SYDE552/BIOL487 Assignment: Convolutional Networks

25 marks total
Due Feb 12

1. Show that

$$\frac{\partial E}{\partial a_k} = y_k - t_k$$

for the binary classification output function,

$$y_k = \frac{1}{1+\exp(-a_k)}$$

with error function,

$$E = -\sum_{k=1}^{K} t_k \ln y_k + (1-t_k)\ln(1-y_k)$$

2. One of the earliest convolutional networks is from LeCun et al. (1990) Handwritten Digit Recognition with a Back-Propagation Network. In *NIPS*89* (pp. 396–404). This is also one of the smallest convolutional networks. Consider a modified version of this network that is identical except that the output units are replaced with the multiclass classification response and error function,

$$y_k = \frac{\exp(a_k)}{\sum_l \exp(a_l)}$$

$$E = -\sum_{k=1}^{K} t_k \ln y_k$$

As we discussed in lectures,

$$\frac{\partial E}{\partial a_k} = y_k - t_k$$

for the **output units**. Calculate the partial derivatives of the error with respect to the activations of the **other** units in the **last convolutional and subsampling layers**, and also for each of the **weights** in the inputs to these layers. Show your work.

3. Briefly describe the method LeCun et al. used to train this network.