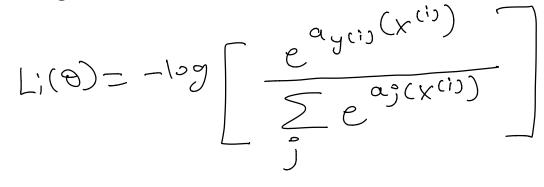
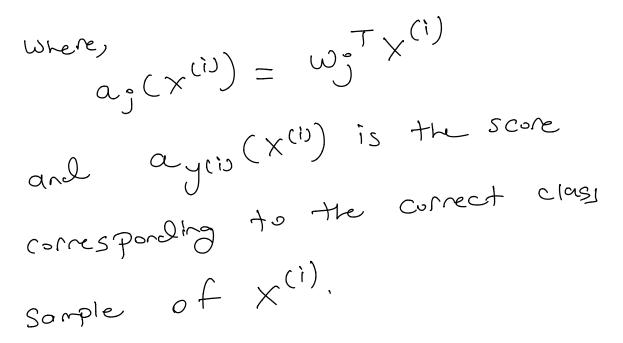
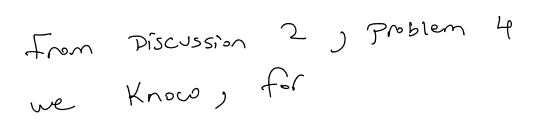
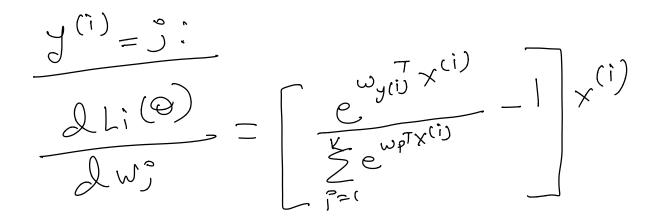
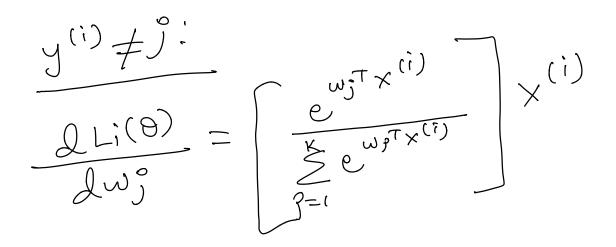
Suppose the tuple (x(i), y(i)) consists of the *î*th training example and we define the loss for the *î*th training example as follows:











) of example; Suppose we have 3 classes.  $C = \frac{2}{2} [, 2, 3]$ and suppose we have a training sample  $(\chi^{(i)}, 2)$ . Then  $\frac{QLi}{QW2} = \begin{bmatrix} \omega_2 T \chi^{(i)} \\ \frac{e}{3e^{\omega_p T \chi^{(i)}}} \\ \frac{e}{3e^{\omega_p T \chi^{(i)}}} \end{bmatrix}$  $\frac{d Li}{d w_{1}} = \left( \frac{e}{\frac{2}{3}e^{w_{1}T \times (i)}} \right) \times (i)$   $\frac{e}{\frac{3}{3}e^{w_{2}T \times (i)}} = \left( \frac{e}{\frac{3}{3}e^{w_{2}T \times (i)}} \right)$ 

 $\frac{d Li}{d w_3} = \left( \begin{array}{c} \omega_3^T \chi^{(i)} \\ \frac{e}{3} e^{\omega_3^T \chi^{(i)}} \\ \frac{3}{2} e^{\omega_3^T \chi^{(i)}} \\ \frac{3}{2} e^{\omega_3^T \chi^{(i)}} \\ \frac{3}{2} e^{\omega_3^T \chi^{(i)}} \end{array} \right)$