Machine Intelligence I
Exercise sheet 3
2 Seriously cool guys

H3.1: Binary Classification

a)

$$e^{(\alpha)} = -\left[y_T^{(\alpha)} \ln y(\underline{x}^{(\alpha)}; \underline{w}) + (1-y_T^{(\alpha)}) \ln(1-y(\underline{x}^{(\alpha)}; \underline{w}))\right]$$

used: $\frac{d}{dx}(f(x)g(x)) = \frac{df(x)}{dx}g(x) + \frac{dg(x)}{dx}f(x)$

$$\frac{d}{dx}(\ln(f(x)) = \frac{f'(x)}{f(x)}$$

$$\frac{\partial e^{(\alpha)}}{\partial y(\underline{x}^{(\alpha)}; \underline{w})} = -\left[\frac{y_T^{(\alpha)}}{y(\underline{x}^{(\alpha)}; \underline{w})} + \frac{(1-y_T^{(\alpha)})(-1)}{1-y(\underline{x}^{(\alpha)}; \underline{w})}\right]$$

$$= \frac{y(\underline{x}^{(\alpha)}; \underline{w}) - y_T^{(\alpha)} y(\underline{x}^{(\alpha)}; \underline{w}) - y_T^{(\alpha)} + y_T^{(\alpha)} y(\underline{x}^{(\alpha)}; \underline{w})}{y(\underline{x}^{(\alpha)}; \underline{w}) \cdot (1-y(\underline{x}^{(\alpha)}; \underline{w}))}$$

$$= \frac{y(\underline{x}^{(\alpha)}; \underline{w}) - y_T^{(\alpha)}}{y(\underline{x}^{(\alpha)}; \underline{w})(1-y(\underline{x}^{(\alpha)}; \underline{w})} \checkmark$$

b) $f(h_1^2) = \frac{1}{1+e^{-h_1^2}}$ used: $\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - g'(x)f(x)}{(g(x))^2}$

$$f'(h_1^2) = \frac{-(-e^{-h_1^2})}{(1+e^{-h_1^2})^2} = \frac{1}{(1+e^{-h_1^2})^2} e^{-h_1^2}$$

$\rightarrow f(h_1^2)^2$

$$= f(h_1^2)^2 \cdot e^{-h_1^2} = f(h_1^2)^2 (e^{-h_1^2} + 1 - 1) = f(h_1^2)^2 \left(\frac{1}{f(h_1^2)} - 1\right)$$

$e^{-h_1^2} + 1 \rightarrow \frac{1}{f(h_1^2)}$

$$\Rightarrow f'(h_1^2) = f(h_1^2)(1-f(h_1^2)) \checkmark$$

c)

$$\frac{\partial y}{\partial w_{ij}^{v,v-1}} = \frac{\partial y}{\partial h_i^v} \frac{\partial h_i^v}{\partial w_{ij}^{v,v-1}} = \frac{\partial y}{\partial h_i^v} S_j^{v-1}$$

$$\frac{\partial e^{(\alpha)}}{\partial w_{1j}^{21}} = \frac{\partial e^{(\alpha)}}{\partial y(\underline{x}^{(\alpha)}; \underline{w})} \cdot \frac{\partial y(\underline{x}^{(\alpha)}; \underline{w})}{\partial w_{1j}^{21}}$$

$$= \frac{\partial e^{(\alpha)}}{\partial y(\underline{x}^{(\alpha)}; \underline{w})} \cdot \frac{\partial y(\underline{x}^{(\alpha)}; \underline{w})}{\partial h_1^2} \cdot S_j^1$$

(a) (b)

$$= \frac{y(\underline{x}^{(\alpha)}; \underline{w}) - y_T^{(\alpha)}}{y(\underline{x}^{(\alpha)}; \underline{w})(1 - y(\underline{x}^{(\alpha)}; \underline{w})} \cdot \frac{y(\underline{x}^{(\alpha)}; \underline{w})(1 - y(\underline{x}^{(\alpha)}; \underline{w})}{1} S_j^1$$

$$= (y(\underline{x}^{(\alpha)}; \underline{w}) - y_T^{(\alpha)}) S_j^1 \checkmark$$