Deep Learning101

Softmax와 Cross Entropy 미분

Ans: $\hat{y} - y$

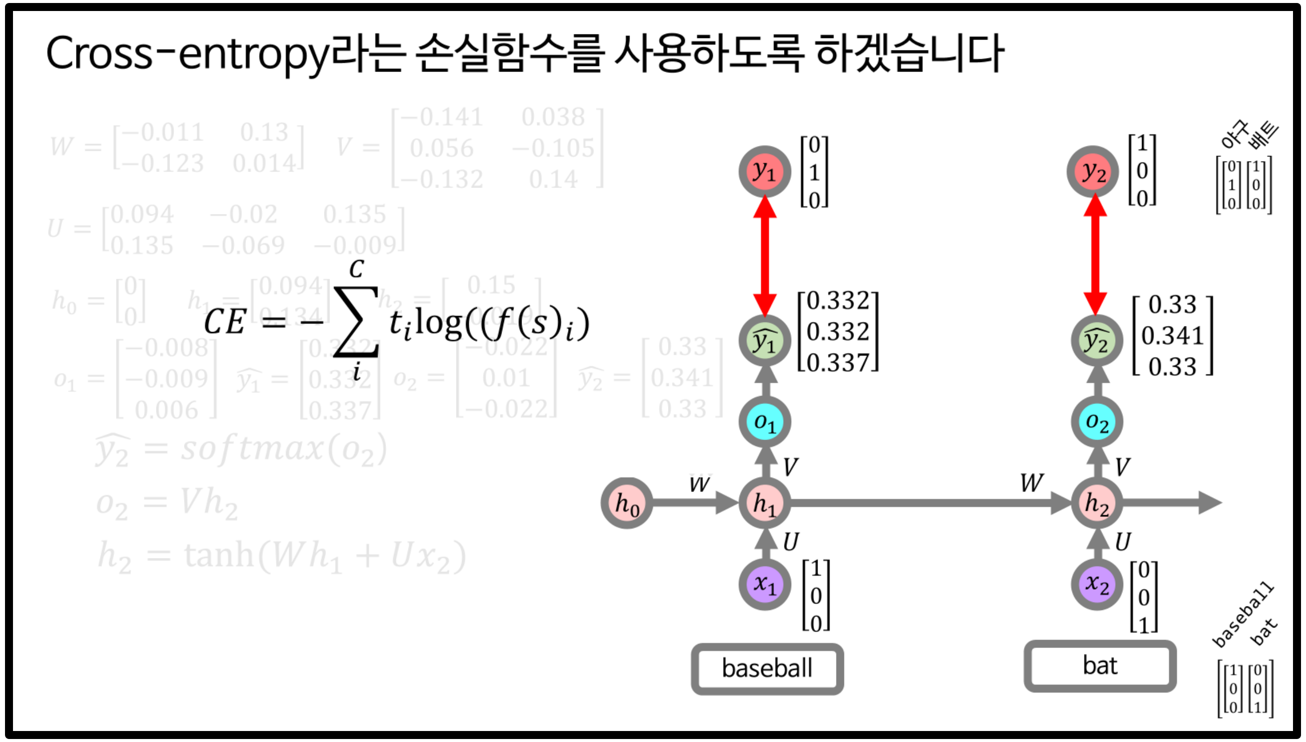
신박AI

안녕하세요 여러분! 신박AI입니다

오늘은 지난 영상에서 잠시 소개드렸던,
Softmax와 Cross-entropy의 미분값에
대해 말씀을 드리고자 합니다

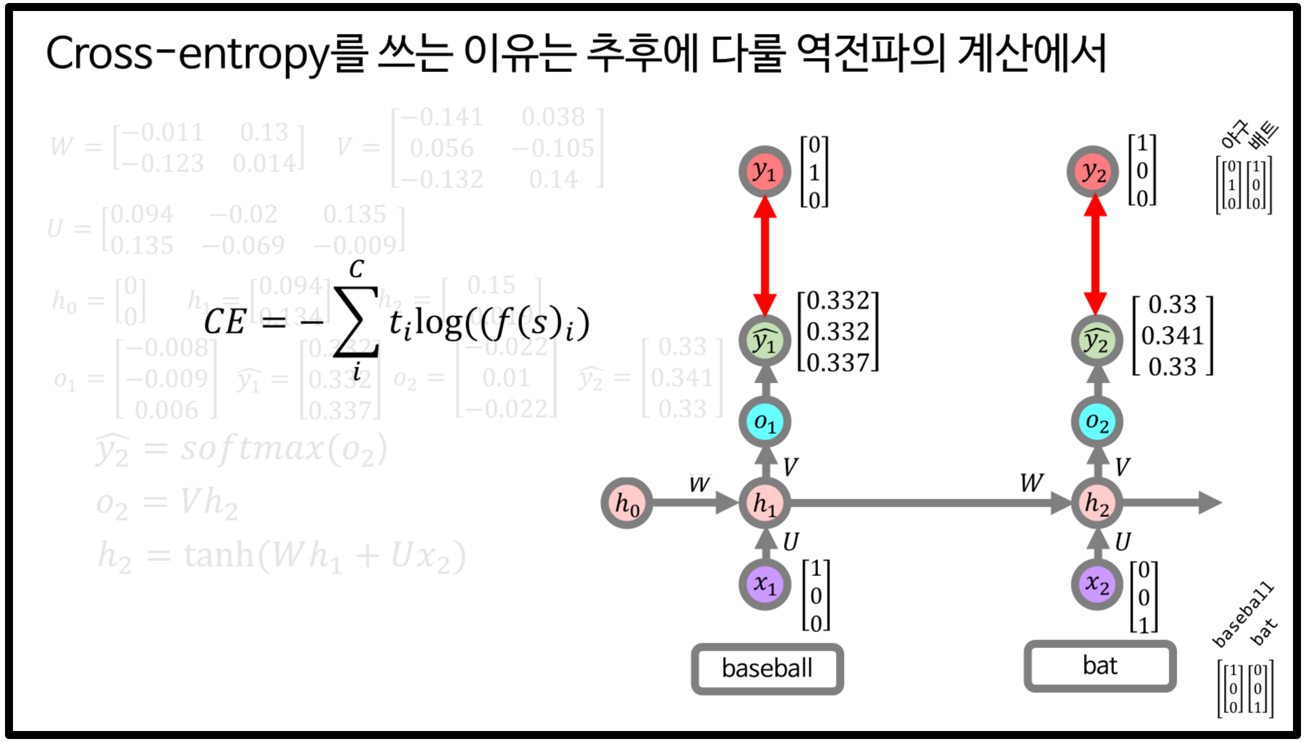
신박AI

우리는 지난 RNN영상에서 Cross-entropy를 손실함수로 쓰는 이유를,

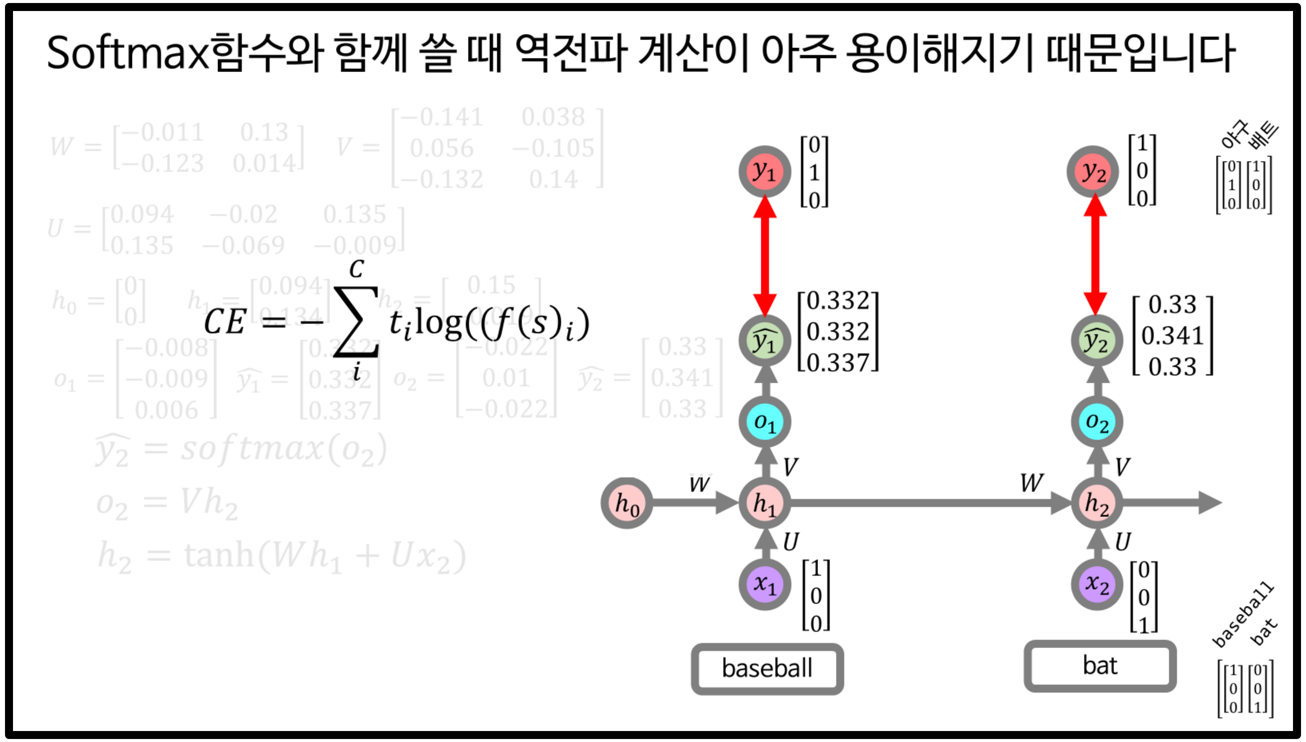


신박AI

역전파 계산에서,


Cross-entropy를 쓰는 이유는 추후에 다룰 역전파의 계산에서
$CE = -\sum_{i}^{C} t_i \log((f(s)_i)$
$\widehat{y_2} = softmax(o_2)$
$o_2 = Vh_2$
$h_2 = \tanh(Wh_1 + Ux_2)$
baseball
bat
야구
배트

Softmax함수와 함께 쓸 때 그 계산이 아주 용이해지기 때문이라고 말씀을 드렸습니다.



