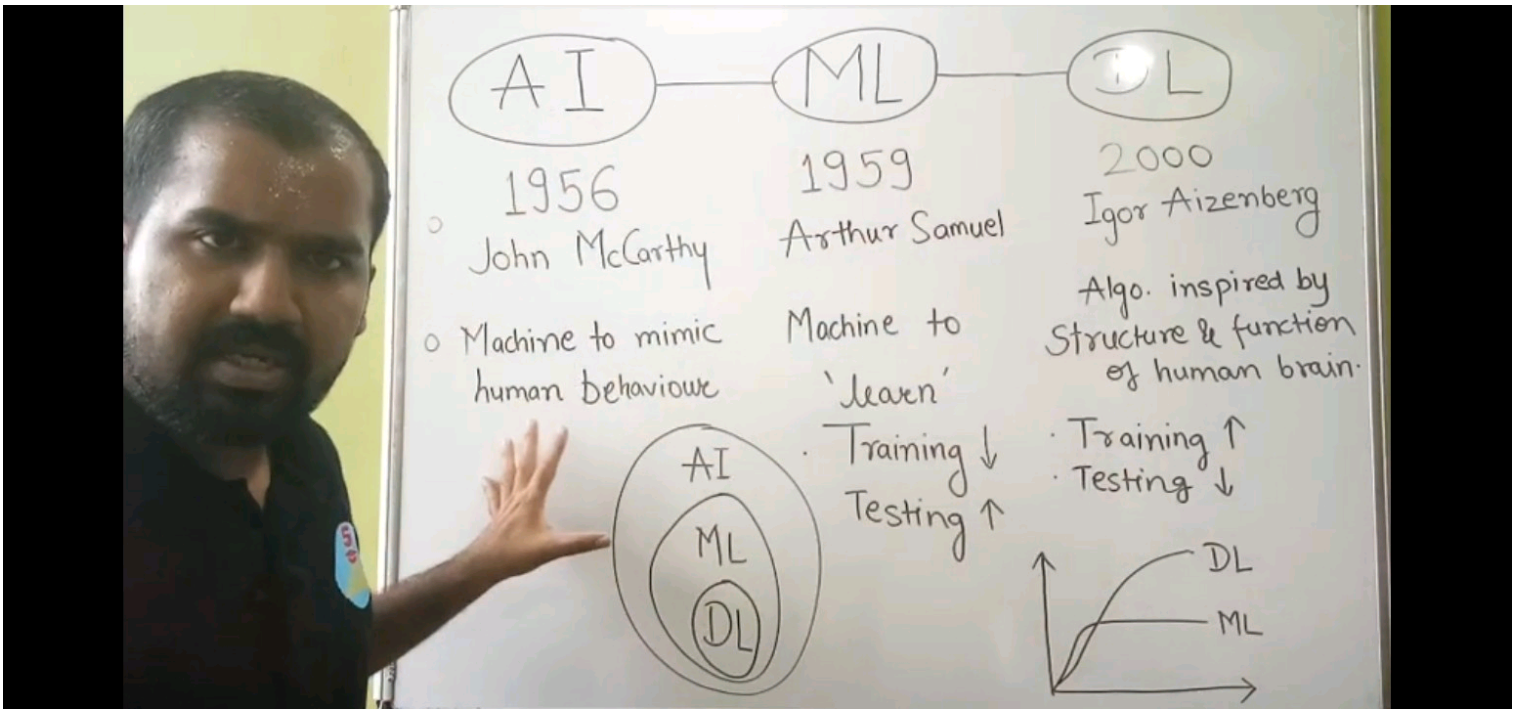
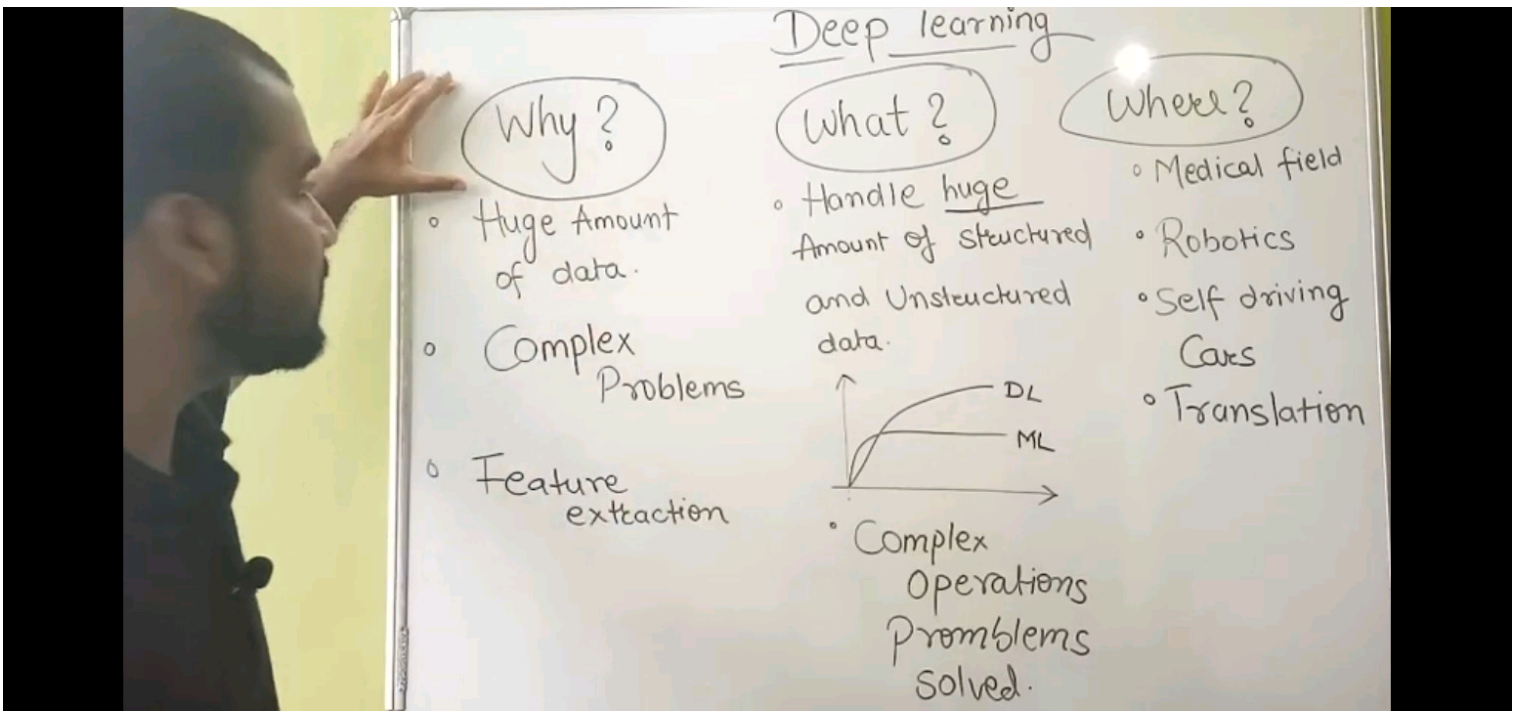
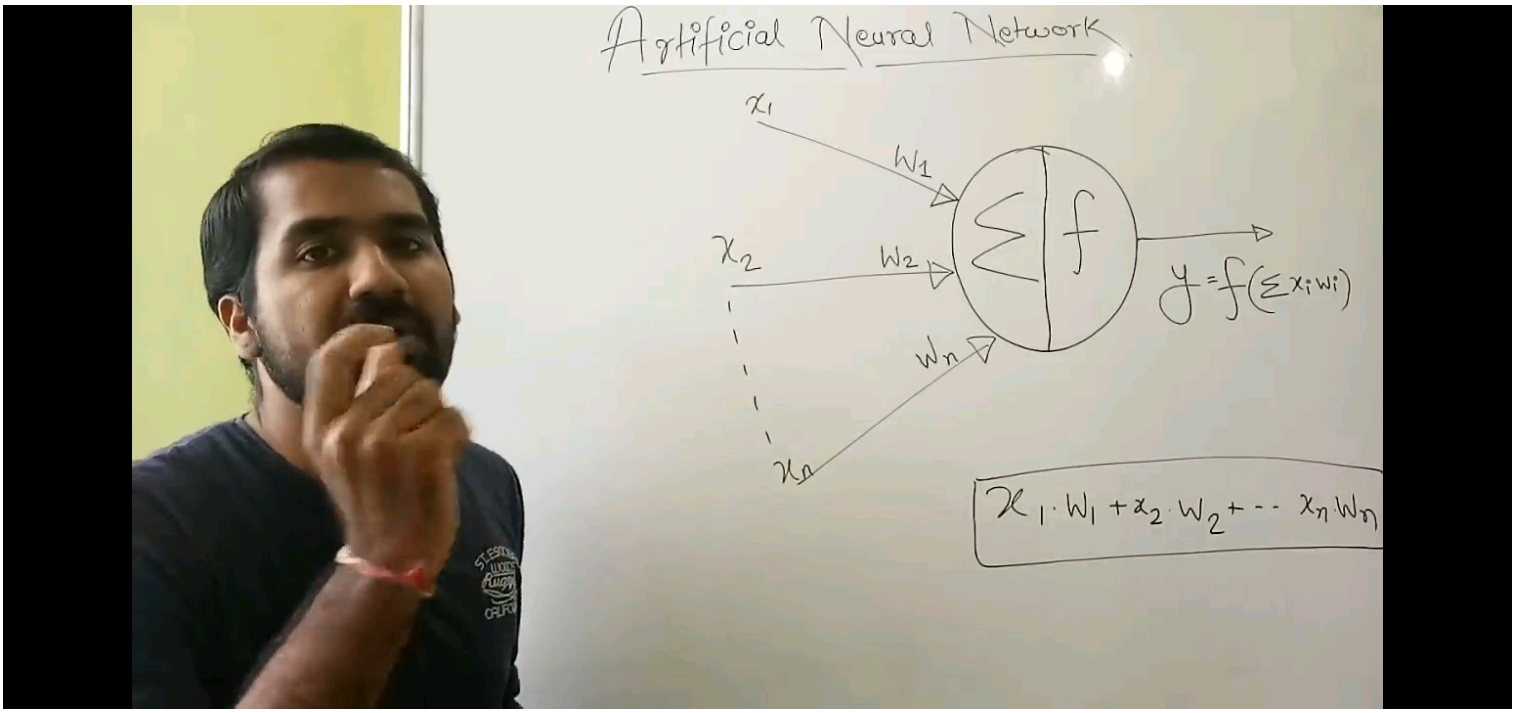
