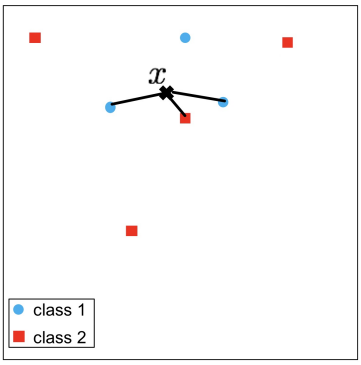
1.) x와 가장 가까운 요소 3가지를 보자 (∵k=3)



■ ● ● 순으로 가까운 요소이므로,
3개중 2개를 차지하는 ● 에 의해 x는 class 1으로 estimated 된다.

2)

s = Wx +b, $x = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$, $W = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$

2-a)

$$s = Wx + b$$
$$= \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$
$$= \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$
$$= \begin{bmatrix} 3 \\ 6 \\ 7 \end{bmatrix}$$

first element of s represents the score for "cat"
so the score for "cat" is 3

2-b)

$$s = Wx + b$$
$$= \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$
$$= \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$
$$= \begin{bmatrix} 3 \\ 6 \\ 7 \end{bmatrix}$$

second element of s represents the score for "dog"
so the score for "dog" is 6

2-c)

$$S = Wx + b$$

$$= \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} 3 \\ 6 \\ 7 \end{bmatrix}$$

third element of s represents the score for "ship"
so the score for "ship" is 7

3)

$$L = \sum_{j \neq y_i} \left(\max(0, s_j - s_{y_i} + 1) \right)$$

$$s_1 = \begin{bmatrix} 3 \\ 5 \\ -1 \end{bmatrix}, \quad s_2 = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}, \quad s_3 = \begin{bmatrix} 2 \\ 2 \\ -3 \end{bmatrix}$$

$$L_{s_1} = \max(5 - 3 + 1, 0) + \max(-1 - 3 + 1, 0) = 3$$

$$L_{s_2} = \max(1 - 4 + 1, 0) + \max(2 - 4 + 1, 0) = 0$$

$$L_{s_3} = \max(2 - (-3) + 1, 0) + \max(2 - (-3) + 1, 0) = 12$$

$$L = \frac{1}{3}(L_{s_1} + L_{s_2} + L_{s_3}) = \frac{1}{3}(3 + 0 + 12) = 5$$

4)

	Predicted class: Positive	Predicted class: Negative
actual class: Positive	85 = TP	15 = FN
actual class: Negative	890 = FP	10 = TN

$$\text{accuracy} = \frac{TP + TN}{TP + FP + FN + TN}$$

$$= \frac{85 + 10}{85 + 890 + 15 + 10}$$

$$= 0.095$$

$$\text{precision} = \frac{TP}{TP + FP}$$

$$= \frac{85}{85 + 890}$$

$$= 0.0872$$

$$\text{recall} = \frac{TP}{TP + FN}$$

$$= \frac{85}{85 + 15}$$

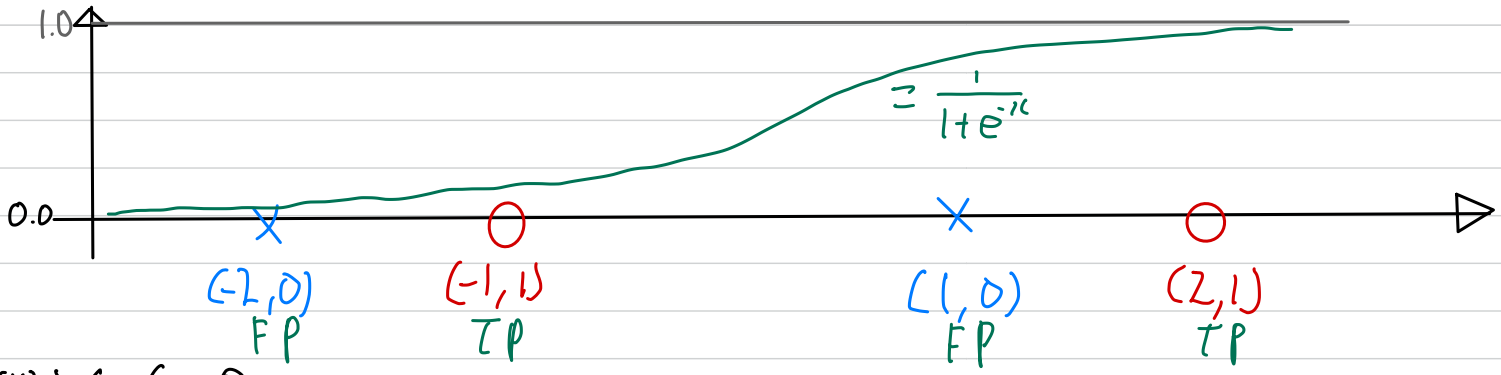
$$0.85$$

$$\text{F-score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

$$= \frac{2 \times 0.0872 \times 0.85}{0.0872 + 0.85}$$

$$= 0.1582$$

5)