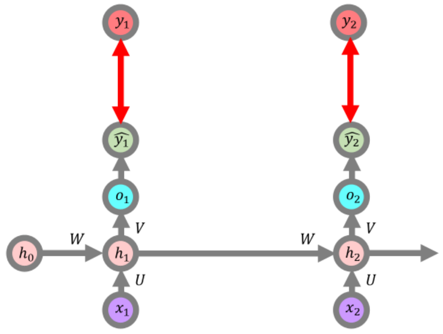
그리고 그 계산값은 $\hat{y} - y$ 이 된다고만 말씀을 드렸고, 그 자세한 도출 과정은 시간 관계상 생략했었습니다.

이 부분은 제가 다른 영상에서 도출과정을 보여드리려고 합니다.
많이 기대해주세요 ;)

$$\frac{\partial L}{\partial V} = \frac{\partial L_1}{\partial V} + \frac{\partial L_2}{\partial V}$$

$$\frac{\partial L}{\partial V} = \frac{\partial L_1}{\partial \hat{y}_1}\frac{\partial \hat{y}_1}{\partial o_1}\frac{\partial o_1}{\partial V} + \frac{\partial L_2}{\partial \hat{y}_2}\frac{\partial \hat{y}_2}{\partial o_2}\frac{\partial o_2}{\partial V}$$

$$\frac{\partial L}{\partial V} = (\hat{y}_1 - y_1)\frac{\partial o_1}{\partial V} + (\hat{y}_2 - y_2)\frac{\partial o_2}{\partial V}$$



LSTM 영상에서도 결과만 보여드리고 과정은 생략하였습니다.

$\hat{y}_t - y$는 $\partial L/\partial Z_t$ 가 됩니다

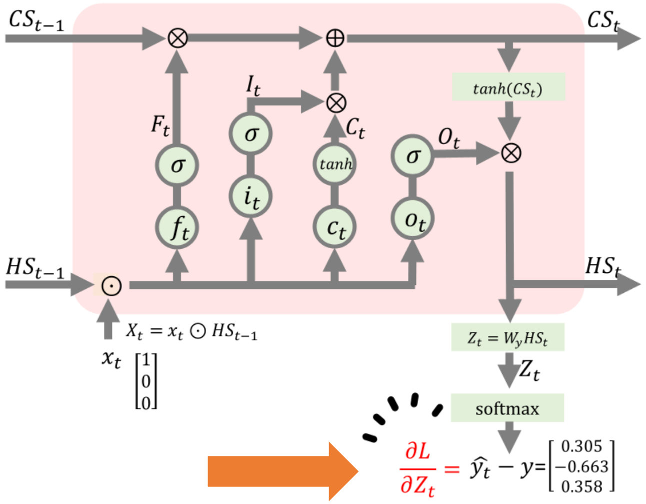
Forget Gate: $F_t = \sigma(W_f X_t)$
$F_t = \sigma(f_t)$
$f_t = W_f X_t$

Input Gate: $I_t = \sigma(W_i X_t)$
$I_t = \sigma(i_t)$
$i_t = W_i X_t$

Candidate Gate: $C_t = tanh(W_c X_t)$
$C_t = tanh(c_t)$
$c_t = W_c X_t$

Output Gate: $O_t = \sigma(W_o X_t)$
$O_t = \sigma(o_t)$
$o_t = W_o X_t$

$Z_t = W_y HS_t$



오늘 이 부분에 대해서 자세히 설명 드리고자 합니다.

$\hat{y}_t - y$는 $\partial L/\partial Z_t$ 가 됩니다

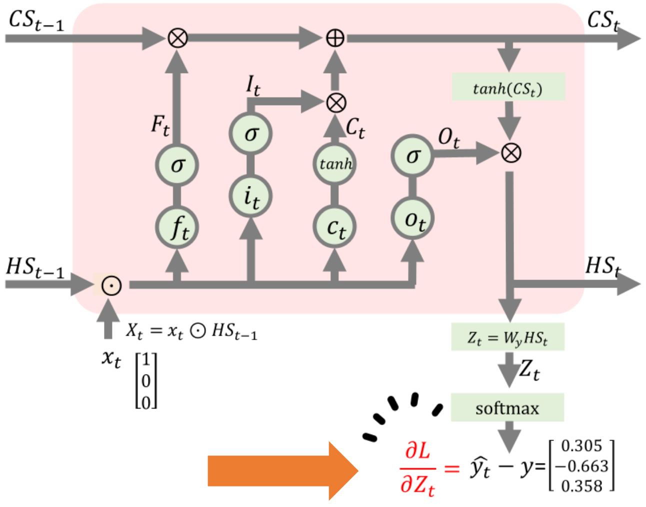
Forget Gate: $F_t = \sigma(W_f X_t)$
$F_t = \sigma(f_t)$
$f_t = W_f X_t$

Input Gate: $I_t = \sigma(W_i X_t)$
$I_t = \sigma(i_t)$
$i_t = W_i X_t$

Candidate Gate: $C_t = tanh(W_c X_t)$
$C_t = tanh(c_t)$
$c_t = W_c X_t$

Output Gate: $O_t = \sigma(W_o X_t)$
$O_t = \sigma(o_t)$
$o_t = W_o X_t$

$Z_t = W_y HS_t$



신박AI

Softmax와 Cross-entropy의 조합은
RNN이나 LSTM뿐만 아니라,

다중 클래스를 분류하는 여러 신경망에서
대표적으로 쓰이는 조합이기 때문에,

그 도출 과정을 알아두는 것은 여러모로
유익할 수 있다고 생각이 듭니다.

그러면 바로 시작하도록 하겠습니다.

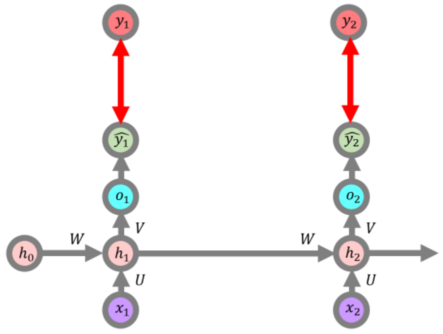
$$\frac{\partial L_1}{\partial \widehat{y_1}} \frac{\partial \widehat{y_1}}{\partial o_1}$$

이 부분은 제가 다른 영상에서 도출과정을 보여드리려고 합니다.
많이 기대해주세요 ;)

$$\frac{\partial L}{\partial V} = \frac{\partial L_1}{\partial V} + \frac{\partial L_2}{\partial V}$$

$$\frac{\partial L}{\partial V} = \frac{\partial L_1}{\partial \widehat{y_1}} \frac{\partial \widehat{y_1}}{\partial o_1} \frac{\partial o_1}{\partial V} + \frac{\partial L_2}{\partial \widehat{y_2}} \frac{\partial \widehat{y_2}}{\partial o_2} \frac{\partial o_2}{\partial V}$$

$$\frac{\partial L}{\partial V} = (\widehat{y_1} - y_1) \frac{\partial o_1}{\partial V} + (\widehat{y_2} - y_2) \frac{\partial o_2}{\partial V}$$



우리는 $\frac{\partial L_1}{\partial \widehat{y_1}} \frac{\partial \widehat{y_1}}{\partial o_1}$ 이 편미분 값이

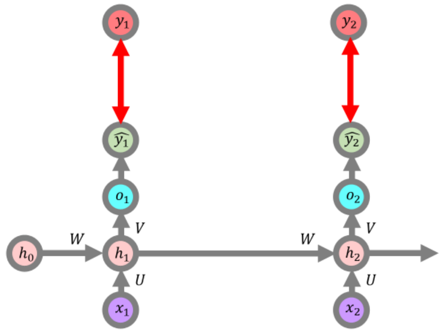
$$\frac{\partial L_1}{\partial \widehat{y_1}} \frac{\partial \widehat{y_1}}{\partial o_1}$$

이 부분은 제가 다른 영상에서 도출과정을 보여드리려고 합니다.
많이 기대해주세요 ;)

$$\frac{\partial L}{\partial V} = \frac{\partial L_1}{\partial V} + \frac{\partial L_2}{\partial V}$$

$$\frac{\partial L}{\partial V} = \frac{\partial L_1}{\partial \widehat{y_1}} \frac{\partial \widehat{y_1}}{\partial o_1} \frac{\partial o_1}{\partial V} + \frac{\partial L_2}{\partial \widehat{y_2}} \frac{\partial \widehat{y_2}}{\partial o_2} \frac{\partial o_2}{\partial V}$$

$$\frac{\partial L}{\partial V} = (\widehat{y_1} - y_1) \frac{\partial o_1}{\partial V} + (\widehat{y_2} - y_2) \frac{\partial o_2}{\partial V}$$



