

5.4 This exercise seems similar to Exercise 4.16.

Using cross entropy, we get:

$$\begin{aligned} -\ln p(\mathbf{t}|\mathbf{X}, \mathbf{w}) &= -\sum_{n=1}^N \{ (t_n(1-\epsilon) + (1-t_n)\epsilon) \ln y(\mathbf{x}_n, \mathbf{w}) \\ &\quad + ((1-t_n)(1-\epsilon) + t_n\epsilon) \ln(1-y(\mathbf{x}_n, \mathbf{w})) \} \end{aligned}$$

Note : The official solution manual is incorrect here. Let's assume that $\epsilon = 1/2$. In that case, according to the official solution, the cross entropy becomes:

$$\begin{aligned} E(\mathbf{w}) &= -\sum_{n=1}^N \{ t_n \ln[(1-\epsilon)y(\mathbf{x}_n, \mathbf{w}) + \epsilon(1-y(\mathbf{x}_n, \mathbf{w}))] \\ &\quad + (1-t_n) \ln[1 - (1-\epsilon)y(\mathbf{x}_n, \mathbf{w}) - \epsilon(1-y(\mathbf{x}_n, \mathbf{w}))] \} \\ &= -\sum_{n=1}^N \{ t_n \ln[(1/2)y(\mathbf{x}_n, \mathbf{w}) + (1/2)(1-y(\mathbf{x}_n, \mathbf{w}))] \\ &\quad + (1-t_n) \ln[1 - (1/2)y(\mathbf{x}_n, \mathbf{w}) - (1/2)(1-y(\mathbf{x}_n, \mathbf{w}))] \} \\ &= -\sum_{n=1}^N \{ t_n \ln[(1/2)y(\mathbf{x}_n, \mathbf{w}) + (1/2) - (1/2)y(\mathbf{x}_n, \mathbf{w})] \\ &\quad + (1-t_n) \ln[1 - (1/2)y(\mathbf{x}_n, \mathbf{w}) - (1/2) + (1/2)y(\mathbf{x}_n, \mathbf{w})] \} \\ &= -\sum_{n=1}^N \{ t_n \ln[(1/2)] + (1-t_n) \ln[1/2] \} \\ &= -\sum_{n=1}^N \{ t_n \ln[(1/2)] + \ln[1/2] - t_n \ln[1/2] \} \\ &= -\sum_{n=1}^N \{ \ln[1/2] \} \\ &= N \ln[2] \end{aligned}$$

This would imply that for $\epsilon = 1/2$, the cross-entropy does not depend on the function $y(\mathbf{x}_n, \mathbf{w})$ at all. This is not correct.