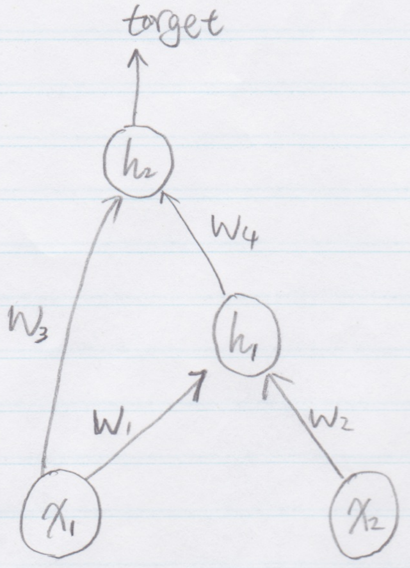
CSC321 T5

true label : t

A2 Part 3 :



$$h_2 = \frac{1}{1+\exp(-z_2)} = \text{target} \equiv \hat{t}$$

$$z_2 = w_3 x_1 + w_4 h_1$$

$$h_1 = \frac{1}{1+\exp(z_1)}$$

$$z_1 = w_1 x_1 + w_2 x_2$$

$W = [w_1 \ \ w_2 \ \ w_3 \ \ w_4]$, $w_k \in [-1 : 0.25 : 1]$, 9 values

$$D = \begin{bmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \\ & \vdots \\ x_{N,1} & x_{N,2} \end{bmatrix}$$

Search space of W : $9^4 = 6561$

$$\text{cost} = -\sum_{i=1}^{N}\left[t_i \log \hat{t}_i + (1-t_i)\log(1-\hat{t}_i)\right]$$

Likelihood: $P(D|W) = \exp(-\text{cost})$

Bayes rule:

$$P(W|D) = \frac{P(W)\,P(D|W)}{P(D)}$$

where $P(D) = \sum_{W} P(W)\,P(D|W)$

Prediction on test data x_j :

Bayesian estimate :

$$\hat{t}_j^{(Bayes)} \equiv p_{Bayes}(t_j | x_j) = \sum_W p(W|D) \overbrace{p(t_j | x_j, W)}^{\hat{t}_j}$$

Maximum a-posteriori (MAP) :

$$W_{MAP} = \underset{W}{\mathrm{argmax}}\ p(W|D)$$

$$= \underset{W}{\mathrm{argmax}}\ \frac{p(W)\,p(D|W)}{p(D)}$$

$$= \underset{W}{\mathrm{argmax}}\ p(W)\,p(D|W)$$

$$\hat{t}_j^{(MAP)} \equiv p_{MAP}(t_j | x_j) = p(t_j | x_j, W_{MAP})$$

Maximum likelihood (ML) :

$$W_{ML} = \underset{W}{\mathrm{argmax}}\ p(D|W)$$

$$\hat{t}_j^{(ML)} \equiv p_{ML}(t_j | x_j) = p(t_j | x_j, W_{ML})$$

If $p(W) \sim \mathrm{uniform}(W)$, then:

$$W_{MAP} = W_{ML} = \underset{W}{\mathrm{argmax}}\ p(W)\,p(D|W),$$

so

$$\hat{t}^{(MAP)} = \hat{t}^{(ML)}$$

Evaluation function (Contrastive divergence) :

$$\text{ErrorSum} = -\sum_{i=1}^{N} \left[t_i \log(\hat{t}_i) + (1-t_i) \log(\hat{t}_i) \right] + \sum_{i=1}^{N} \left[t_i \log(t_i) + (1-t_i) \log(t_i) \right]$$

$$\text{ErrorPerBit} = \frac{\text{ErrorSum}}{N \log(2)}$$