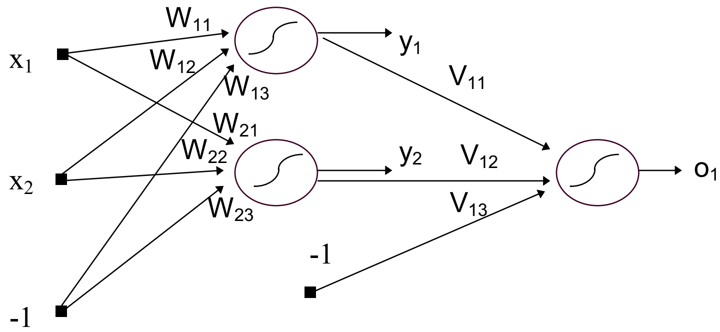
FIT5186 Lecture 4 Example

Learning the weights for the XOR problem:



Initialise weights: $\mathbf{W} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ $\mathbf{V} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$ $f(net) = \dfrac{1}{1+e^{-net}}$ $(\lambda = 1, c = 1)$

BACKPROPAGATION LEARNING RULE:

STEP 1: Input vector (0,0,–1) desired output $d_1 = 0$

STEP 2: $net_1^h = 0 \quad net_2^h = 0$

STEP 3: $y_1 = 0.5 \quad y_2 = 0.5$

STEP 4: $net_1^0 = 0$

STEP 5: $o_1 = 0.5$

STEP 6: $r_1^o = o_1(d_1 - o_1)(1 - o_1) = -0.125$

STEP 7: $r_1^h = 0 \quad r_2^h = 0$

STEP 8: $\mathbf{V}^T = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + -0.125 \begin{pmatrix} 0.5 \\ 0.5 \\ -1 \end{pmatrix} = \begin{pmatrix} -0.0625 \\ -0.0625 \\ 0.125 \end{pmatrix}$

STEP 9: $\mathbf{W_1} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \mathbf{W}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

STEP 10: error $E \leftarrow E + \sum_{k=1}^{K} (r_k^o)^2 = 0.0156$

STEP 1: Input vector (1,0,–1) desired output $d_1 = 1$

STEP 2: $net_1^h = 0 \quad net_2^h = 0$

STEP 3: $y_1 = 0.5 \quad y_2 = 0.5$

STEP 4: $net_1^0 = -0.1875$

STEP 5: $o_1 = 0.4533$

STEP 6: $r_1^o = o_1(d_1 - o_1)(1 - o_1) = 0.1355$

STEP 7: $r_1^h = -0.002117 \quad r_2^h = -0.002117$

STEP 8: $\mathbf{V}^T = \begin{pmatrix} 0.00525 \\ 0.00525 \\ -0.0105 \end{pmatrix}$

STEP 9: $\mathbf{W_1} = \begin{pmatrix} -0.002117 \\ 0 \\ 0.002117 \end{pmatrix} \qquad \mathbf{W}_2 = \begin{pmatrix} -0.002117 \\ 0 \\ 0.002117 \end{pmatrix}$

STEP 10: error $E \leftarrow E + \sum_{k=1}^{K} (r_k^o)^2 = 0.034$

STEP 1 : Input vector (0,1,−1) desired output $d_1 = 1$
STEP 2 : $net_1^h = -0.002117 \quad net_2^h = -0.002117$
STEP 3 : $y_1 = 0.4995 \quad y_2 = 0.4995$
STEP 4 : $net_1^0 = 0.0157$
STEP 5 : $o_1 = 0.5039$
STEP 6 : $r_1^o = o_1(d_1 - o_1)(1 - o_1) = 0.1240$
STEP 7 : $r_1^h = 0.00016 \quad r_2^h = 0.00016$
STEP 8 : $\mathbf{V}^T = \begin{pmatrix} 0.067188 \\ 0.067188 \\ -0.1345 \end{pmatrix}$
STEP 9 : $\mathbf{W_1} = \begin{pmatrix} -0.002117 \\ 0.00016 \\ 0.001957 \end{pmatrix} \quad \mathbf{W}_2 = \begin{pmatrix} -0.002117 \\ 0.00016 \\ 0.001957 \end{pmatrix}$
STEP 10 : error $E \leftarrow E + \sum_{k=1}^{K} (r_k^o)^2 = 0.049$

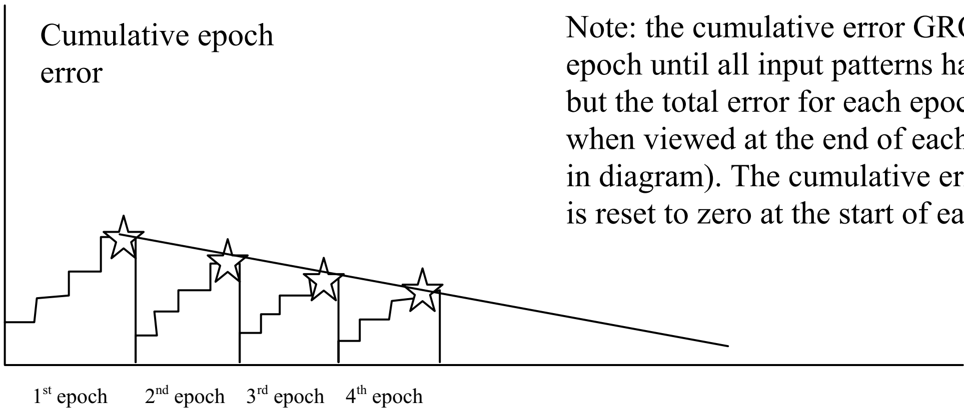
STEP 1 : Input vector (1,1,−1) desired output $d_1 = 0$
STEP 2 : $net_1^h = -0.003914 \quad net_2^h = -0.003914$
STEP 3 : $y_1 = 0.4990 \quad y_2 = 0.4990$
STEP 4 : $net_1^0 = 0.20155$
STEP 5 : $o_1 = 0.5502$
STEP 6 : $r_1^o = o_1(d_1 - o_1)(1 - o_1) = -0.1362$
STEP 7 : $r_1^h = -0.00229 \quad r_2^h = -0.00229$
STEP 8 : $\mathbf{V}^T = \begin{pmatrix} -0.00078 \\ -0.00078 \\ 0.0017 \end{pmatrix}$
STEP 9 : $\mathbf{W_1} = \begin{pmatrix} -0.0044 \\ -0.00213 \\ 0.00427 \end{pmatrix} \quad \mathbf{W}_2 = \begin{pmatrix} -0.0044 \\ -0.00213 \\ 0.00427 \end{pmatrix}$
STEP 10 : error $E \leftarrow E + \sum_{k=1}^{K} (r_k^o)^2 = 0.0676$

At end of 1st epoch,

if E < 0.00001 (or some other suitably small tolerance) then STOP training, otherwise reset E=0 and repeat for another epoch.

Since E = 0.0676 (> tolerance level), continue repeating epochs until error is below tolerance level.

•



Note: the cumulative error GROWS during each epoch until all input patterns have been presented, but the total error for each epoch DECREASES when viewed at the end of each epoch (see stars in diagram). The cumulative error for each epoch is reset to zero at the start of each new epoch.