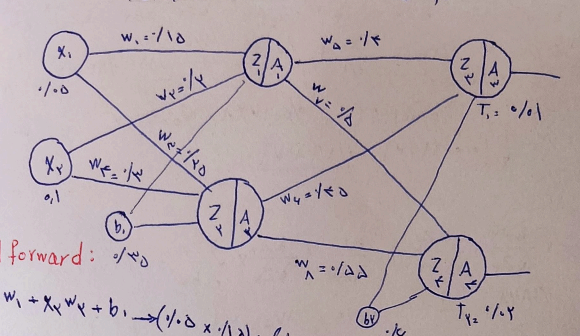
بسمه تعالی

سؤال اول :　　علی سلیم خانی　　۴۰۱۴۶۳۱۴۱



1. feed forward: ۰/۳۵

$Z_1=x_1w_1+x_2w_2+b_1 \rightarrow (۰/۰۵\times۰/۱۵)+(۰/۱\times۰/۲)+۰/۳۵=۰/۳۷۷۵$　$A_1=\text{sigmoid}(z_1)=\frac{1}{1+e^{-z_1}}$

$Z_2=x_2w_4+x_1w_3+b_1 \rightarrow (۰/۰۵*۰/۲۵)+(۰/۱\times۰/۳)+۰/۳۵=۰/۳۹۲۵$　$A_1=۰/۵۹۳۲$

$Z_3=A_1w_5+A_2w_6+b_2 \rightarrow (۰/۵۹۳۲\times۰/۴)+(۰/۵۹۶۸\times۰/۴۵)+۰/۶=۱/۱۰۵۸۴$　$A_2=\text{sigmoid}(z_2)=۰/۵۹۶۸$

$A_3=\text{sigmoid}(z_3)=۰/۷۵۱۳$

$Z_4=A_1w_7+A_2w_8+b_2 \rightarrow ۱/۲۲۴۸۴$　$A_4=\text{sigmoid}(z_4)=۰/۷۷۲۹$

$\text{Loss}=\frac{1}{2}(T_1-A_3)^2+\frac{1}{2}(T_2-A_4)^2=۰/۲۷۴۷۹۲۸۴۵+۰/۲۸۳۴۲۹۲۰۵=۰/۵۵۸۱۹۲۰۵$

2. backPropagation:

$\star\ \frac{\partial L}{\partial w_5}=\frac{\partial L}{\partial A_3}\times\frac{\partial A_3}{\partial z_3}\times\frac{\partial z_3}{\partial w_5}$

$\frac{\partial L}{\partial A_3} \Rightarrow L=\frac{1}{2}(A_3-T_1)^2=2\times\frac{1}{2}(A_3-T_1)\times1=(A_3-T_1)=۰/۷۴۱۳$

$\frac{\partial A_3}{\partial z_3} \Rightarrow A_3=\text{sigmoid}(z_3)=A_3(1-A_3)=۰/۱۸۶۸$　$\frac{\partial L}{\partial w_5}=(۰/۷۴۱۳)\times(۰/۱۸۶۸)+(۰/۵۹۳۲)=۰/۰۸۲۱$

$\frac{\partial z_3}{\partial w_5} \Rightarrow z_3=A_1w_5+A_2w_6+b_2=A_1=۰/۵۹۳۲$

$w_5'=w_5-\alpha\frac{\partial L}{\partial w_5}$

$w_5'=۰/۴-\frac{1}{2}(۰/۰۸۲۱)=\boxed{\cancel{۳۹۵۸۹۵}}$ ~~۳/۹۵۸۹۵~~ $\boxed{۰/۳۵۸۹۵}$

$\star\ \frac{\partial L}{\partial w_6}=\frac{\partial L}{\partial A_3}\times\frac{\partial A_3}{\partial z_3}\times\frac{\partial z_3}{\partial w_6}=۰/۷۴۱۳\times۰/۱۸۶۸\times A_2=۰/۰۸۲۶$