$\hat{y} - y$, 이 값으로 도출되는 과정을 알아보는 것입니다.

$$\frac{\partial L_1}{\partial \widehat{y_1}} \frac{\partial \widehat{y_1}}{\partial o_1}$$

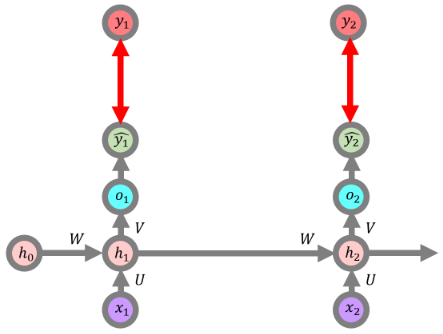
$$\hat{y} - y$$

이 부분은 제가 다른 영상에서 도출과정을 보여드리려고 합니다.
많이 기대해주세요 ;)

$$\frac{\partial L}{\partial V} = \frac{\partial L_1}{\partial V} + \frac{\partial L_2}{\partial V}$$

$$\frac{\partial L}{\partial V} = \frac{\partial L_1}{\partial \widehat{y_1}} \frac{\partial \widehat{y_1}}{\partial o_1} \frac{\partial o_1}{\partial V} + \frac{\partial L_2}{\partial \widehat{y_2}} \frac{\partial \widehat{y_2}}{\partial o_2} \frac{\partial o_2}{\partial V}$$

$$\frac{\partial L}{\partial V} = (\widehat{y_1} - y_1) \frac{\partial o_1}{\partial V} + (\widehat{y_2} - y_2) \frac{\partial o_2}{\partial V}$$



신박AI

여기에서 $\partial L_1/\partial\widehat{y_1}$ 은 구분하자면 Cross Entropy의 편미분 값이고,

$$\boxed{\frac{\partial L_1}{\partial\widehat{y_1}}}\frac{\partial\widehat{y_1}}{\partial o_1}$$

이 부분은 Softmax의 편미분 값입니다.

$$\frac{\partial L_1}{\partial \widehat{y_1}} \boxed{\frac{\partial \widehat{y_1}}{\partial o_1}}$$

신박AI

여기서 우리는 Softmax의 편미분 부터 구해보도록 하겠습니다.

$$\frac{\partial L_1}{\partial \widehat{y_1}} \boxed{\frac{\partial \widehat{y_1}}{\partial o_1}}$$

신박AI

일반적인 설명을 위해 숫자를 문자 i, j로 바꾸어 보겠습니다.

$$\frac{\partial L}{\partial \hat{y}} \boxed{\frac{\partial \widehat{y_i}}{\partial o_j}}$$

신박AI

Softmax의 공식은 다음과 같습니다.

$$\frac{\partial \hat{y}_i}{\partial o_j}$$

$$\hat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}}$$

신박AI

쉬운 설명을 위해 다음과 같이 세 개의 값들로 이루어진 행렬로 가정하겠습니다.

$$\frac{\partial \widehat{y_i}}{\partial o_j}$$

$$\widehat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

신박AI

Softmax 값인 $\widehat{y_i}$ 은 다음과 같이 계산할 수 있습니다.

$$\frac{\partial \widehat{y_i}}{\partial o_j}$$

$$\widehat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

신박AI

먼저 $\partial\widehat{y_1}/\partial o_1$ 부터 계산해보도록 하겠습니다.

$$\frac{\partial\widehat{y}_i}{\partial o_j}$$

$$\widehat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix}\widehat{y_1}\\\widehat{y_2}\\\widehat{y_3}\end{bmatrix} = softmax\left(\begin{bmatrix}o_1\\o_2\\o_3\end{bmatrix}\right)$$

$$\begin{bmatrix}\widehat{y_1}\\\widehat{y_2}\\\widehat{y_3}\end{bmatrix} = \begin{bmatrix}\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\\\frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}}\\\frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}}\end{bmatrix}$$

$$\frac{\partial\widehat{y_1}}{\partial o_1} =$$

신박AI

$\partial\widehat{y_1}/\partial o_1$ 는 다음과 같이 바꾸어 쓸 수가 있습니다.

$$\frac{\partial \hat{y}_i}{\partial o_j}$$

$$\hat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$\frac{\partial \widehat{y_1}}{\partial o_1} = \frac{\partial}{\partial o_1}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

신박AI

$\partial\widehat{y_1}/\partial o_1$ 를 구하기 위해서는 다음 미분공식이 필요합니다

$$\frac{\partial\widehat{y_i}}{\partial o_j}$$

$$\widehat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix}\widehat{y_1}\\\widehat{y_2}\\\widehat{y_3}\end{bmatrix} = softmax\left(\begin{bmatrix}o_1\\o_2\\o_3\end{bmatrix}\right)$$

$$\begin{bmatrix}\widehat{y_1}\\\widehat{y_2}\\\widehat{y_3}\end{bmatrix} = \begin{bmatrix}\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\\\frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}}\\\frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}}\end{bmatrix}$$

$$\frac{\partial\widehat{y_1}}{\partial o_1} = \frac{\partial}{\partial o_1}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

신박AI

여기서 e^{o_1} 을 $g(x)$ 로 보고, $e^{o_1} + e^{o_2} + e^{o_3}$ 을 $h(x)$로 본다면,

$$\frac{\partial \widehat{y_i}}{\partial o_j}$$

$$\widehat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$\frac{\partial \widehat{y_1}}{\partial o_1} = \frac{\partial}{\partial o_1}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

그러면, 다음과 같이 식을 대입할 수 있고,

$$\frac{\partial \widehat{y}_i}{\partial o_j}$$

$$\widehat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$\frac{\partial \widehat{y_1}}{\partial o_1} = \frac{\partial}{\partial o_1}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$= \frac{(e^{o_1})'(e^{o_1} + e^{o_2} + e^{o_3}) - e^{o_1}(e^{o_1} + e^{o_2} + e^{o_3})'}{(e^{o_1} + e^{o_2} + e^{o_3})(e^{o_1} + e^{o_2} + e^{o_3})}$$

신박AI

e^x의 미분값은 자기 자신이기 때문에,

$$\frac{\partial \hat{y}_i}{\partial o_j}$$

$$\hat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \hat{y_1} \\ \hat{y_2} \\ \hat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \hat{y_1} \\ \hat{y_2} \\ \hat{y_3} \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$f(x) = e^x \longrightarrow f'(x) = e^x$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx} f(x) = \frac{d}{dx} \frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$\frac{\partial \hat{y_1}}{\partial o_1} = \frac{\partial}{\partial o_1}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$= \frac{\boxed{(e^{o_1})'}(e^{o_1}+ e^{o_2} + e^{o_3}) - e^{o_1}(e^{o_1}+ e^{o_2} + e^{o_3})'}{(e^{o_1} + e^{o_2} + e^{o_3})(e^{o_1} + e^{o_2} + e^{o_3})}$$

신박AI

이렇게 바꾸어 쓸 수가 있습니다

$$\frac{\partial \hat{y}_i}{\partial o_j}$$

$$\hat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$f(x) = e^x \longrightarrow f'(x) = e^x$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx} f(x) = \frac{d}{dx} \frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$\frac{\partial \hat{y}_1}{\partial o_1} = \frac{\partial}{\partial o_1}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$= \frac{\boxed{e^{o_1}}\,(e^{o_1}+e^{o_2}+e^{o_3}) - e^{o_1}(e^{o_1}+e^{o_2}+e^{o_3})'}{(e^{o_1}+e^{o_2}+e^{o_3})(e^{o_1}+e^{o_2}+e^{o_3})}$$

신박AI

또한 이 부분은

$$\frac{\partial \hat{y}_i}{\partial o_j}$$

$$\hat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$\frac{\partial \hat{y}_1}{\partial o_1} = \frac{\partial}{\partial o_1}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$= \frac{e^{o_1}\,(e^{o_1}+e^{o_2}+e^{o_3}) - e^{o_1}\boxed{(e^{o_1}+e^{o_2}+e^{o_3})'}}{(e^{o_1}+e^{o_2}+e^{o_3})(e^{o_1}+e^{o_2}+e^{o_3})}$$

신박AI

o_1에 관한 편미분만 고려하기 때문에 e^{o_2}과 e^{o_3}은 그냥 사라지게 됩니다.

$$\frac{\partial \hat{y}_i}{\partial o_j}$$

$$\hat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx} f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$\frac{\partial \hat{y}_1}{\partial o_1} = \frac{\partial}{\partial o_1}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$= \frac{e^{o_1}\,(e^{o_1}+ e^{o_2} + e^{o_3}) - e^{o_1}\boxed{(e^{o_1}+ \cancel{e^{o_2}} + \cancel{e^{o_3}})'}}{(e^{o_1} + e^{o_2} + e^{o_3})(e^{o_1} + e^{o_2} + e^{o_3})}$$

신박AI

그래서 정리하면, $\partial\widehat{y_1}/\partial o_1$ 는 다음과 같이 바꾸어 쓸 수가 있습니다.