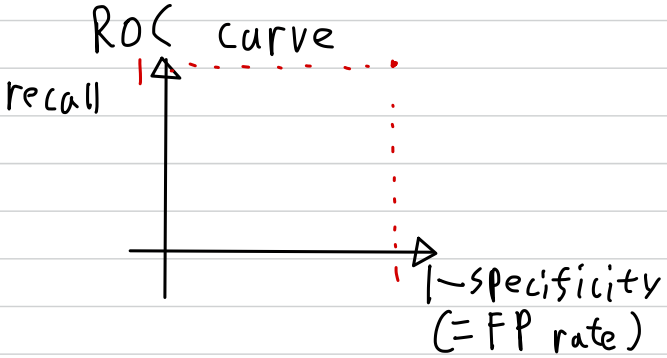
$f(x) > 5,\ 5 = 0$

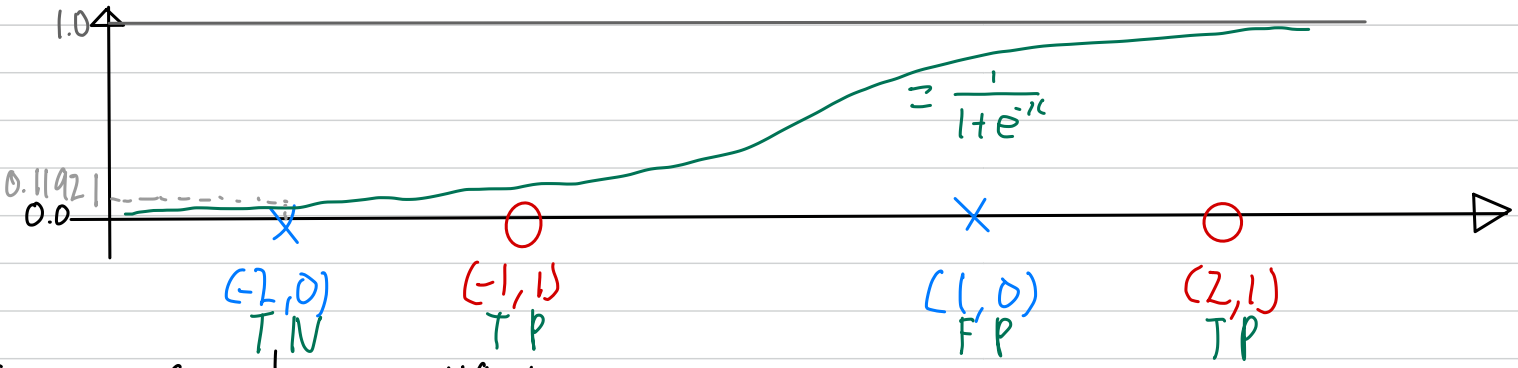
	Predicted as Positive	Predicted as Negative
Ground Truth Positive	2 = TP	0 = FN
Grounf Truth Negative	2 = FP	0 = TN

$$\text{recall} = \frac{TP}{TP+FN} = \frac{2}{2+0} = 1$$

$$\text{FP rate} = \frac{FP}{TN+FP} = \frac{2}{0+2} = 1$$

$\therefore (\text{FP rate}, \text{recall}) = (1, 1)$





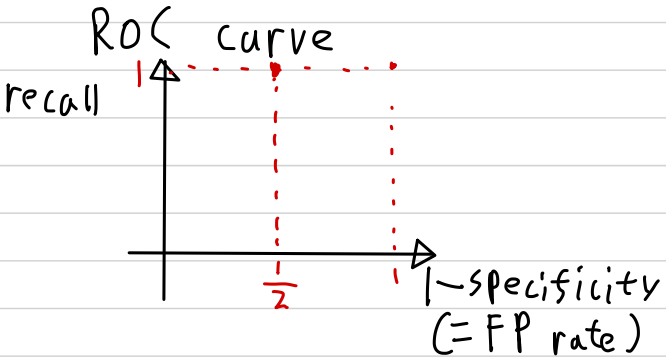
$f(x) > 5,\ 5 = \frac{1}{1+e^{-(-2)}} \approx 0.11921$

