

BP In XOR-NET

Binary Neural Networks

Formula: $I \times W \approx (I \oplus B) \alpha$, B is binary,

I : Input, W : weight, α : argmin, \oplus : no mult, Conv, $B = \text{Sign}(W)$
 $W \approx \alpha B$

① FORWARD

1) Binary weight \tilde{W}

For each $\tilde{W} = \text{Sign}(W)$

① $h_1 = \tilde{W}_1 \times i_1 + \tilde{W}_2 \times i_2 \dots = \text{net}_{h_1}$

Hidden

$h_1 = f(\text{net}_{h_1})$, $f(x)$ is activation

$h_1 = \text{Batch Norm}(h_1)$

$\tilde{h}_1 = \text{Sign}(h_1)$

② $o_1 = \tilde{h}_1 \times \tilde{W}_{h_1} + \tilde{h}_2 \times \tilde{W}_{h_2} = \text{net}_{o_1}$

Output

For each, $\tilde{W}_{h_i} = \text{Sign}(W_{h_i})$

$o_{1i} = f(\text{net}_{o_1})$

② BP

①
$$\frac{\partial E_{\text{total}}}{\partial \tilde{W}_{h_1}} = \frac{\partial E_{\text{total}}}{\partial \text{net}_{h_1}} \times \frac{\partial \text{net}_{h_1}}{\partial \tilde{W}_{h_1}} \times \frac{\partial \text{net}_{h_1}}{\partial \tilde{W}_{h_1}}$$

$$= \frac{\partial E_{\text{total}}}{\partial \text{net}_{h_1}} \times \frac{\partial \text{net}_{h_1}}{\partial \tilde{W}_{h_1}} \times H \tanh(x)$$

$H \tanh(x) = \text{clip}(x, -1, 1) = \max(-1, \min(1, x))$

BUT UPDATE is to UPDATE before binary

$W_{h_1, \text{new}} = W_{h_1} - \eta \frac{\partial E_{\text{total}}}{\partial \tilde{W}_{h_1}}$, $W_{h_1, \text{new}}$ will binary in Forward Propagation