$$\frac{\partial \widehat{y_i}}{\partial o_j}$$

$$\widehat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$\frac{\partial \widehat{y_1}}{\partial o_1} = \frac{\partial}{\partial o_1}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$= \frac{e^{o_1}(e^{o_1}+e^{o_2}+e^{o_3}) - e^{o_1}e^{o_1}}{(e^{o_1}+e^{o_2}+e^{o_3})(e^{o_1}+e^{o_2}+e^{o_3})}$$

신박AI

e^{o_1} 을 이렇게 옮겨서 다음과 같이 항들을 새롭게 어레인지 해 보면

$$\frac{\partial \widehat{y}_i}{\partial o_j}$$

$$\widehat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$\frac{\partial \widehat{y_1}}{\partial o_1} = \frac{\partial}{\partial o_1}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$= \frac{e^{o_1}(e^{o_1}+e^{o_2}+e^{o_3}-\ e^{o_1})}{(e^{o_1}+e^{o_2}+e^{o_3})(e^{o_1}+e^{o_2}+e^{o_3})}$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

신박AI

$\partial \widehat{y_1} / \partial o_1$ 는 다음과 같이 바꾸어 쓸 수가 있습니다.

$$\frac{\partial \widehat{y}_i}{\partial o_j}$$

$$\widehat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$\frac{\partial \widehat{y_1}}{\partial o_1} = \frac{\partial}{\partial o_1}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx} f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$= \frac{e^{o_1}}{(e^{o_1} + e^{o_2} + e^{o_3})} \frac{(e^{o_1}+e^{o_2} + e^{o_3} - e^{o_1})}{(e^{o_1} + e^{o_2} + e^{o_3})}$$

그리고 이렇게 항들을 분리하면,

$$\frac{\partial \hat{y}_i}{\partial o_j}$$

$$\hat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx} f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$\frac{\partial \hat{y}_1}{\partial o_1} = \frac{\partial}{\partial o_1}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$= \frac{e^{o_1}}{(e^{o_1}+e^{o_2}+e^{o_3})} \frac{(e^{o_1}+e^{o_2}+e^{o_3} - e^{o_1})}{(e^{o_1}+e^{o_2}+e^{o_3})}$$

$$= \frac{e^{o_1}}{(e^{o_1}+e^{o_2}+e^{o_3})} \left(\frac{(e^{o_1}+e^{o_2}+e^{o_3})}{(e^{o_1}+e^{o_2}+e^{o_3})} - \frac{e^{o_1}}{(e^{o_1}+e^{o_2}+e^{o_3})} \right)$$

신박AI

$\partial\widehat{y_1}/\partial o_1$ 는 결국 $\widehat{y_1}\,(1-\widehat{y_1})$ 로 간단하게 바꿀 수 있습니다

$$\frac{\partial \hat{y}_i}{\partial o_j}$$

$$\hat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix}\widehat{y_1}\\ \widehat{y_2}\\ \widehat{y_3}\end{bmatrix} = softmax\left(\begin{bmatrix}o_1\\ o_2\\ o_3\end{bmatrix}\right)$$

$$\begin{bmatrix}\widehat{y_1}\\ \widehat{y_2}\\ \widehat{y_3}\end{bmatrix} = \begin{bmatrix}\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}}\\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}}\end{bmatrix}$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$\frac{\partial \widehat{y_1}}{\partial o_1} = \frac{\partial}{\partial o_1}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$= \frac{e^{o_1}}{(e^{o_1}+e^{o_2}+e^{o_3})}\,\frac{(e^{o_1}+e^{o_2}+e^{o_3} - e^{o_1})}{(e^{o_1}+e^{o_2}+e^{o_3})}$$

$$= \underbrace{\frac{e^{o_1}}{(e^{o_1}+e^{o_2}+e^{o_3})}}_{\widehat{y_1}}\left(\underbrace{\frac{(e^{o_1}+e^{o_2}+e^{o_3})}{(e^{o_1}+e^{o_2}+e^{o_3})}}_{1} - \underbrace{\frac{e^{o_1}}{(e^{o_1}+e^{o_2}+e^{o_3})}}_{\widehat{y_1}}\right) = \widehat{y_1}\,(1-\widehat{y_1})$$

이 전개는 $\partial\widehat{y_2}/\partial o_2$ 의 경우도 마찬가지입니다

$$\frac{\partial \widehat{y}_i}{\partial o_j}$$

$$\widehat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx} f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$\frac{\partial \widehat{y_2}}{\partial o_2} = \frac{\partial}{\partial o_2}\left(\frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$= \frac{e^{o_2}}{(e^{o_1}+e^{o_2}+e^{o_3})} \frac{(e^{o_1}+e^{o_2}+e^{o_3} - e^{o_2})}{(e^{o_1}+e^{o_2}+e^{o_3})}$$

$$= \underbrace{\frac{e^{o_2}}{(e^{o_1}+e^{o_2}+e^{o_3})}}_{\widehat{y_2}} \left(\underbrace{\frac{(e^{o_1}+e^{o_2}+e^{o_3})}{(e^{o_1}+e^{o_2}+e^{o_3})}}_{1} - \underbrace{\frac{e^{o_2}}{(e^{o_1}+e^{o_2}+e^{o_3})}}_{\widehat{y_2}} \right) = \widehat{y_2}\,(1 - \widehat{y_2})$$

신박AI

$\partial \widehat{y_3} / \partial o_3$ 도 마찬가지로 $\widehat{y_3}\ (1 - \widehat{y_3})$ 로 바꿀 수 있습니다

$$\frac{\partial \widehat{y_i}}{\partial o_j}$$

$$\widehat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx} f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$\frac{\partial \widehat{y_3}}{\partial o_3} = \frac{\partial}{\partial o_3}\left(\frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$= \frac{e^{o_3}}{(e^{o_1} + e^{o_2} + e^{o_3})} \frac{(e^{o_1}+e^{o_2} + e^{o_3} - e^{o_3})}{(e^{o_1} + e^{o_2} + e^{o_3})}$$

$$= \underbrace{\frac{e^{o_3}}{(e^{o_1} + e^{o_2} + e^{o_3})}}_{\widehat{y_3}} \left(\underbrace{\frac{(e^{o_1} + e^{o_2} + e^{o_3})}{(e^{o_1} + e^{o_2} + e^{o_3})}}_{1} - \underbrace{\frac{e^{o_3}}{(e^{o_1} + e^{o_2} + e^{o_3})}}_{\widehat{y_3}} \right) = \widehat{y_3}\ (1 - \widehat{y_3})$$

신박AI

그래서 일반적으로 $\partial \hat{y}_i / \partial o_j = \hat{y}_i\,(1-\hat{y}_i)$ 가 됩니다 ($i=j$일 경우)

$$\frac{\partial \hat{y}_i}{\partial o_j} = \begin{cases} \hat{y}_i\,(1-\hat{y}_i), if\ i=j \\ \end{cases} \qquad \hat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$\frac{\partial \hat{y}_i}{\partial o_j} = \frac{\partial}{\partial o_j}\left(\frac{e^{o_i}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

그렇다면 $i \neq j$ 의 경우라면 어떨까요?

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\ (1 - \widehat{y_i}), if\ i = j \\ \end{cases}$$

$$\widehat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \frac{\partial}{\partial o_j}\left(\frac{e^{o_i}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

신박AI

$\partial\widehat{y_1}/\partial o_2$ 를 한번 구해보도록 하겠습니다.

$$\frac{\partial\widehat{y_i}}{\partial o_j} = \begin{cases}\widehat{y_i}\,(1-\widehat{y_i}), if\ i = j \\ \end{cases}$$

$$\widehat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix}\widehat{y_1}\\ \widehat{y_2}\\ \widehat{y_3}\end{bmatrix} = softmax\left(\begin{bmatrix}o_1\\ o_2\\ o_3\end{bmatrix}\right)$$

$$\begin{bmatrix}\widehat{y_1}\\ \widehat{y_2}\\ \widehat{y_3}\end{bmatrix} = \begin{bmatrix}\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}}\\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}}\end{bmatrix}$$

$$\frac{\partial\widehat{y_1}}{\partial o_2} = \frac{\partial}{\partial o_2}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

신박AI

미분공식에 의해서 $\partial\widehat{y_1}/\partial o_2$ 는 다음과 같이 전개할 수 있고,

$$\frac{\partial\widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\,(1-\widehat{y_i}), if\ i=j \\ \end{cases} \quad \widehat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix}\widehat{y_1}\\ \widehat{y_2}\\ \widehat{y_3}\end{bmatrix} = softmax\left(\begin{bmatrix}o_1\\ o_2\\ o_3\end{bmatrix}\right)$$

$$\begin{bmatrix}\widehat{y_1}\\ \widehat{y_2}\\ \widehat{y_3}\end{bmatrix} = \begin{bmatrix}\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}}\\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}}\end{bmatrix}$$

$$\frac{\partial\widehat{y_1}}{\partial o_2} = \frac{\partial}{\partial o_2}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$= \frac{(e^{o_1})'(e^{o_1}+e^{o_2}+e^{o_3}) - e^{o_1}(e^{o_1}+e^{o_2}+e^{o_3})'}{(e^{o_1}+e^{o_2}+e^{o_3})(e^{o_1}+e^{o_2}+e^{o_3})}$$

