Homework 4

Pana Wanitchollakit 6532136721

T1

$$
\begin{aligned}
x_1 &= \max(1.0 * 0.3 + 0.1, 0) = 0.4 \\
y_1 &= 0.4 * (-0.2) - 0.3 = -0.38 \\
z &= \max(-0.38 + 1.0, 0) = 0.62
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial z}{\partial y_1} &= 1 \\
\frac{\partial z}{\partial x_1} &= \frac{\partial z}{\partial y_1}\frac{\partial y_1}{\partial x_1} = 1 * w_1
\end{aligned}
$$

$\therefore$

$$
\begin{aligned}
\frac{\partial z}{\partial w_0} &= \frac{\partial z}{\partial x_1}\frac{\partial x_1}{\partial w_0} = w_1 * x_0 = -0.2 \\
\frac{\partial z}{\partial w_1} &= \frac{\partial z}{\partial y_1}\frac{\partial y_1}{\partial w_1} = 1 * x_1 = 0.4 \\
\frac{\partial z}{\partial b_0} &= \frac{\partial z}{\partial x_1}\frac{\partial x_1}{\partial b_0} = w_1 * 1 = -0.2 \\
\frac{\partial z}{\partial b_1} &= \frac{\partial z}{\partial y_1}\frac{\partial y_1}{\partial b_1} = 1 * 1 = 1
\end{aligned}
$$

T2

$$
\begin{aligned}
\textbf{Input}: &\ (32, 30) \\
\textbf{A}: &\ (32, 1024) \\
\textbf{B}: &\ (32, 512) \\
\textbf{C}: &\ (32, 1)
\end{aligned}
$$

T3

$$1024(30 + 1) + 512(1024 + 1) + 1(512 + 1) = 557,057 \text{ parameters}$$

T4

Proof.

$$L = -\sum_j y_j \log P(y = j)$$
$$L = -\sum_j y_j (h_j - \log \sum_k \exp(h_k))$$

$\therefore$

$$\frac{\partial L}{\partial h_i} = -\sum_j y_j \left(\frac{\partial}{\partial h_i} h_j - \frac{\frac{\partial}{\partial h_i} \sum_k \exp(h_k)}{\sum_k \exp(h_k)} \right)$$
$$\frac{\partial L}{\partial h_i} = -\sum_j y_j \left(\frac{\partial}{\partial h_i} h_j - \frac{\exp(h_i)}{\sum_k \exp(h_k)} \right)$$
$$\frac{\partial L}{\partial h_i} = -y_i + \sum_j y_j \frac{\exp(h_i)}{\sum_k \exp(h_k)}$$
$$\frac{\partial L}{\partial h_i} = -y_i + \frac{\exp(h_i)}{\sum_k \exp(h_k)} \qquad \because \sum_j y_j = 1$$
$$\frac{\partial L}{\partial h_i} = P(y = i) - y_i$$

□