3.4 Given the error function :

$$E_D(\mathbf{w}) = \frac{1}{2}\sum_{n=1}^{N}\{y(\mathbf{x}_n, \mathbf{w}) - t_n\}^2$$

$$= \frac{1}{2}\sum_{n=1}^{N}\left\{\left(w_0 + \sum_{i=1}^{D} w_i(x_{ni} + \epsilon_i)\right) - t_n\right\}^2$$

Expanding the square, we get:

$$= \frac{1}{2}\sum_{n=1}^{N}\left\{\left(w_0 + \sum_{i=1}^{D} w_i(x_{ni} + \epsilon_i)\right)^2 + t_n^2 - 2\left(w_0 + \sum_{i=1}^{D} w_i(x_{ni} + \epsilon_i)\right) t_n\right\}$$

If we take expectation w.r.t ϵ, we get:

$$\mathbb{E}_\epsilon[E_D] = \frac{1}{2}\sum_{n=1}^{N}\left\{\mathbb{E}_\epsilon\left[\left(w_0 + \sum_{i=1}^{D} w_i(x_{ni} + \epsilon_i)\right)^2\right] + t_n^2 - 2\left(w_0 + \sum_{i=1}^{D} w_i x_{ni}\right) t_n\right\}$$

$$= \frac{1}{2}\sum_{n=1}^{N}\left\{\mathbb{E}_\epsilon\left[\left(w_0 + \sum_{i=1}^{D} w_i x_{ni} + \sum_{i=1}^{D} w_i \epsilon_i\right)^2\right] + t_n^2 - 2\left(w_0 + \sum_{i=1}^{D} w_i x_{ni}\right) t_n\right\}$$

$$\left(w_0 + \sum_{i=1}^{D} w_i x_{ni} + \sum_{i=1}^{D} w_i \epsilon_i\right)^2$$

$$= \left(w_0 + \sum_{i=1}^{D} w_i x_{ni}\right)^2 + \left(\sum_{i=1}^{D} w_i \epsilon_i\right)^2 + 2\left(w_0 + \sum_{i=1}^{D} w_i x_{ni}\right)\left(\sum_{i=1}^{D} w_i \epsilon_i\right)$$

Taking expectation w.r.t ϵ, the third term vanishes, giving us:

$$= \left(w_0 + \sum_{i=1}^{D} w_i x_{ni}\right)^2 + \mathbb{E}_\epsilon\left[\left(\sum_{i=1}^{D} w_i \epsilon_i\right)^2\right]$$

$$= \left(w_0 + \sum_{i=1}^{D} w_i x_{ni} \right)^2 + \mathbb{E}_\epsilon \left[\left(\sum_{i,j} w_i w_j \epsilon_i \epsilon_j \right) \right]$$

$$= \left(w_0 + \sum_{i=1}^{D} w_i x_{ni} \right)^2 + \left(\sum_{i,j} w_i w_j \mathbb{E}_\epsilon[\epsilon_i \epsilon_j] \right)$$

$$= \left(w_0 + \sum_{i=1}^{D} w_i x_{ni} \right)^2 + \left(\sum_{i,j} w_i w_j \delta_{ij} \sigma^2 \right)$$

Substituting this result in the expression for $\mathbb{E}_\epsilon[E_D]$, we get:

$$\mathbb{E}_\epsilon[E_D] = \frac{1}{2} \sum_{n=1}^{N} \left\{ \left(w_0 + \sum_{i=1}^{D} w_i x_{ni} \right)^2 + \left(\sum_{i,j} w_i w_j \delta_{ij} \sigma^2 \right) + t_n^2 - 2 \left(w_0 + \sum_{i=1}^{D} w_i x_{ni} \right) t_n \right\}$$

$$= \frac{1}{2} \sum_{n=1}^{N} \left\{ \left(w_0 + \sum_{i=1}^{D} w_i x_{ni} \right)^2 + t_n^2 - 2 \left(w_0 + \sum_{i=1}^{D} w_i x_{ni} \right) t_n \right\} + \frac{N\sigma^2}{2} \left(\sum_{i,j} w_i w_j \delta_{ij} \right)$$

$$= \frac{1}{2} \sum_{n=1}^{N} \left\{ \left(w_0 + \left(\sum_{i=1}^{D} w_i x_{ni} \right) - t_n \right)^2 \right\} + \frac{N\sigma^2}{2} \left(\sum_{i,j} w_i w_j \delta_{ij} \right)$$

The first term minimizes the sum-of-squares error for noise-free input variables, and the second term is the weight-decay regularization term, in which the bias parameter w_0 is omitted from the regularizer.