신박AI

e^{o_1} 는 보시는 바와 같이 o_2 와는 아무런 연관이 없기 때문에,

$$\frac{\partial \hat{y_i}}{\partial o_j} = \begin{cases} \hat{y_i}\,(1-\hat{y_i}), if\ i=j \\ \end{cases} \qquad \hat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix}\hat{y_1}\\ \hat{y_2}\\ \hat{y_3}\end{bmatrix} = softmax\left(\begin{bmatrix}o_1\\ o_2\\ o_3\end{bmatrix}\right)$$

$$\begin{bmatrix}\hat{y_1}\\ \hat{y_2}\\ \hat{y_3}\end{bmatrix} = \begin{bmatrix}\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}}\\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}}\end{bmatrix}$$

$$\frac{\partial \hat{y_1}}{\partial o_2} = \frac{\partial}{\partial o_2}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$= \frac{\boxed{(e^{o_1})'}(e^{o_1}+e^{o_2}+e^{o_3}) - e^{o_1}(e^{o_1}+e^{o_2}+e^{o_3})'}{(e^{o_1}+e^{o_2}+e^{o_3})(e^{o_1}+e^{o_2}+e^{o_3})}$$

신박AI

$(e^{o_1})'$ 즉 $(= \partial e^{o_1}/\partial o_2)$ 은 0가 됩니다.

$$\frac{\partial \hat{y}_i}{\partial o_j} = \begin{cases} \hat{y}_i\,(1-\hat{y}_i), if\ i = j \\ \end{cases} \qquad \hat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$\frac{\partial \hat{y}_1}{\partial o_2} = \frac{\partial}{\partial o_2}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$= \frac{\boxed{0}\,(e^{o_1} + e^{o_2} + e^{o_3}) - e^{o_1}(e^{o_1} + e^{o_2} + e^{o_3})'}{(e^{o_1} + e^{o_2} + e^{o_3})(e^{o_1} + e^{o_2} + e^{o_3})}$$

신박AI

그러면 이 부분은 다 0이 되고,

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\,(1-\widehat{y_i}), if\ i=j \\ \end{cases}$$

$$\widehat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$\frac{\partial \widehat{y_1}}{\partial o_2} = \frac{\partial}{\partial o_2}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$= \frac{\cancel{\boxed{0}\ (e^{o_1}+e^{o_2}+e^{o_3})} - e^{o_1}(e^{o_1}+e^{o_2}+e^{o_3})'}{(e^{o_1}+e^{o_2}+e^{o_3})(e^{o_1}+e^{o_2}+e^{o_3})}$$

신박AI

또한 이 부분은

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\,(1-\widehat{y_i}), if\ i=j \\ \end{cases}$$

$$\widehat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$\frac{\partial \widehat{y_1}}{\partial o_2} = \frac{\partial}{\partial o_2}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$= \frac{\cancel{\boxed{0}\ (e^{o_1}+e^{o_2}+e^{o_3})} - e^{o_1}\boxed{(e^{o_1}+e^{o_2}+e^{o_3})'}}{(e^{o_1}+e^{o_2}+e^{o_3})(e^{o_1}+e^{o_2}+e^{o_3})}$$

신박AI

o_2 와 관련이 있는 부분은 e^{o_2} 밖에 없기 때문에

$$\frac{\partial \hat{y}_i}{\partial o_j} = \begin{cases} \hat{y}_i\,(1-\hat{y}_i), if\ i=j \\ \end{cases}$$

$$\hat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$\frac{\partial \hat{y}_1}{\partial o_2} = \frac{\partial}{\partial o_2}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx} f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$= \frac{\cancel{\boxed{0}\ (e^{o_1}+e^{o_2}+e^{o_3})} - e^{o_1}\boxed{(e^{o_1}+e^{o_2}+e^{o_3})'}}{(e^{o_1}+e^{o_2}+e^{o_3})(e^{o_1}+e^{o_2}+e^{o_3})}$$

신박AI

o_2 와 관련이 없는 부분들은 이렇게 없애버릴 수 있습니다.

$$\frac{\partial \hat{y}_i}{\partial o_j} = \begin{cases} \hat{y}_i\ (1-\hat{y}_i), if\ i=j \\ \end{cases}$$

$$\hat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix}\hat{y}_1\\ \hat{y}_2\\ \hat{y}_3\end{bmatrix} = softmax\left(\begin{bmatrix}o_1\\ o_2\\ o_3\end{bmatrix}\right)$$

$$\begin{bmatrix}\hat{y}_1\\ \hat{y}_2\\ \hat{y}_3\end{bmatrix} = \begin{bmatrix}\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}}\\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}}\end{bmatrix}$$

$$\frac{\partial \hat{y}_1}{\partial o_2} = \frac{\partial}{\partial o_2}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$= \frac{\cancel{\boxed{0}\ (e^{o_1}+e^{o_2}+e^{o_3})} - e^{o_1}\boxed{(\cancel{e^{o_1}}+e^{o_2}+\cancel{e^{o_3}})'}}{(e^{o_1}+e^{o_2}+e^{o_3})(e^{o_1}+e^{o_2}+e^{o_3})}$$

신박AI

그러면 $\partial\widehat{y_1}/\partial o_2$ 를 다음과 같이 새로이 정리할 수 있습니다.

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\,(1-\widehat{y_i}), if\ i=j \\ \end{cases} \qquad \widehat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix}\widehat{y_1}\\ \widehat{y_2}\\ \widehat{y_3}\end{bmatrix} = softmax\left(\begin{bmatrix}o_1\\ o_2\\ o_3\end{bmatrix}\right)$$

$$\begin{bmatrix}\widehat{y_1}\\ \widehat{y_2}\\ \widehat{y_3}\end{bmatrix} = \begin{bmatrix}\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}}\\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}}\end{bmatrix}$$

$$\frac{\partial \widehat{y_1}}{\partial o_2} = \frac{\partial}{\partial o_2}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$= - \frac{e^{o_1}}{(e^{o_1} + e^{o_2} + e^{o_3})} \frac{e^{o_2}}{(e^{o_1} + e^{o_2} + e^{o_3})}$$

그러면 각각의 항들을 다음처럼 $\widehat{y_1}$, $\widehat{y_2}$로 바꾸어 생각하면,

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\ (1 - \widehat{y_i}), if\ i = j \\ \end{cases}$$

$$\widehat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \widehat{y_3} \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx} f(x) = \frac{d}{dx} \frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$\frac{\partial \widehat{y_1}}{\partial o_2} = \frac{\partial}{\partial o_2}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$= -\frac{e^{o_1}}{(e^{o_1} + e^{o_2} + e^{o_3})} \frac{e^{o_2}}{(e^{o_1} + e^{o_2} + e^{o_3})}$$

신박AI

$\partial\widehat{y_1}/\partial o_2 = -\widehat{y_1}\cdot\widehat{y_2}$가 됩니다

$$\frac{\partial\widehat{y_i}}{\partial o_j} = \begin{cases}\widehat{y_i}\,(1-\widehat{y_i}), if\ i=j \\ \end{cases}$$

$$\widehat{y_i} = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix}\widehat{y_1}\\ \widehat{y_2}\\ \widehat{y_3}\end{bmatrix} = softmax\left(\begin{bmatrix}o_1\\ o_2\\ o_3\end{bmatrix}\right)$$

$$\begin{bmatrix}\widehat{y_1}\\ \widehat{y_2}\\ \widehat{y_3}\end{bmatrix} = \begin{bmatrix}\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}}\\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}}\end{bmatrix}$$

$$\frac{\partial\widehat{y_1}}{\partial o_2} = \frac{\partial}{\partial o_2}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx}f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$= -\frac{e^{o_1}}{(e^{o_1}+e^{o_2}+e^{o_3})}\,\frac{e^{o_2}}{(e^{o_1}+e^{o_2}+e^{o_3})}$$

$$= -\widehat{y_1}\cdot\widehat{y_2}$$

신박AI

그래서 일반적인 $i \neq j$ 의 경우, $\partial \hat{y}_i / \partial o_j = -\hat{y}_i \cdot \hat{y}_j$가 됩니다

$$\frac{\partial \hat{y}_i}{\partial o_j} = \begin{cases} \hat{y}_i\,(1-\hat{y}_i), & if\ i = j \\ \hat{y}_i \cdot \hat{y}_j, & if\ i \neq j \end{cases} \qquad \hat{y}_i = \frac{e^{o_i}}{\sum_{k=1}^{N} e^{o_k}} \longrightarrow \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = softmax\left(\begin{bmatrix} o_1 \\ o_2 \\ o_3 \end{bmatrix}\right)$$

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \end{bmatrix} = \begin{bmatrix} \frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_2}}{e^{o_1}+e^{o_2}+e^{o_3}} \\ \frac{e^{o_3}}{e^{o_1}+e^{o_2}+e^{o_3}} \end{bmatrix}$$

$$\frac{\partial \hat{y}_1}{\partial o_2} = \frac{\partial}{\partial o_2}\left(\frac{e^{o_1}}{e^{o_1}+e^{o_2}+e^{o_3}}\right)$$

$$f(x) = \frac{g(x)}{h(x)} \longrightarrow f'(x) = \frac{d}{dx} f(x) = \frac{d}{dx}\frac{g(x)}{h(x)} = \frac{g'(x)h(x) - g(x)h'(x)}{\{h(x)\}^2}$$

$$= -\frac{e^{o_1}}{(e^{o_1} + e^{o_2} + e^{o_3})}\ \frac{e^{o_2}}{(e^{o_1} + e^{o_2} + e^{o_3})}$$

$$= -\hat{y}_1 \cdot \hat{y}_2$$

신박AI

자 여기까지가 Softmax 기울기를 구하는 과정입니다.