$z_3=A_1w_5+A_2w_6+b_2\rightarrow A_2$

$w_6'=w_6-\alpha\frac{\partial L}{\partial w_6}=۰/۴۵-\frac{1}{2}(۰/۰۸۲۶)$　$w_6'=۰/۴۰۸۷$

$\star\ \frac{\partial L}{\partial w_7}=\frac{\partial L}{\partial A_4}\times\frac{\partial A_4}{\partial z_4}\times\frac{\partial z_4}{\partial w_7}=۰/۰۷۸۴$

$L=\frac{1}{2}(A_4-T_2)^2=\frac{1}{2}\times2\times(A_4-T_2)\times1=A_4-T_2=۰/۷۵۲۹$

$w_7'=w_7-\alpha\frac{\partial L}{\partial w_7}=۰/۵-\frac{1}{2}(۰/۰۷۸۴)=\boxed{۰/۴۶۰۸۵}$

$A_4=\text{sigmoid}(z_4)=A_4(1-A_4)=۰/۱۷۵۵$

$z_4=A_1w_7+A_2w_8+b_2=A_1=۰/۵۹۳۲$

$\star\ \frac{\partial L}{\partial w_8}=\frac{\partial L}{\partial A_4}\times\frac{\partial A_4}{\partial z_4}\times\frac{\partial z_4}{\partial w_8}=۰/۷۵۲۹\times۰/۱۷۵۵\times A_2=۰/۰۷۸۸$　$w_8'=w_8-\alpha\frac{\partial L}{\partial w_8}=۰/۵۵-\frac{1}{2}(۰/۰۷۸۸)=۰/۵۱۰۶$

$\star\ \frac{\partial L}{\partial w_1}=\frac{\partial L}{\partial A_3}\times\frac{\partial A_3}{\partial z_3}\times\frac{\partial z_3}{\partial A_1}\times\frac{\partial A_1}{\partial z_1}\times\frac{\partial z_1}{\partial w_1}+\frac{\partial L}{\partial A_4}\times\frac{\partial A_4}{\partial z_4}\times\frac{\partial z_4}{\partial A_1}\times\frac{\partial A_1}{\partial z_1}\times\frac{\partial z_1}{\partial w_1}$

(A_3-T_1) ۰/۷۴۱۳　$A_3(1-A_3)$ ۰/۱۸۶۸　w_5 ۰/۴　$A_1(1-A_1)$ ۰/۲۴۱۳　x_1 ۰/۰۵　(A_4-T_2) ۰/۷۵۲۹　$A_4(1-A_4)$ ۰/۱۷۵۵　w_7 ۰/۵　$A_1(1-A_1)$ ۰/۲۴۱۳　x_1 ۰/۰۵

$z_3=A_1w_5+A_2w_6$　$z_1=x_1w_1+x_2w_2$　$z_4=A_1w_7+A_2w_8$　$z_1=x_1w_1+x_2w_2$

$۰/۰۰۰۶۶۸۲+۰/۰۰۰۷۹۷۰=۰/۰۰۱۴۶۵۲$

$w_1'=w_1-\alpha\frac{\partial L}{\partial w_1}=۰/۱۵-\frac{1}{2}(۰/۰۰۱۴۶۵۲)=۰/۱۴۹۲۶$

$\star\ \frac{\partial L}{\partial w_2}=\frac{\partial L}{\partial A_3}\times\frac{\partial A_3}{\partial z_3}\times\frac{\partial z_3}{\partial A_1}\times\frac{\partial A_1}{\partial z_1}\times\frac{\partial z_1}{\partial w_2}+\frac{\partial L}{\partial A_4}\times\frac{\partial A_4}{\partial z_4}\times\frac{\partial z_4}{\partial A_1}\times\frac{\partial A_1}{\partial z_1}\times\frac{\partial z_1}{\partial w_2}$