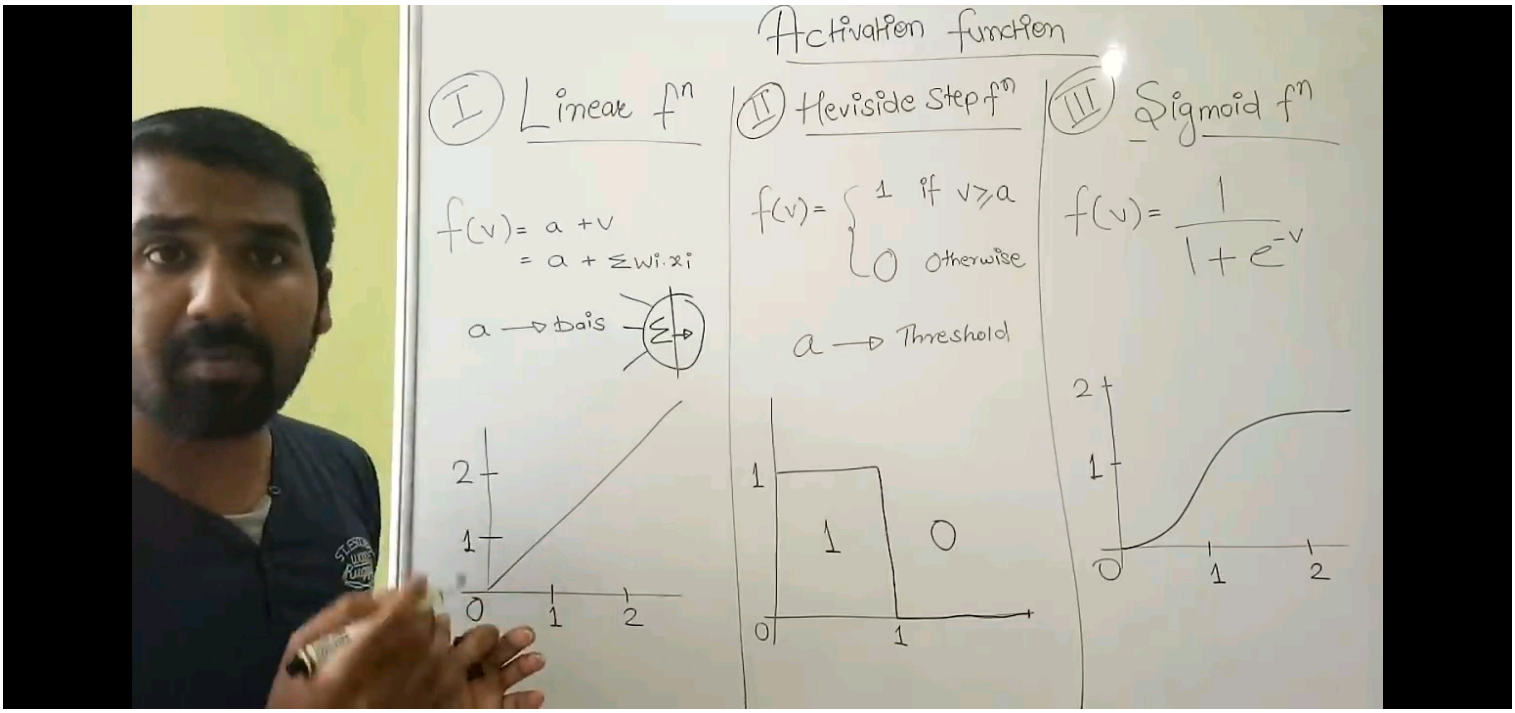
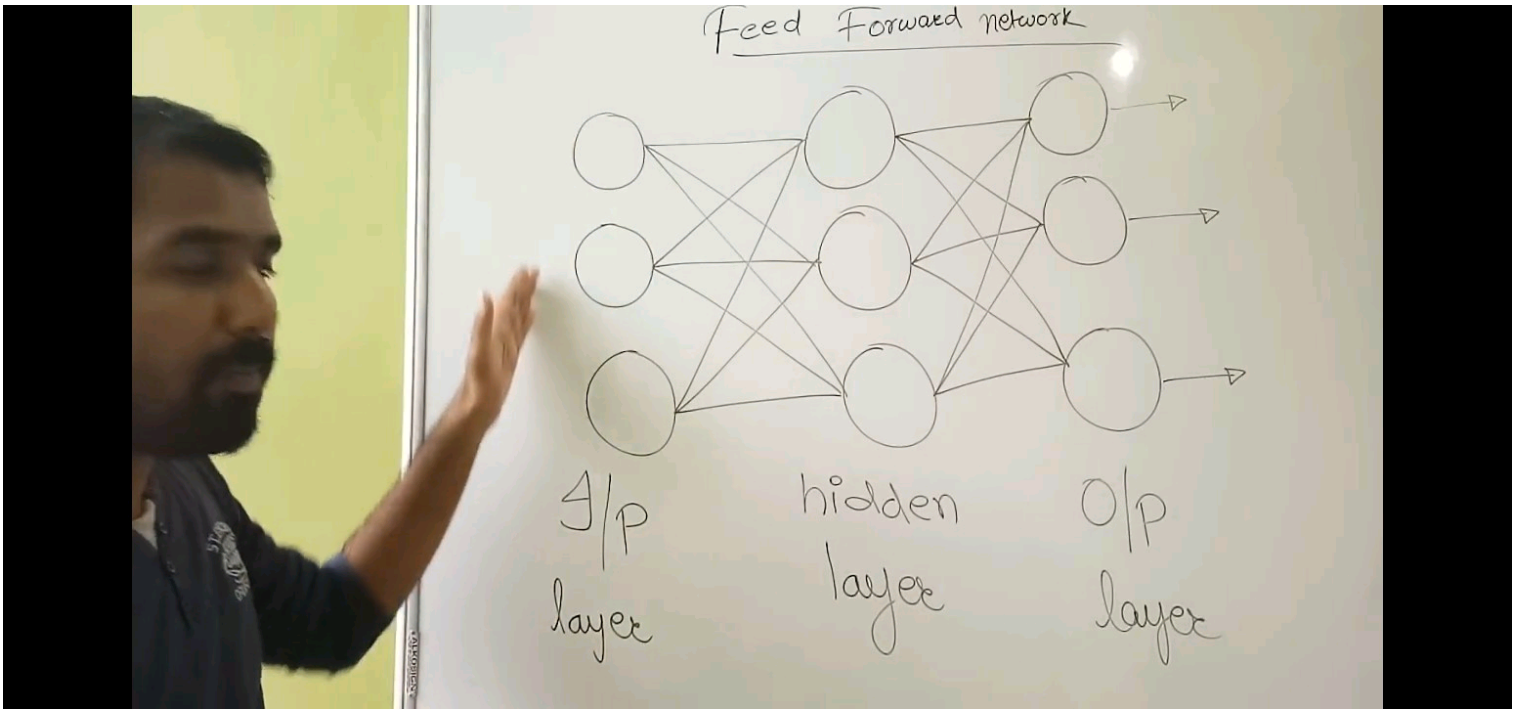