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\,(1 - \widehat{y_i}), & if\ i = j \\ -\widehat{y_i} \cdot \widehat{y_j}, & if\ i \neq j \end{cases}$$

신박AI

자 여기까지가 Softmax 기울기를 구하는 과정입니다.

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\,(1-\widehat{y_i}), & if\ i = j \\ -\widehat{y_i} \cdot \widehat{y_j}, & if\ i \neq j \end{cases}$$

신박AI

Softmax 기울기는 이렇게 구석에 잠시 두고,

$$\frac{\partial \hat{y}_i}{\partial o_j} = \begin{cases} \hat{y}_i\,(1-\hat{y}_i), & if\ i = j \\ -\hat{y}_i \cdot \hat{y}_j, & if\ i \neq j \end{cases}$$

신박AI

이제는 이 결과를 바탕으로 Cross-Entropy + Softmax의 편미분을 구해보도록 하겠습니다

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\,(1 - \widehat{y_i}), & if\ i = j \\ -\widehat{y_i} \cdot \widehat{y_j}, & if\ i \neq j \end{cases}$$

$$\frac{\partial L}{\partial o_j}$$

신박AI

크로스 엔트로피 손실 함수는 다음과 같습니다

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\ (1 - \widehat{y_i}), if\ i = j \\ -\widehat{y_i} \cdot\ \widehat{y_j}, \qquad if\ i \neq j \end{cases}$$

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial L}{\partial o_j}$$

신박AI

$\partial L/\partial o_j$값은 다음과 같이 구할 수 있습니다

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\,(1-\widehat{y_i}), if\ i = j \\ -\widehat{y_i}\cdot\ \widehat{y_j}, \qquad if\ i \neq j \end{cases}$$

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = \frac{\partial}{\partial o_j}(-\sum_{k=1}^{N} y_k\, log(\widehat{y_k}))$$

신박AI

그리고 편미분 부분을 보기 좋게 안쪽으로 넣어보겠습니다

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\,(1-\widehat{y_i}), if\ i=j \\ -\widehat{y_i}\cdot\ \widehat{y_j}, \qquad if\ i \neq j \end{cases}$$

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = \frac{\partial}{\partial o_j}(-\sum_{k=1}^{N} y_k\, log(\widehat{y_k}))$$

신박AI

그리고 편미분 부분을 보기 좋게 안쪽으로 넣어보겠습니다

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\ (1 - \widehat{y_i}), if\ i = j \\ -\widehat{y_i} \cdot \widehat{y_j}, \quad\quad if\ i \neq j \end{cases}$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \frac{\partial log(\widehat{y_k})}{\partial o_j}$$

신박AI

그리고 체인룰을 가미하면 다음과 같이 바꿀 수 있습니다.

$$\frac{\partial \hat{y_i}}{\partial o_j} = \begin{cases} \hat{y_i}\,(1-\hat{y_i}), if\ i=j \\ -\hat{y_i}\cdot\ \hat{y_j}, \qquad if\ i \neq j \end{cases}$$

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \frac{\partial log(\widehat{y_k})}{\partial o_j}$$

신박AI

그리고 체인룰을 가미하면 다음과 같이 바꿀 수 있습니다.

$$\frac{\partial \hat{y}_i}{\partial o_j} = \begin{cases} \hat{y}_i \, (1 - \hat{y}_i), if \;\; i = j \\ -\hat{y}_i \cdot \; \hat{y}_j, \qquad\quad if \;\; i \neq j \end{cases}$$

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \frac{\partial log(\widehat{y_k})}{\partial \widehat{y_k}} \frac{\partial \widehat{y_k}}{\partial o_j}$$

신박AI

그러면 이 부분은 log 함수의 미분 공식에 의해서..

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\,(1-\widehat{y_i}), if\ i=j \\ -\widehat{y_i}\cdot\ \widehat{y_j}, \quad\quad if\ i\neq j \end{cases}$$

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \frac{\partial log(\widehat{y_k})}{\partial \widehat{y_k}} \frac{\partial \widehat{y_k}}{\partial o_j}$$

$$\frac{d(\log(x))}{dx} = (log(x))' = \frac{1}{x}$$

신박AI

이 부분은..

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\,(1 - \widehat{y_i}), if\ i = j \\ -\widehat{y_i} \cdot \widehat{y_j}, \qquad if\ i \neq j \end{cases}$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \boxed{\frac{\partial log(\widehat{y_k})}{\partial \widehat{y_k}}} \frac{\partial \widehat{y_k}}{\partial o_j}$$

$$\frac{d(\log(x))}{dx} = (log(x))' = \frac{1}{x}$$

신박AI

이렇게 바꿀 수 있습니다.

$$\frac{\partial \hat{y}_i}{\partial o_j} = \begin{cases} \hat{y}_i\,(1-\hat{y}_i), if\ i=j \\ -\hat{y}_i \cdot \hat{y}_j, \quad\quad if\ i \neq j \end{cases}$$

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \boxed{\frac{1}{\widehat{y_k}}} \frac{\partial \widehat{y_k}}{\partial o_j}$$

$$\frac{d(\log(x))}{dx} = (log(x))' = \frac{1}{x}$$

신박AI

그 다음 이 부분은 이미 우리가 앞에서 구한 바가 있습니다.

$$\boxed{\frac{\partial \hat{y}_i}{\partial o_j}} = \begin{cases} \hat{y}_i\,(1-\hat{y}_i), if\ i = j \\ -\hat{y}_i \cdot\ \hat{y}_j, \qquad if\ i \neq j \end{cases}$$

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \quad \frac{1}{\widehat{y_k}} \quad \boxed{\frac{\partial \widehat{y_k}}{\partial o_j}}$$

$$\frac{d(\log(x))}{dx} = (log(x))' = \frac{1}{x}$$

그런데 $\partial\widehat{y_k}/\partial o_j$ 는 k값의 변화에 따라 $i = j$ 가 될 수도 있고, $i \neq j$ 가 될 수도 있기 때문에,

$$\boxed{\frac{\partial\widehat{y}_i}{\partial o_j}} = \begin{cases} \widehat{y}_i\,(1-\widehat{y}_i), if\ i=j \\ -\widehat{y}_i \cdot \widehat{y}_j, \quad if\ i \neq j \end{cases}$$

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \quad \frac{1}{\widehat{y_k}} \quad \boxed{\frac{\partial\widehat{y_k}}{\partial o_j}} \qquad \frac{d(\log(x))}{dx} = (log(x))' = \frac{1}{x}$$

Σ로만 연결 되어 있기 때문에, 이렇게 나누어 쓸 수가 있습니다

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\boxed{\frac{\partial \widehat{y_i}}{\partial o_j}} = \begin{cases} \widehat{y_i}\,(1-\widehat{y_i}), if\ i = j \\ -\widehat{y_i} \cdot \widehat{y_j}, \qquad if\ i \neq j \end{cases}$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \quad \frac{1}{\widehat{y_k}} \quad \boxed{\frac{\partial \widehat{y_k}}{\partial o_j}}$$

$$\frac{d(\log(x))}{dx} = (log(x))' = \frac{1}{x}$$

$$= \quad -y_j \frac{1}{\widehat{y_j}} \widehat{y_j}\,(1-\widehat{y_j})$$ ← *when k = j*

$$+ \left(-\textstyle\sum_{k\neq j}^{N} y_k \frac{1}{\widehat{y_k}} \left(-\widehat{y_j} \cdot \widehat{y_k}\right)\right)$$ ← *when k ≠ j*

그러면 각각의 분자 분모를 상쇄하고 나면..

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\ (1-\widehat{y_i}), if\ i = j \\ -\widehat{y_i} \cdot \widehat{y_j}, \quad\quad if\ i \neq j \end{cases}$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \quad \frac{1}{\widehat{y_k}} \quad \frac{\partial \widehat{y_k}}{\partial o_j}$$

$$\frac{d(\log(x))}{dx} = (log(x))' = \frac{1}{x}$$

$$= \quad -y_j \frac{1}{\cancel{\widehat{y_j}}} \cancel{\widehat{y_j}}\ (1-\widehat{y_j})$$ ← $when\ k = j$

$$+ \left(-\sum_{k\neq j}^{N} y_k \frac{1}{\cancel{\widehat{y_k}}} \left(-\widehat{y_j} \cdot \cancel{\widehat{y_k}}\right)\right)$$ ← $when\ k \neq j$

신박AI

다음과 같이 정리할 수 있습니다.

$$\frac{\partial \hat{y}_i}{\partial o_j} = \begin{cases} \hat{y}_i\,(1-\hat{y}_i), if\ i = j \\ -\hat{y}_i \cdot \hat{y}_j, \quad\quad if\ i \neq j \end{cases}$$

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \quad \frac{1}{\widehat{y_k}} \quad \frac{\partial \widehat{y_k}}{\partial o_j}$$

$$\frac{d(\log(x))}{dx} = (log(x))' = \frac{1}{x}$$

$$= \quad -y_j \frac{1}{\cancel{\hat{y}_j}} \cancel{\hat{y}_j}\,(1-\hat{y}_j)$$

$\leftarrow$ $when\ k = j$

$$+ \left(-\textstyle\sum_{k\neq j}^{N} y_k \frac{1}{\cancel{\widehat{y_k}}} \left(-\hat{y}_j \cdot \cancel{\widehat{y_k}}\right)\right)$$