۰/۷۴۱۳　۰/۱۸۶۸　w_5 ۰/۴　۰/۲۴۱۳　x_2 ۰/۱　۰/۷۵۲۹　۰/۱۷۵۵　w_7 ۰/۵　۰/۲۴۱۳　x_2 ۰/۱

$۰/۰۰۱۳۳۶+۰/۰۰۱۵۹۴=۰/۰۰۲۹۳$　$w_2'=۰/۲-\frac{1}{2}(۰/۰۰۲۹۳)=۰/۱۹۸۵$

$\star\ \frac{\partial L}{\partial w_3}=\frac{\partial L}{\partial A_3}\times\frac{\partial A_3}{\partial z_3}\times\frac{\partial z_3}{\partial A_2}\times\frac{\partial A_2}{\partial z_2}\times\frac{\partial z_2}{\partial w_3}+\frac{\partial L}{\partial A_4}\times\frac{\partial A_4}{\partial z_4}\times\frac{\partial z_4}{\partial A_2}\times\frac{\partial A_2}{\partial z_2}\times\frac{\partial z_2}{\partial w_3}$

۰/۷۴۱۳　۰/۱۸۶۸　w_6 ۰/۴۵　x_1 ۰/۰۵　+　۰/۷۵۲۹　۰/۱۷۵۵　w_8 ۰/۵۵　$A_2(1-A_2)$ ۰/۲۴۰۶۴　x_1 ۰/۰۵

$z_2=x_1w_3+x_2w_4$　$A_2(1-A_2)$ ۰/۲۴۰۶۴

$z_4=A_1w_7+A_2w_8$

$w_3'=w_3-\frac{1}{2}(۰/۰۰۱۶۲۳۹)=۰/۲۴۹۱۸۸$

$۰/۰۰۰۷۴۹۶+۰/۰۰۰۸۷۴۳=۰/۰۰۱۶۲۳۹$

$\star\ \frac{\partial L}{\partial w_4}=\frac{\partial L}{\partial A_3}\times\frac{\partial A_3}{\partial z_3}\times\frac{\partial z_3}{\partial A_2}\times\frac{\partial A_2}{\partial z_2}\times\frac{\partial z_2}{\partial w_4}+\frac{\partial L}{\partial A_4}\times\frac{\partial A_4}{\partial z_4}\times\frac{\partial z_4}{\partial A_2}\times\frac{\partial A_2}{\partial z_2}\times\frac{\partial z_2}{\partial w_4}$

۰/۷۴۱۳　۰/۱۸۶۸　w_6 ۰/۴۵　۰/۲۴۰۶۴　x_2 ۰/۱　+　۰/۷۵۲۹　۰/۱۷۵۵　w_8 ۰/۵۵　۰/۲۴۰۶۴　x_2 ۰/۱

$۰/۰۰۱۴۹۹+۰/۰۰۱۷۴۸=۰/۰۰۳۲۴۷$

$w_4'=w_4-\alpha\frac{\partial L}{\partial w_4}=۰/۳-\frac{1}{2}(۰/۰۰۳۲۴۷)=۰/۲۹۸۳۷$

$\frac{\partial L}{\partial b_2}=\frac{\partial L}{\partial A_3}\times\frac{\partial A_3}{\partial z_3}\times\frac{\partial z_3}{\partial b_2}+\frac{\partial L}{\partial A_4}\times\frac{\partial A_4}{\partial z_4}\times\frac{\partial z_4}{\partial b_2}=۰/۲۷۰۶۰۸$

۰/۷۴۱۳　۰/۱۸۶۸　1　۰/۷۵۲۹　۰/۱۷۵۵ × 1

$b_2'=b_2-\alpha\frac{\partial L}{\partial b_2}$

$z_3=A_1w_5+A_2w_6+b_2\rightarrow1$

$z_4=A_1w_7+A_2w_8+b_2\rightarrow1$

$۰/۶-\frac{1}{2}(۰/۲۷۰۶۰۸)=۰/۴۶۴۶۹۶$

$\frac{\partial L}{\partial b_1}=\frac{\partial L}{\partial A_3}\times\frac{\partial A_3}{\partial z_3}\times\frac{\partial z_3}{\partial A_1}\times\frac{\partial A_1}{\partial z_1}\times\frac{\partial z_1}{\partial b_1}+\frac{\partial L}{\partial A_4}\times\frac{\partial A_4}{\partial z_4}\times\frac{\partial z_4}{\partial A_2}\times\frac{\partial A_2}{\partial z_2}\times\frac{\partial z_2}{\partial b_1}$

