Assignment

Q: Given overall network function

$$y_k(\boldsymbol{x}, \boldsymbol{w}) = f(\sum_{j=0}^{M} w_{kj}^{(2)} h(\sum_{i=0}^{D} w_{ji}^{(1)} x_i))$$

and cross-entropy error function for the multiclass classification problem

$$E(\boldsymbol{w}) = -\ln\left(\prod_{k=1}^{K} y_k^{t_k}\right) = -\sum_{k=1}^{K} t_k \ln y_k$$

where $t_k = \{0, 1\}^K$ and $\sum_{k=1}^{K} t_k = 1$, $f(z_k) = \frac{\exp(z_k)}{\sum_{l=1}^{K} \exp(z_l)}$ is the softmax function, $h(a) = \tanh(a) = \frac{\exp(a) - \exp(-a)}{\exp(a) + \exp(-a)}$. Compute the derivatives of error function w.r.t the first-layer and second-layer weights with Back Propagation

$$\frac{\partial E}{\partial w_{ji}^{(1)}} \text{ and } \frac{\partial E}{\partial w_{kj}^{(2)}}$$