$\leftarrow$ $when\ k \neq j$

$$= \quad -y_j(1-\hat{y}_j) + \textstyle\sum_{k\neq j}^{N} y_k\,\hat{y}_j$$

신박AI

계속해서 정리를 다음과 같이 하면,

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\,(1-\widehat{y_i}), if\ i=j \\ -\widehat{y_i} \cdot \widehat{y_j}, \quad\quad if\ i \neq j \end{cases}$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \quad \frac{1}{\widehat{y_k}} \quad \frac{\partial \widehat{y_k}}{\partial o_j}$$

$$\frac{d(\log(x))}{dx} = (log(x))' = \frac{1}{x}$$

$$= \quad -y_j \frac{1}{\cancel{\widehat{y_j}}} \cancel{\widehat{y_j}}\,(1-\widehat{y_j}) \qquad \leftarrow \color{red}{when\ k=j}$$

$$+\,(-\textstyle\sum_{k\neq j}^{N} y_k \frac{1}{\cancel{\widehat{y_k}}}\, \widehat{y_j} \cdot \cancel{\widehat{y_k}}\,) \qquad \leftarrow \color{red}{when\ k \neq j}$$

$$= \quad -y_j(1-\widehat{y_j}) + \textstyle\sum_{k\neq j}^{N} y_k\, \widehat{y_j}$$

$$= \quad -y_j + y_j\, \widehat{y_j} \ + \textstyle\sum_{k\neq j}^{N} y_k\, \widehat{y_j}$$

신박AI

이렇게 정리를 할 수가 있습니다.

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\ (1-\widehat{y_i}), if\ i=j \\ -\widehat{y_i} \cdot \widehat{y_j}, \quad if\ i \neq j \end{cases}$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \quad \frac{1}{\widehat{y_k}} \quad \frac{\partial \widehat{y_k}}{\partial o_j} \qquad \frac{d(\log(x))}{dx} = (log(x))' = \frac{1}{x}$$

$$= -y_j \frac{1}{\cancel{\widehat{y_j}}} \cancel{\widehat{y_j}} (1-\widehat{y_j}) \quad \leftarrow when\ k=j$$

$$+ (-\textstyle\sum_{k\neq j}^{N} y_k \frac{1}{\cancel{\widehat{y_k}}} \widehat{y_j} \cdot \cancel{\widehat{y_k}}\ ) \quad \leftarrow when\ k \neq j$$

$$= -y_j(1-\widehat{y_j}) + \textstyle\sum_{k\neq j}^{N} y_k \widehat{y_j}$$

$$= -y_j + y_j \widehat{y_j} + \textstyle\sum_{k\neq j}^{N} y_k \widehat{y_j}$$

$$= -y_j + \widehat{y_j}(y_j + \textstyle\sum_{k\neq j}^{N} y_k)$$

신박AI

그런데 이 부분은 우리가 one-hot encoding으로 실제값을 하기로 했었기 때문에,

$$\frac{\partial \widehat{y_i}}{\partial o_j} = \begin{cases} \widehat{y_i}\,(1-\widehat{y_i}), if\ i=j \\ -\widehat{y_i}\cdot\widehat{y_j}, \qquad if\ i\neq j \end{cases}$$

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \quad \frac{1}{\widehat{y_k}} \quad \frac{\partial \widehat{y_k}}{\partial o_j} \qquad \frac{d(\log(x))}{dx} = (log(x))' = \frac{1}{x}$$

$$= -y_j \frac{1}{\cancel{\widehat{y_j}}} \cancel{\widehat{y_j}}\,(1-\widehat{y_j})$$ ← *when $k = j$*

$$+ \left(-\textstyle\sum_{k\neq j}^{N} y_k \frac{1}{\cancel{\widehat{y_k}}}\, \widehat{y_j}\cdot\cancel{\widehat{y_k}}\right)$$ ← *when $k \neq j$*

$$= -y_j(1-\widehat{y_j}) + \textstyle\sum_{k\neq j}^{N} y_k\,\widehat{y_j}$$

$$= -y_j + y_j\,\widehat{y_j} + \textstyle\sum_{k\neq j}^{N} y_k\,\widehat{y_j}$$

$$= -y_j + \widehat{y_j}\boxed{(y_j + \textstyle\sum_{k\neq j}^{N} y_k)}$$

신박AI

one-hot encoding은 모든 벡터의 합이 언제나 1이 됩니다.

$$\frac{\partial \hat{y_i}}{\partial o_j} = \begin{cases} \hat{y_i}\,(1-\hat{y_i}), if\ i=j \\ -\hat{y_i}\cdot\ \hat{y_j}, \qquad if\ i \neq j \end{cases}$$

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \quad \frac{1}{\widehat{y_k}} \quad \frac{\partial \widehat{y_k}}{\partial o_j} \qquad \frac{d(\log(x))}{dx} = (log(x))' = \frac{1}{x}$$

to	=[1,0,0,0,0,0,0,0,0,…,0]
be	=[0,1,0,0,0,0,0,0,0,…,0]
or	=[0,0,1,0,0,0,0,0,0,…,0]
not	=[0,0,0,1,0,0,0,0,0,…,0]
to	=[1,0,0,0,0,0,0,0,0,…,0]
be	=[0,1,0,0,0,0,0,0,0,…,0]
that	=[0,0,0,0,1,0,0,0,0,…,0]
is	=[0,0,0,0,0,1,0,0,0,…,0]
the	=[0,0,0,0,0,0,1,0,0,…,0]
question	=[0,0,0,0,0,0,0,1,0,…,0]

$$= -y_j \frac{1}{\cancel{\widehat{y_j}}} \cancel{\widehat{y_j}}\,(1-\widehat{y_j})$$ ← *when* $k = j$

$$+\,(-\textstyle\sum_{k\neq j}^{N} y_k \frac{1}{\cancel{\widehat{y_k}}}\ \widehat{y_j}\cdot\ \cancel{\widehat{y_k}}\,)$$ ← *when* $k \neq j$

$$= -y_j(1-\widehat{y_j}) + \textstyle\sum_{k\neq j}^{N} y_k\,\widehat{y_j}$$

$$= -y_j + y_j\,\widehat{y_j} + \textstyle\sum_{k\neq j}^{N} y_k\,\widehat{y_j}$$

$$= -y_j + \widehat{y_j}\boxed{(y_j + \textstyle\sum_{k\neq j}^{N} y_k)}$$

"1"

신박AI

그래서 마지막으로 정리하면 $\partial L/\partial o_j$은 $\hat{y}_j - y_j$ 로 최종 정리할 수가 있습니다!

$$L = -\sum_{k=1}^{N} y_k log(\widehat{y_k})$$

$$\frac{\partial \hat{y}_i}{\partial o_j} = \begin{cases} \hat{y}_i (1 - \hat{y}_i), if\ i = j \\ -\hat{y}_i \cdot \hat{y}_j, \quad if\ i \neq j \end{cases}$$

$$\frac{\partial L}{\partial o_j}$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \frac{1}{\widehat{y_k}} \frac{\partial \widehat{y_k}}{\partial o_j}$$

$$\frac{d(\log(x))}{dx} = (log(x))' = \frac{1}{x}$$

$$= -y_j \frac{1}{\cancel{\hat{y}_j}} \cancel{\hat{y}_j} (1 - \hat{y}_j) \quad \textcolor{red}{\leftarrow when\ k = j}$$

$$+ (-\textstyle\sum_{k \neq j}^{N} y_k \frac{1}{\cancel{\widehat{y_k}}} \hat{y}_j \cdot \cancel{\widehat{y_k}}) \quad \textcolor{red}{\leftarrow when\ k \neq j}$$

$$= -y_j (1 - \hat{y}_j) + \textstyle\sum_{k \neq j}^{N} y_k \hat{y}_j$$

$$= -y_j + y_j \hat{y}_j + \textstyle\sum_{k \neq j}^{N} y_k \hat{y}_j$$

$$= -y_j + \hat{y}_j (y_j + \textstyle\sum_{k \neq j}^{N} y_k) \textcolor{red}{= \hat{y}_j - y_j}$$