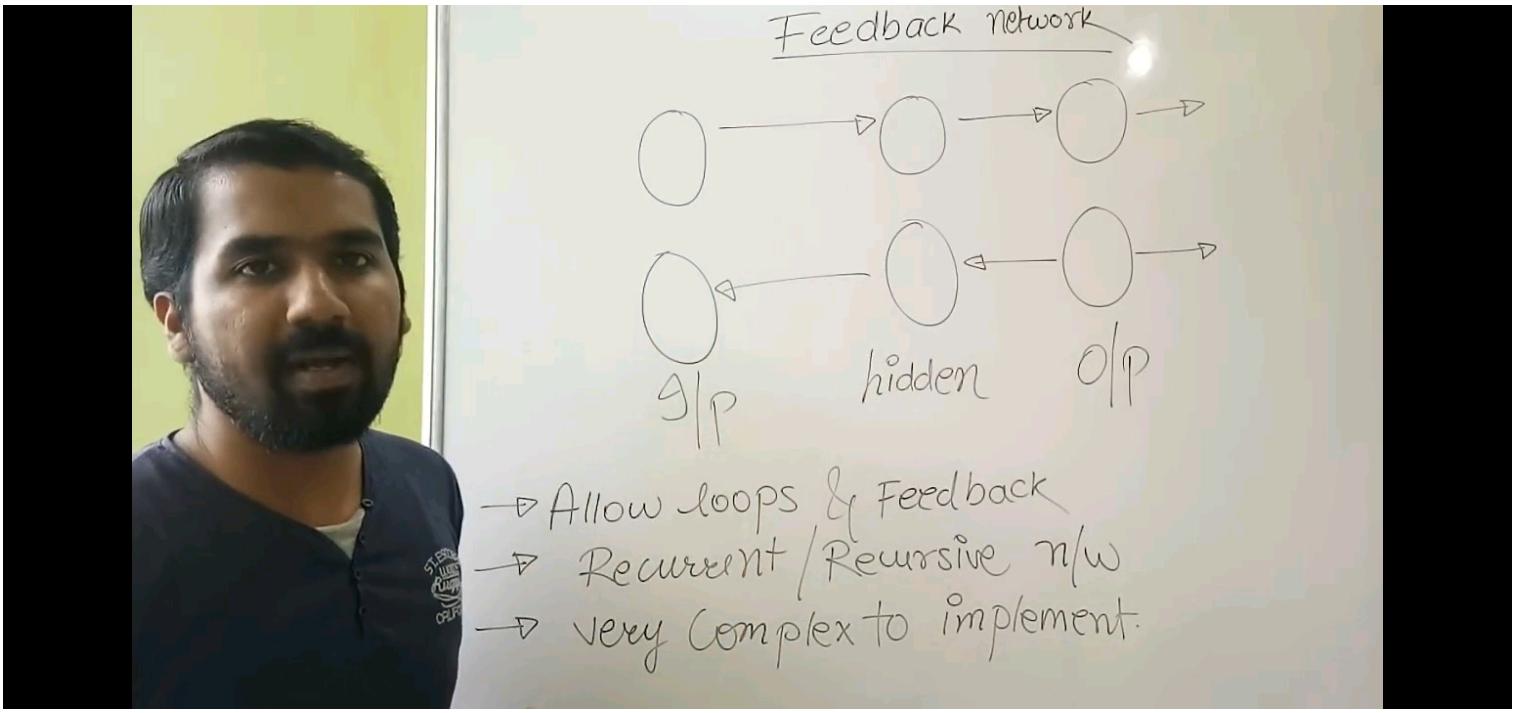

AI
ML
DL
1956
John McCarthy
Machine to mimic human behaviour
1959
Arthur Samuel
Machine to 'learn'
Training ↓
Testing ↑
2000
Igor Aizenberg
Algo. inspired by Structure & function of human brain.