$+\frac{\partial L}{\partial A_3}\times\frac{\partial A_3}{\partial z_3}\times\frac{\partial z_3}{\partial A_2}\times\frac{\partial A_2}{\partial z_2}\times\frac{\partial z_2}{\partial b_1}+\frac{\partial L}{\partial A_4}\times\frac{\partial A_4}{\partial z_4}\times\frac{\partial z_4}{\partial A_1}\times\frac{\partial A_1}{\partial z_1}\times\frac{\partial z_1}{\partial b_1}$

$=۲/۱۲۶۴۲۷۵۵۲۶۲۴۳$　$w'=۰/۳۵-(\frac{1}{2}\times۲/۱۲۶...)=-۰/۷۱۳۲۱۳$

سؤال دوم : اگر فرض کنیم که ~~a_i sigmoid~~ $a_i=\text{sigmoid}(z_i)$

الف) $\frac{\partial f(x)}{\partial z_1^{[2]}}=\frac{\partial f(x)}{\partial a_1^{[2]}}\times\frac{\partial a_1^{[2]}}{\partial z_1^{[2]}}=w_1^{[3]}\times a_1^{[2]}(1-a_1^{[2]})$

ب) $z^{[2]}=\begin{bmatrix}z_1^{[2]}\\ z_2^{[2]}\end{bmatrix}\Rightarrow z_2^{[2]}=\frac{\partial f(x)}{\partial a_2^{[2]}}\times\frac{\partial a_2^{[2]}}{z_2^{[2]}}=w_2^{[3]}\times a_2^{[2]}(1-a_2^{[2]})$

$\rightarrow\begin{bmatrix}w_1^{[3]}a_1^{[2]}(1-a_1^{[2]})\\ w_2^{[3]}a_2^{[2]}(1-a_2^{[2]})\end{bmatrix}$

ج) $\frac{\partial f(x)}{\partial z^{[1]}}$ and $z^{[1]}=\begin{bmatrix}z_1^{[1]}\\ z_2^{[1]}\end{bmatrix}$ then $\frac{\partial f(x)}{\partial z^{[1]}}=\begin{bmatrix}\frac{\partial f(x)}{\partial z_1^{[1]}}\\ \frac{\partial f(x)}{\partial z_2^{[1]}}\end{bmatrix}$

$=\frac{\partial f(x)}{\partial z_1^{[1]}}=\frac{\partial f(x)}{\partial a_1^{[2]}}\times\frac{\partial a_1^{[2]}}{\partial z_1^{[2]}}\times\frac{\partial z_1^{[2]}}{\partial a_1^{[1]}}\times\frac{\partial a_1^{[1]}}{\partial z_1^{[1]}}+\frac{\partial f(x)}{\partial a_2^{[2]}}\times\frac{\partial a_2^{[2]}}{\partial z_2^{[2]}}\times\frac{\partial z_2^{[2]}}{\partial a_1^{[1]}}\times\frac{\partial a_1^{[1]}}{\partial z_1^{[1]}}$

$w_1^{[3]}$　$a_1^{[2]}(1-a_1^{[2]})$　$w_{11}^{[2]}$　$a_1^{[1]}(1-a_1^{[1]})$　$w_2^{[3]}$　$a_2^{[2]}(1-a_2^{[2]})$　$w_{21}^{[2]}$　$a_1^{[1]}(1-a_1^{[1]})$

$\Rightarrow\frac{\partial f(x)}{\partial z_1^{[1]}}=w_1^{[3]}\times a_1^{[2]}(1-a_1^{[2]})\times w_{11}^{[2]}\times a_1^{[1]}(1-a_1^{[1]})+w_2^{[3]}\times a_2^{[2]}(1-a_2^{[2]})\times w_{21}^{[2]}\times a_1^{[1]}(1-a_1^{[1]})$ let's call it $\boxed{t_1}$

