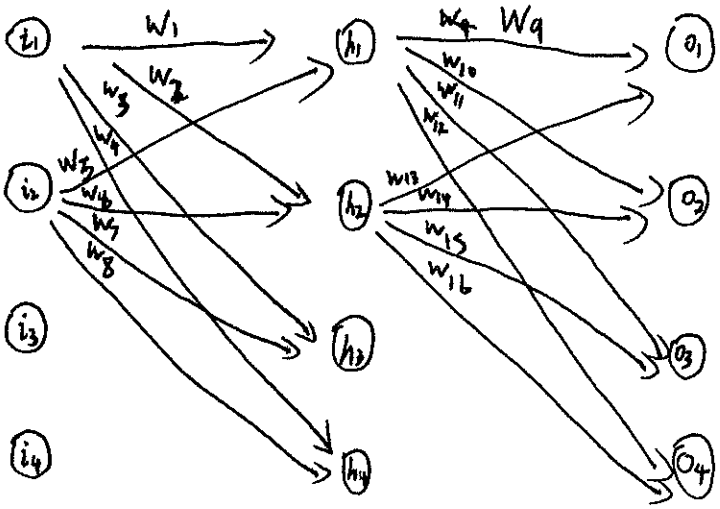
BP In Traditional Networks



INPUT

INPUT: $i_1 = 0.05$, $i_2 = 0.1$, i_3, i_4

LABEL OUTPUT: $o_1 = 0.01$, $o_2 = 0.99$, o_3, o_4

Initial weight:

$w_1 = 0.01$, $w_2 = 0.4$, $w_3 = 0.5$, $w_4 = 0.6$,

$w_5 = 0.08$, $w_6 = 0.04$, $w_7 = 0.5$, $w_8 = 0.8$, ...

Activation Function: sigmoid $= \frac{1}{1+e^{-x}}$

η: learning rate

STEP 1: INPUT → Hidden → OUTPUT (Forward Propaga...)

h_1: $w_1 \times i_1 + w_3 \times i_2 + w_{i_3 \to h_1} \times i_3 + w_{i_4 \to h_1} \times i_4 = net_{h_1}$

$out_{h_1} = \frac{1}{1+e^{-net_{h_1}}}$

h_2: $w_2 \times i_1 + w_6 \times i_2 + w_{i_3 \to h_2} \times i_3 + w_{i_4 \to h_2} \times i_4 = net_{h_2}$

~~out~~ $out_{h_2} = \frac{1}{1+e^{-net_{h_2}}}$

⋮

O_1:

$out_{h_1} = i_{h_1}$, $out_{h_2} = i_{h_2}$ ...

$i_{h_1} \times w_9 + i_{h_2} \times w_{13} + i_{h_3} \times w_{h_3 \to o_1} + i_{h_4} \times w_{h_4 \to o_1} = net_{o_1}$

$out_{o_1} = \frac{1}{1+e^{-net_{o_1}}}$

O_2:

$i_{h_1} \times w_{10} + i_{h_2} \times w_{14} + $ ~~$i_{h_3} \times w_{h_3 \to o_2}$~~ $+ i_{h_3} \times w_{h_3 \to o_2} + i_{h_4} \times w_{h_4 \to o_2} = net_{o_2}$

$out_{o_2} = \frac{1}{1+e^{-net_{o_2}}}$

⋮

STEP 1 Forward Propagation Done.

LABEL (O_1, O_2, O_3, O_4 ...)

OUTPUT (out_{o_1}, out_{o_2}, out_{o_3}, out_{o_4} ...)

STEP2 Back Propagation

① Square Error. $E_{total} = \sum (E_{o_1} + E_{o_2} \cdots) = \sum \frac{1}{2}(target - output)^2$

$E_{o_1} = \frac{1}{2}(O_1 - out_{o_1})^2 = \frac{1}{2}(0.01 - out_{o_1})^2 = e_1$

$E_{o_2} = \frac{1}{2}(O_2 - out_{o_2})^2 = \frac{1}{2}(0.99 - out_{o_2})^2 = e_2$

$E_{total} = e_1 + e_2 + \cdots = e_{total}$

② UPDATE Hid → OUTPUT weight

w_9: $\frac{\partial E_{total}}{\partial w_9} = \frac{\partial E_{total}}{\partial out_{o_1}} \times \frac{\partial out_{o_1}}{\partial net_{o_1}} \times \frac{\partial net_{o_1}}{\partial w_9}$

$= \left(-2 \times \frac{1}{2}(target_{o_1} - out_{o_1})^{\#}\right) \times \left(out_{o_1}(1 - out_{o_1})\right)$

$\times \left(out_{h_1} \times w_9^{\circ}\right) = E'_{w_9}$

$E'_{w_9} = -(target_{o_1} - out_{o_1}) \times out_{o_1}(1 - out_{o_1}) = \delta_{o_1} \times out_{h_1}$

UPDATE: $w_{9new} = w_9 - \eta * \frac{\partial E_{total}}{\partial w_9} = w_9 - \eta \delta_{o_1} \times out_{h_1}$.

⋮

③ UPDATE Hid → Input weight

$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial out_{h_1}} \times \frac{\partial out_{h_1}}{\partial net_{h_1}} \times \frac{\partial net_{h_1}}{\partial w_1}$

$= \left(\frac{\partial E_{o_1}}{\partial out_{h_1}} + \frac{\partial E_{o_2}}{\partial out_{h_1}}\right) \times \frac{\partial out_{h_1}}{\partial net_{h_1}} \times \frac{\partial net_{h_1}}{\partial w_1}$

$= (E'_{h_1} + E'_{h_2}) \times out_{h_1}(1 - out_{h_1}) \times i_1 = E'_1$

$\frac{\partial E_{o_1}}{\partial out_{h_1}} = \frac{\partial E_{o_1}}{\partial net_{o_1}} \times \frac{\partial net_{o_1}}{\partial out_{h_1}} = E_{o_1} E'_{h_1}$

UPDATE: $w_{1new} = w_1 - \eta \frac{\partial E_{total}}{\partial w_1} = w_1 - \eta E'_1$

⋮

Back Propagatim Done

KEEP Iteration