Training ↑
Testing ↓
AI
ML
DL
DL
ML

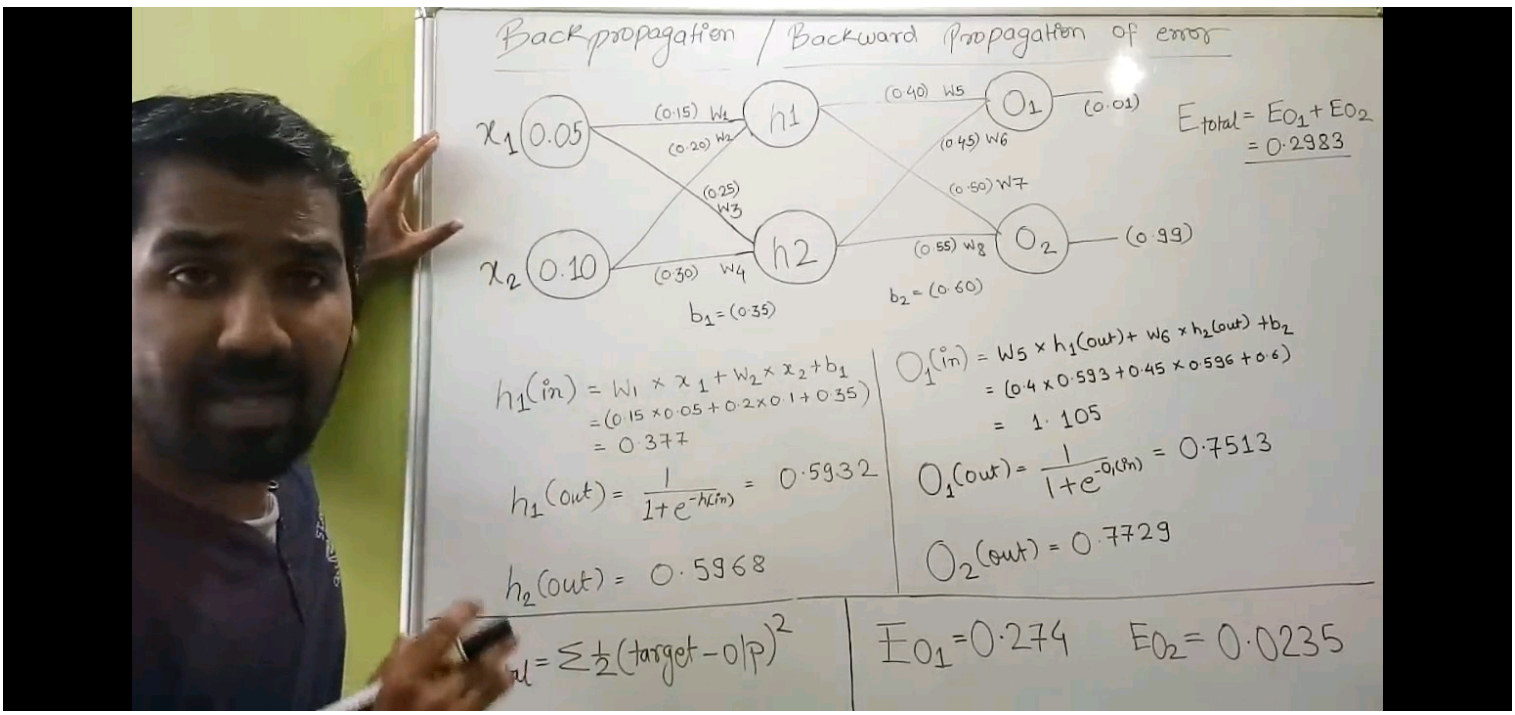

Deep learning
Why ?
Huge Amount of data.
Complex Problems
Feature extraction
What ?