$\frac{\partial f(x)}{\partial z_2^{[1]}}=w_1^{[3]}\times a_1^{[2]}(1-a_1^{[2]})\times w_{12}^{[2]}\times a_2^{[1]}\times(1-a_2^{[1]})+w_2^{[3]}\times a_2^{[2]}(1-a_2^{[2]})\times w_{22}^{[2]}\times a_2^{[1]}(1-a_2^{[1]})$ let's call it $\boxed{t_2}$

then $\frac{\partial f(x)}{\partial z^{[1]}}=\begin{bmatrix}t_1\\ t_2\end{bmatrix}$

د) $\frac{\partial f(x)}{\partial w_{11}^{[1]}}=w_1^{[3]}\times a_1^{[2]}(1-a_1^{[2]})\times w_{11}^{[2]}\times a_1^{[1]}(1-a_1^{[1]})\times x_1+w_2^{[3]}\times a_2^{[2]}(1-a_2^{[2]})\times w_{21}^{[2]}\times a_1^{[1]}(1-a_1^{[1]})\times x_1$　$\rightarrow\frac{\partial z_1^{[1]}}{\partial w_{11}^{[1]}}=x_1\rightarrow z_1^{[1]}=w_{11}^{[1]}x_1+w_{12}^{[1]}x_2$

سؤال سوم :

$\hat{y}=\text{softmax}(W_sx+b_s)$　$\hat{y}_1=\frac{e^{z_1}}{e^{z_1}+e^{z_2}},\ \hat{y}_2=\frac{e^{z_2}}{e^{z_1}+e^{z_2}}$

$z_1=w_1x+b_1$　$z_2=w_2x+b_2$　if $\hat{y}_1\geq\hat{y}_2$ then 0 else 1

$\frac{e^{z_1}}{e^{z_1}+e^{z_2}}\geq\frac{e^{z_2}}{e^{z_1}+e^{z_2}}\xrightarrow{\times\frac{e^{z_1}+e^{z_2}}{1}}e^{z_1}\geq e^{z_2}\xrightarrow{Ln}z_1\geq z_2$

$\rightarrow w_1x+b_1\geq w_2x+b_2\rightarrow\underbrace{(w_1-w_2)}_{wL}x+\underbrace{(b_1-b_2)}_{bL}\geq0$

$wLx+bL\geq0$

از طرفی مثل که نورون سیگموید را همان آنگونه نوشت ← $\hat{y}=\text{sigmoid}(\underbrace{w_Lx+b_L}_{z})\rightarrow\hat{y}\ \cancel{\geq}\ ۰/5\Rightarrow z\leq0$

$\Rightarrow z_1-z_2\geq0\rightarrow0$

$z\leq0\rightarrow0$　$\Rightarrow z=-(z_1-z_2)=z_2-z_1$

$z=(w_2x+b_2)-(w_1x+b_1)\Rightarrow\boxed{z=(w_2-w_1)x+(b_2-b_1)}$

sigmoid: $\frac{1}{1+e^{-z}}$

softmax: $\frac{e^{z_1}}{e^{z_1}+e^{z_2}},\ \frac{e^{z_2}}{e^{z_1}+e^{z_2}}$　از طرفی میتوان معادله softmax را به sigmoid تبدیل کرد　$\xrightarrow{\text{تقسیم بر } e^{z_1}}\boxed{\frac{1}{1+e^{z_2-z_1}}},\ \boxed{\frac{e^{z_2-z_1}}{1+e^{z_2-z_1}}}\rightarrow$ متمم

$1-\frac{e^{z_2-z_1}}{1+e^{z_2-z_1}}=\frac{1}{1+e^{z_2-z_1}}$

$\frac{1}{1+e^{-(z_1-z_2)}}$ ← فرمول سیگموید اگر $z=(z_1-z_2)$

if $z=z_1-z_2$ then $-z$ is z_2-z_1 which is equal to

$z=(w_2-w_1)x+(b_2-b_1)$