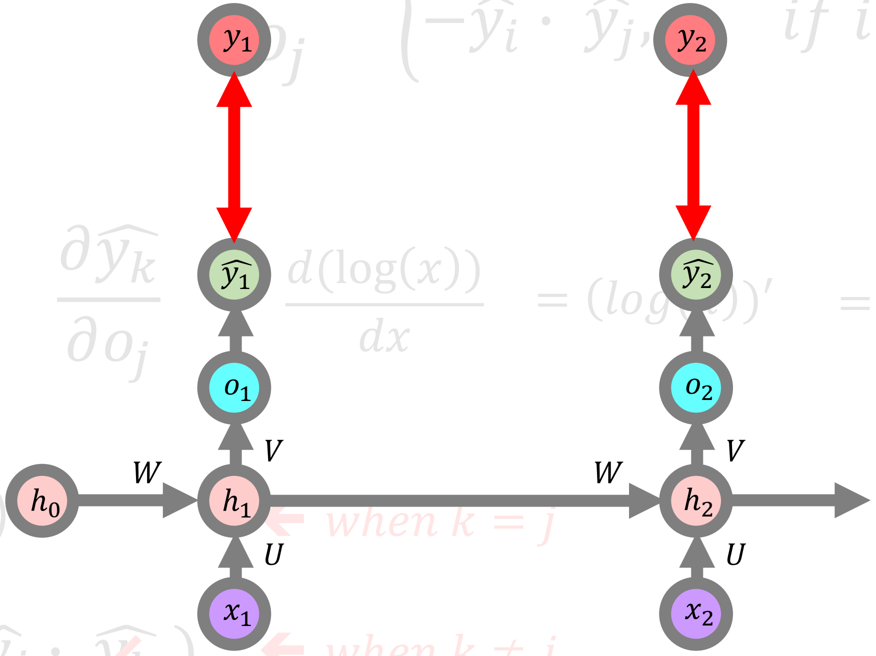
신박AI

여기까지 $\partial L/\partial o$가 $\widehat{y_j} - y_j$가 되는 과정에 대해 전개해 보았습니다

$$\frac{\partial L}{\partial V} = \frac{\partial L_1}{\partial V} + \frac{\partial L_2}{\partial V}$$

$$\frac{\partial L}{\partial V} = \boxed{\frac{\partial L_1}{\partial \widehat{y_1}} \frac{\partial \widehat{y_1}}{\partial o_1}} \frac{\partial o_1}{\partial V} + \boxed{\frac{\partial L_2}{\partial \widehat{y_2}} \frac{\partial \widehat{y_2}}{\partial o_2}} \frac{\partial o_2}{\partial V}$$

$$\frac{\partial L}{\partial V} = \boxed{(\widehat{y_1} - y_1)} \frac{\partial o_1}{\partial V} + \boxed{(\widehat{y_2} - y_2)} \frac{\partial o_2}{\partial V}$$



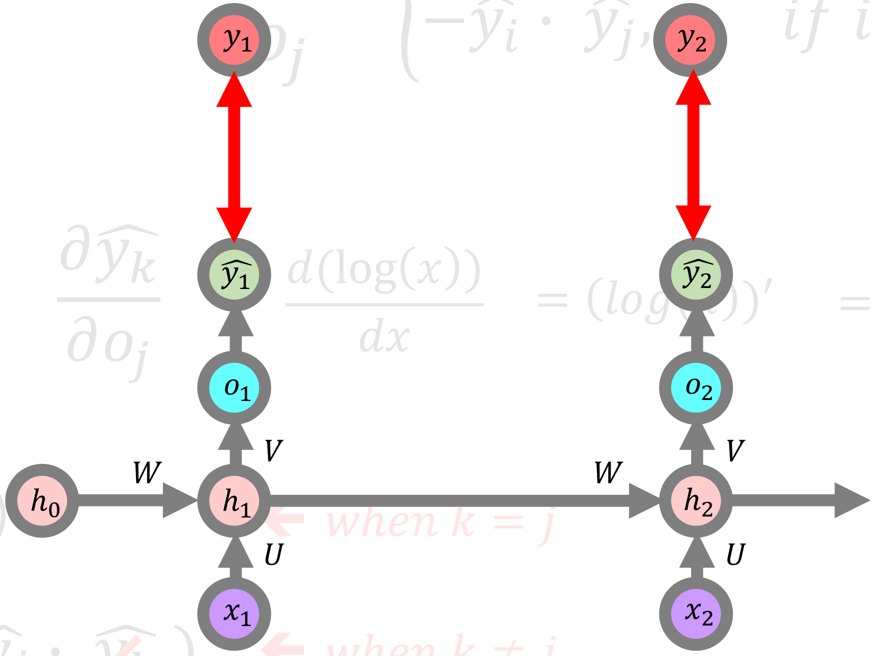
$$= \widehat{y_j} - y_j$$

과정은 좀 복잡하지만 최종적인 그래디언트는 softmax 확률값에 실제값을 빼주기만 하면 되니

$$\frac{\partial L}{\partial V} = \frac{\partial L_1}{\partial V} + \frac{\partial L_2}{\partial V}$$

$$\frac{\partial L}{\partial V} = \boxed{\frac{\partial L_1}{\partial \widehat{y_1}} \frac{\partial \widehat{y_1}}{\partial o_1}} \frac{\partial o_1}{\partial V} + \boxed{\frac{\partial L_2}{\partial \widehat{y_2}} \frac{\partial \widehat{y_2}}{\partial o_2}} \frac{\partial o_2}{\partial V}$$

$$\frac{\partial L}{\partial V} = \boxed{(\widehat{y_1} - y_1)} \frac{\partial o_1}{\partial V} + \boxed{(\widehat{y_2} - y_2)} \frac{\partial o_2}{\partial V}$$



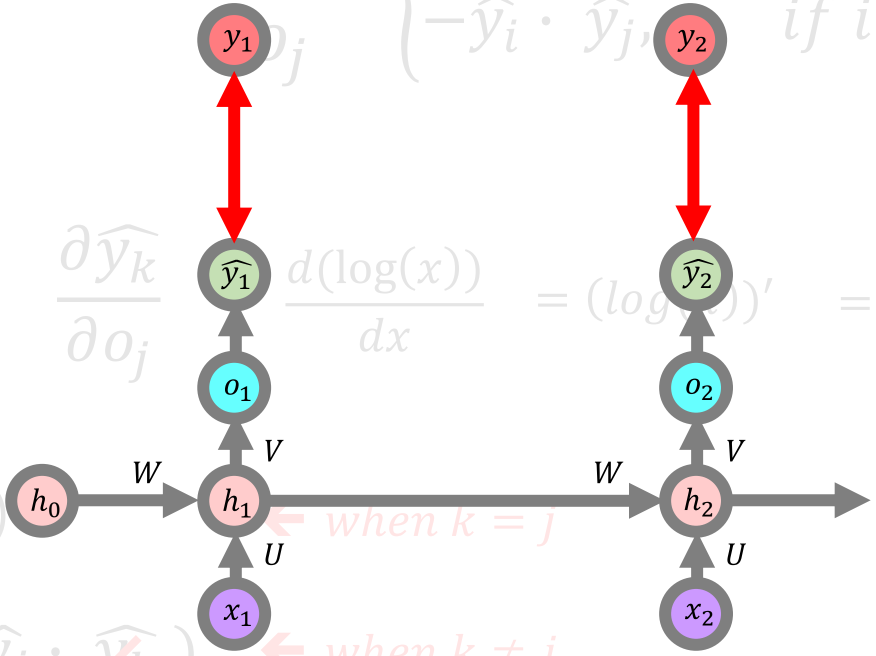
$$= \ \widehat{y_j} - y_j$$

Softmax + Cross-entropy 조합이 많이 활용되고 있는 것 같습니다.

$$\frac{\partial L}{\partial V} = \frac{\partial L_1}{\partial V} + \frac{\partial L_2}{\partial V}$$

$$\frac{\partial L}{\partial V} = \boxed{\frac{\partial L_1}{\partial \hat{y}_1}\frac{\partial \hat{y}_1}{\partial o_1}}\frac{\partial o_1}{\partial V} + \boxed{\frac{\partial L_2}{\partial \hat{y}_2}\frac{\partial \hat{y}_2}{\partial o_2}}\frac{\partial o_2}{\partial V}$$

$$\frac{\partial L}{\partial V} = \boxed{(\hat{y}_1 - y_1)}\frac{\partial o_1}{\partial V} + \boxed{(\hat{y}_2 - y_2)}\frac{\partial o_2}{\partial V}$$



$$\frac{\partial \hat{y}_i}{\partial o_j} = \begin{cases} \hat{y}_i(1-\hat{y}_i), & if\ i=j \\ -\hat{y}_i \cdot \hat{y}_j, & if\ i \neq j \end{cases}$$

$$L = -\sum_{k=1}^{N} y_k log(\hat{y}_k)$$

$$\frac{\partial L}{\partial o_j} = -\sum_{k=1}^{N} y_k \frac{1}{\hat{y}_k}\frac{\partial \hat{y}_k}{\partial o_j} \qquad \frac{d(\log(x))}{dx} = (\log(x))' = \frac{1}{x}$$

$$= -y_j \frac{1}{\cancel{\hat{y}_j}}\cancel{\hat{y}_j}(1-\hat{y}_j)$$ ← when $k = j$

$$+\left(-\sum_{k\neq j}^{N} y_k \frac{1}{\cancel{\hat{y}_k}}\hat{y}_j \cdot \cancel{\hat{y}_k}\right)$$ ← when $k \neq j$

$$= -y_j(1-\hat{y}_j) + \sum_{k\neq j}^{N} y_k \hat{y}_j$$

$$= -y_j + y_j\hat{y}_j + \sum_{k\neq j}^{N} y_k \hat{y}_j$$

$$= -y_j + \hat{y}_j\left(y_j + \sum_{k\neq j}^{N} y_k\right) = \hat{y}_j - y_j$$

신박AI

여기까지가 오늘 제가 준비한 softmax와
cross-entropy의 기울기에 관한
영상입니다.

본의 아니게 변변한 그림 하나 없이
수식만으로 설명 드리게 되어
죄송하게 생각합니다 ☹

다음 영상에서는 보다 더 재미있는 주제로 찾아뵙도록 하겠습니다. 계속 영상 시청 부탁드립니다. ☺

그럼 다음 시간에 또 만나요!

신박AI

감사합니다!

좋은 하루 되세요!!

이 채널은 여러분의 관심과 사랑이 필요합니다

신박AI

‘좋아요’와 ‘구독’버튼은 강의 준비에 큰 힘이 됩니다!

신박AI

그리고 영상 자료를 사용하실때는
출처 ‘신박AI’를 밝혀주세요

신박AI

신박AI