Handle huge Amount of structured and Unstructured data.
DL
ML
Complex Operations Promblems solved.
Where ?
Medical field
Robotics
Self driving Cars
Translation

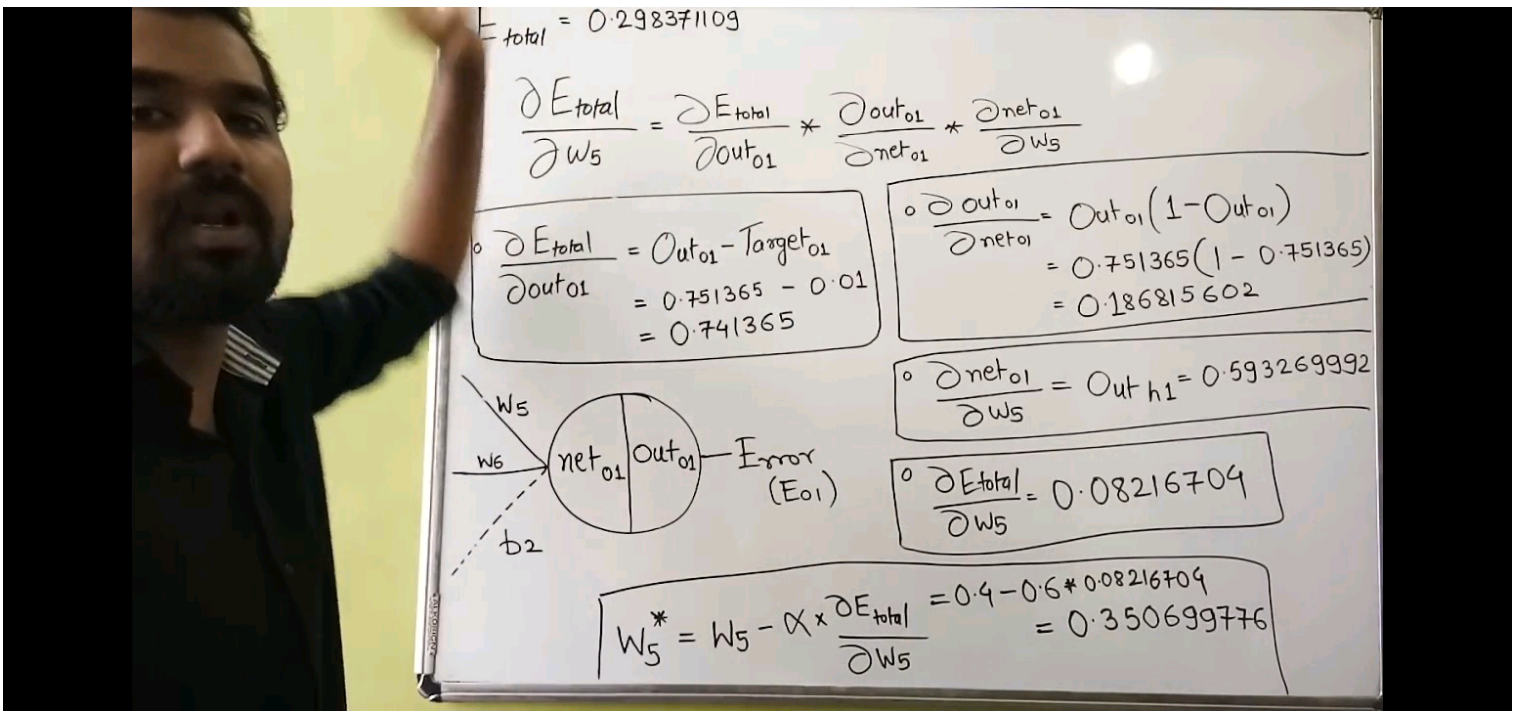

Artificial Neural Network
x_1
w_1
x_2
w_2
w_n
x_n
Σ
f
$y = f(\Sigma x_i w_i)$
$x_1 \cdot w_1 + x_2 \cdot w_2 + \cdots x_n w_n$

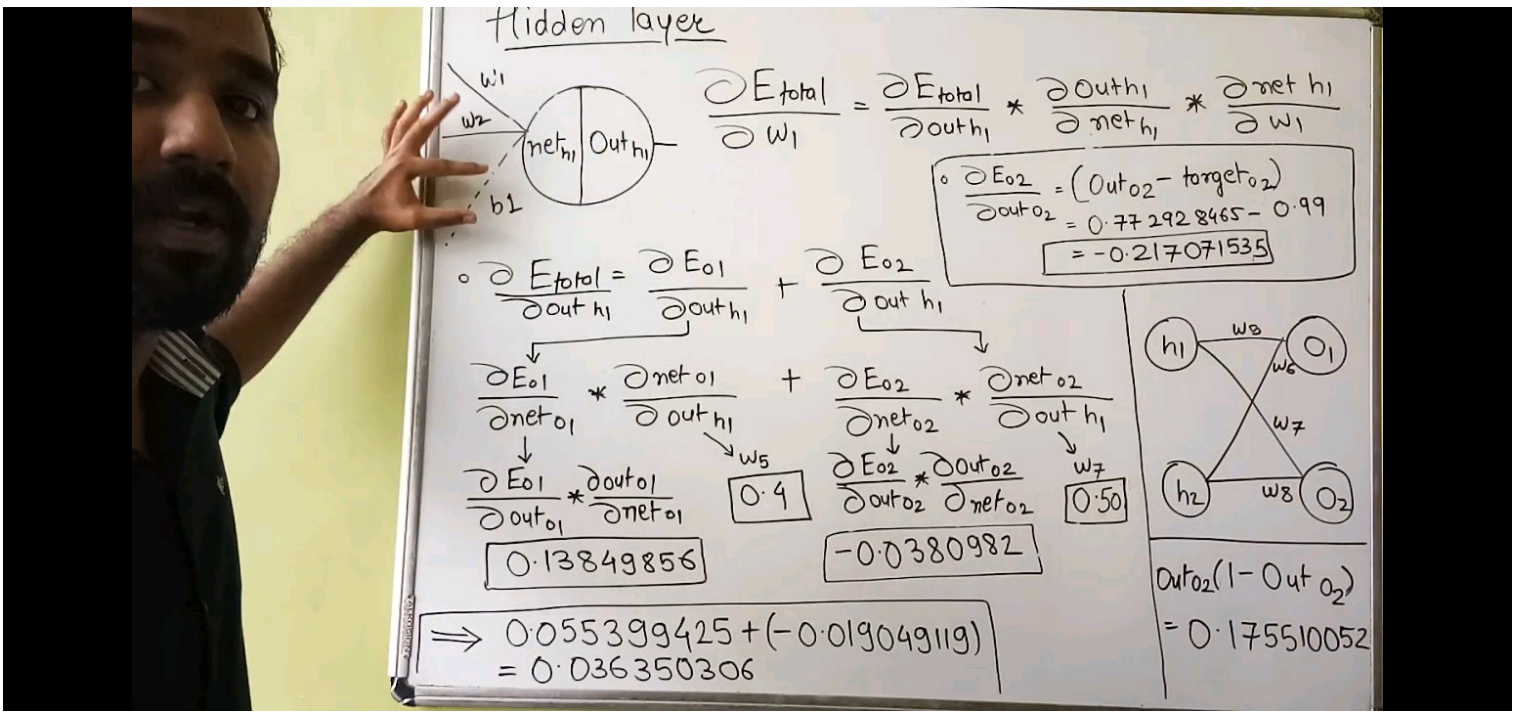

Activation function
I Linear fn
f(v) = a + v
= a + Σwi·xi
a → bais
2
1
0 1 2
II Heaviside Step fn
f(v) = 1 if v ≥ a
0 otherwise
a → Threshold
1
0
1
1 0
III Sigmoid fn
f(v) = 1 / 1 + e^-v
2
1
0 1 2

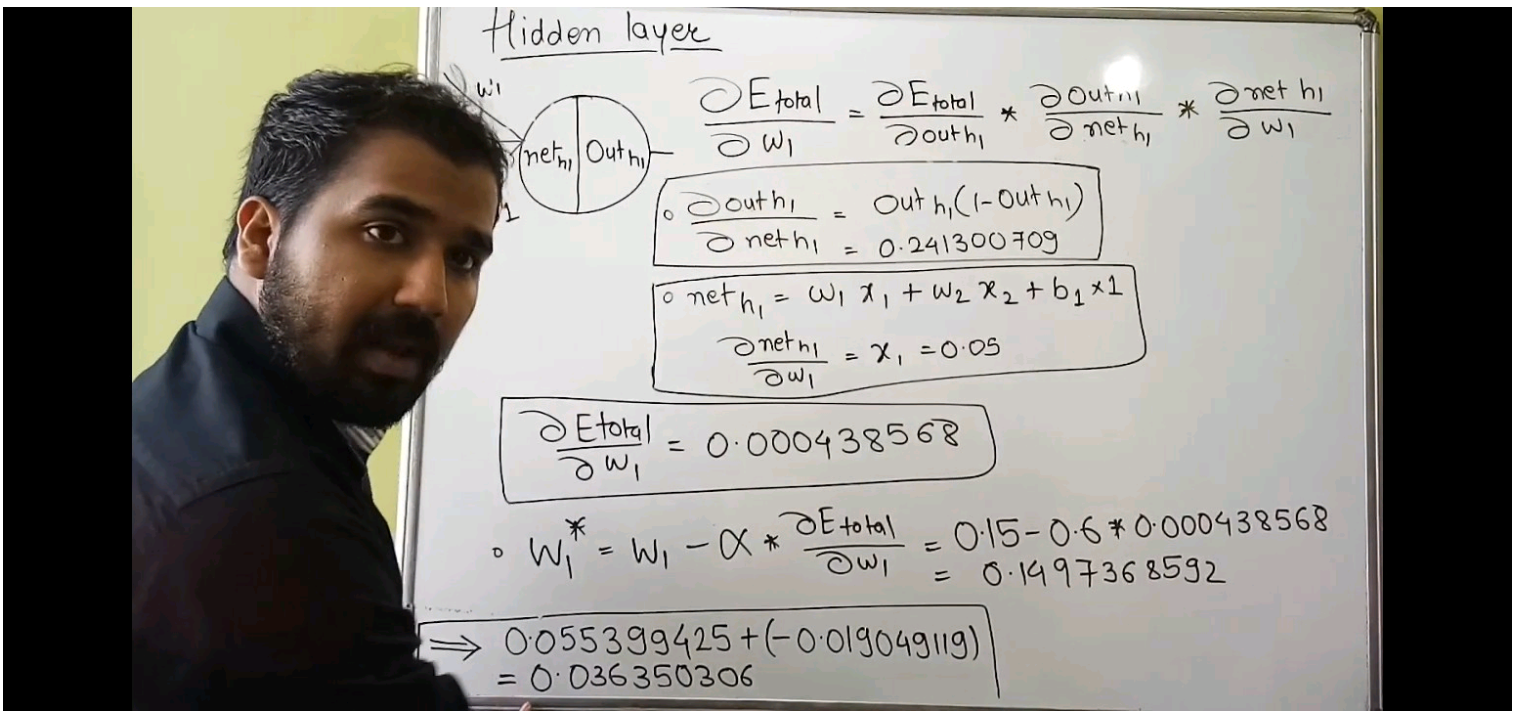

Feed Forward network
I/p
layer
hidden
layer
O/p
layer


Feedback network
i/p
hidden
o/p
→ Allow loops & Feedback
→ Recurrent / Recursive n/w
→ very Complex to implement.


Back propagation / Backward Propagation of error
x1 0.05
x2 0.10
(0.15) W1
(0.20) W2
(0.25) W3
(0.30) W4
h1
h2
b1 = (0.35)
(0.40) W5
(0.45) W6
(0.50) W7
(0.55) W8
