SAPIENZA
UNIVERSITÀ DI ROMA

Advanced Machine Learning - Assignment #1

Group Members
Stefano D'Arrigo 1960500
Paola Antonicoli 1796554
Simone Fiorellino 1960415
Jeremy Sapienza 1960498

Oct 27 2021

1 Question 2

1.1 a)

$$note\ that\ (\Delta_i)_j = \begin{cases} 1, & if j = y_i \\ 0, & otherwise \end{cases} \tag{1}$$

$$\frac{\partial J}{\partial z_i^{(3)}}\left(\theta,\ \{x_i, y_i\}_{i=1}^N\right) = \frac{\partial}{\partial z_i^{(3)}}\ \frac{1}{N}\sum_{i=1}^{N} -\log\left(\frac{\exp\left(z_i^{(3)}\right)_{y_i}}{\sum_{j=1}^{K}\exp\left(z_i^{(3)}\right)_j}\right) =$$

$$\frac{\partial}{\partial z_i^{(3)}}\left[\frac{1}{N}\sum_{i=1}^{N}\log\left(\sum_{j=1}^{K}\exp\left(z_i^{(3)}\right)_j\right) - \log\left(\exp\left(z_i^{(3)}\right)_{y_i}\right)\right] =$$

$$= \frac{1}{N}\sum_{i=1}^{N}\ \left\{\frac{\partial}{\partial z_i^{(3)}}\left[\log\left(\sum_{j=1}^{K}\exp\left(z_i^{(3)}\right)_j\right)\right] - \frac{\partial}{\partial z_i^{(3)}}\left[\log\left(\exp\left(z_i^{(3)}\right)_{y_i}\right)\right]\right\} =$$

$$= \frac{1}{N}\sum_{i=1}^{N}\ \left\{\frac{\partial}{\partial z_i^{(3)}}\left[\log\left(\sum_{j=1}^{K}\exp\left(z_i^{(3)}\right)_j\right)\right] - \frac{\partial}{\partial z_i^{(3)}}\left(z_i^{(3)}\right)_{y_i}\right\} =$$

$$= \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\exp\left(z_i^{(3)}\right)}{\sum_{j=1}^{K}\exp\left(z_i^{(3)}\right)_j} - \Delta_i\right) = \frac{1}{N}\sum_{i=1}^{N}\left(\psi\left(z_i^{(3)}\right) - \Delta_i\right)$$