O1
O2
(0.01)
(0.99)
b2 = (0.60)
E total = EO1 + EO2
= 0.2983
h1(in) = W1 × x1 + W2 × x2 + b1
= (0.15 × 0.05 + 0.2 × 0.1 + 0.35)
= 0.377
h1(out) = 1 / 1+e^-h1(in) = 0.5932
h2(out) = 0.5968
O1(in) = W5 × h1(out) + W6 × h2(out) + b2
= (0.4 × 0.593 + 0.45 × 0.596 + 0.6)
= 1.105
O1(out) = 1 / 1+e^-O1(in) = 0.7513
O2(out) = 0.7729
= Σ ½ (target − o/p)^2
EO1 = 0.274
EO2 = 0.0235


E total = 0.298371109
∂E total / ∂W5 = ∂E total / ∂out01 * ∂out01 / ∂net01 * ∂net01 / ∂W5
∂E total / ∂out01 = Out01 - Target01
= 0.751365 - 0.01
= 0.741365
∂out01 / ∂net01 = Out01 (1 - Out01)
= 0.751365 (1 - 0.751365)
= 0.186815602
∂net01 / ∂W5 = Out h1 = 0.593269992
∂E total / ∂W5 = 0.08216704
W5
W6
net01
Out01
Error (E01)
b2
W5* = W5 - α x ∂E total / ∂W5
= 0.4 - 0.6 * 0.08216704
= 0.350699776


Hidden layer
w1
w2
net h1
Out h1
b1
∂E total / ∂w1 = ∂E total / ∂out h1 * ∂out h1 / ∂net h1 * ∂net h1 / ∂w1
∂E02 / ∂out02 = (Out02 − target02)
= 0.772928465 − 0.99
= −0.217071535
∂E total / ∂out h1 = ∂E01 / ∂out h1 + ∂E02 / ∂out h1
∂E01 / ∂net01 * ∂net01 / ∂out h1 + ∂E02 / ∂net02 * ∂net02 / ∂out h1
∂E01 / ∂out01 * ∂out01 / ∂net01
w5
0.4
0.13849856
∂E02 / ∂out02 * ∂out02 / ∂net02
w7
0.50
−0.0380982
⇒ 0.055399425 + (−0.019049119)
= 0.036350306
h1
h2
O1
O2
w5
w6
w7
w8
Out02(1 − Out02)
= 0.175510052


Hidden layer
w_1
net_{h_1}
Out_{h_1}
1
$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial out_{h_1}} * \frac{\partial out_{h_1}}{\partial net_{h_1}} * \frac{\partial net_{h_1}}{\partial w_1}$
$\frac{\partial out_{h_1}}{\partial net_{h_1}} = out_{h_1}(1 - out_{h_1}) = 0.241300709$
$net_{h_1} = w_1 x_1 + w_2 x_2 + b_1 \times 1$
$\frac{\partial net_{h_1}}{\partial w_1} = x_1 = 0.05$
$\frac{\partial E_{total}}{\partial w_1} = 0.000438568$
$w_1^* = w_1 - \alpha * \frac{\partial E_{total}}{\partial w_1} = 0.15 - 0.6 * 0.000438568 = 0.149736859 2$
$\Rightarrow 0.055399425 + (-0.019049119) = 0.036350306$