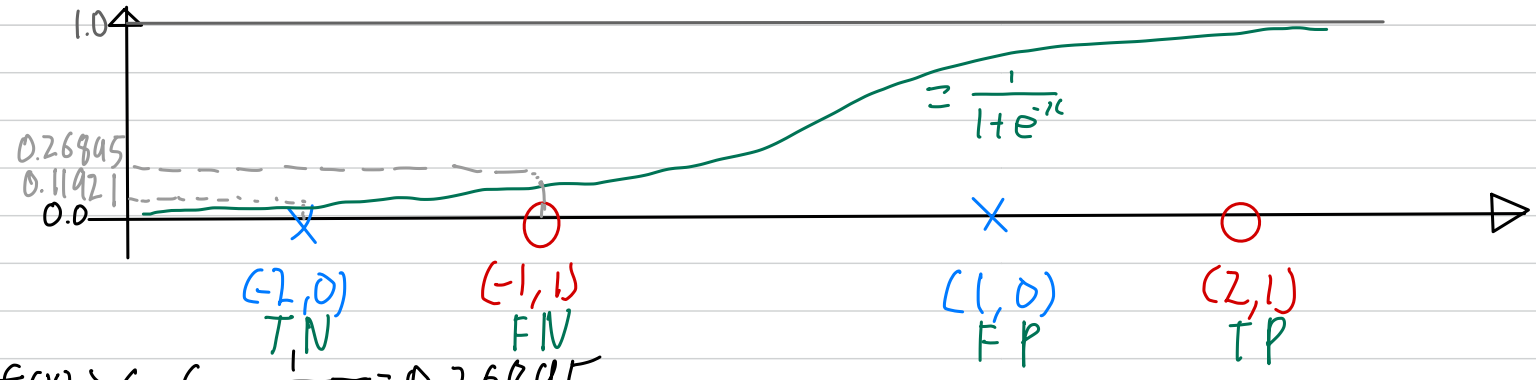
	Predicted as Positive	Predicted as Negative
Ground Truth Positive	2 = TP	0 = FN
Grounf Truth Negative	1 = FP	1 = TN

$$\text{recall} = \frac{TP}{TP+FN} = \frac{2}{2+0} = 1$$

$$\text{FP rate} = \frac{FP}{TN+FP} = \frac{1}{1+1} = \frac{1}{2}$$

$\therefore (\text{FP rate}, \text{recall}) = (\frac{1}{2}, 1)$





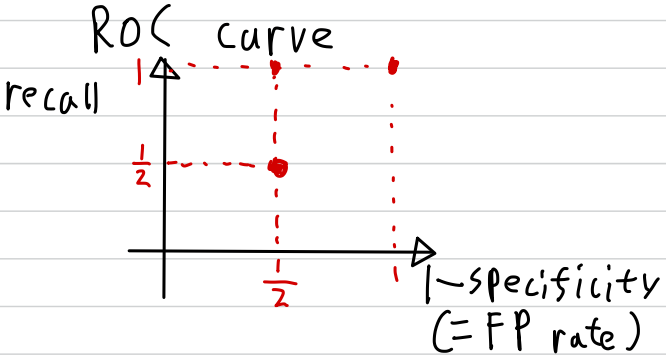
$f(x) > 5,\ S = \frac{1}{1+e^{-(-1)}} = 0.26895$

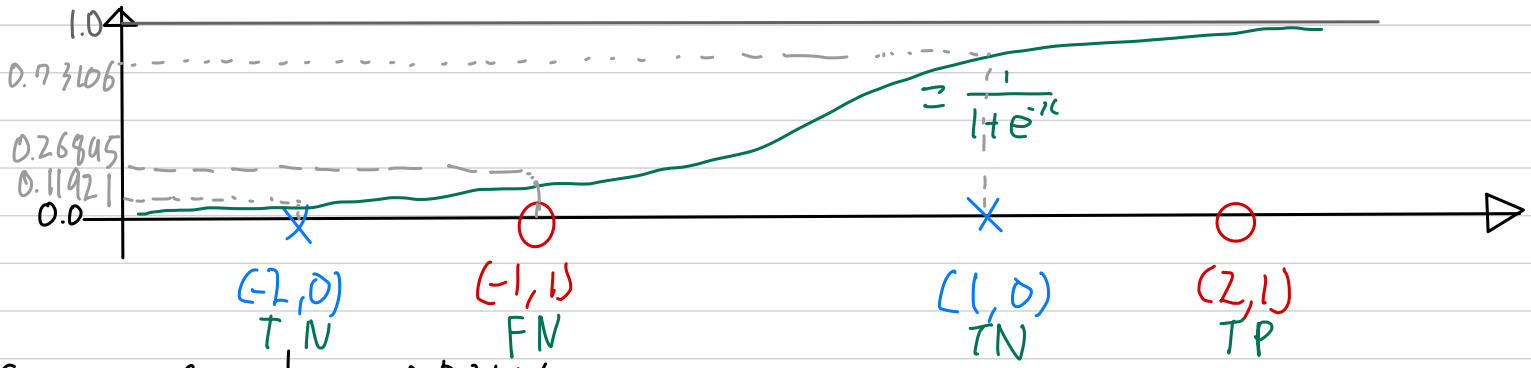
	Predicted as Positive	Predicted as Negative
Ground Truth Positive	1 = TP	1 = FN
Grounf Truth Negative	1 = FP	1 = TN

$$\text{recall} = \frac{TP}{TP+FN} = \frac{1}{1+1} = \frac{1}{2}$$

$$\text{FP rate} = \frac{FP}{TN+FP} = \frac{1}{1+1} = \frac{1}{2}$$

$\therefore (\text{FP rate}, \text{recall}) = (\frac{1}{2}, \frac{1}{2})$





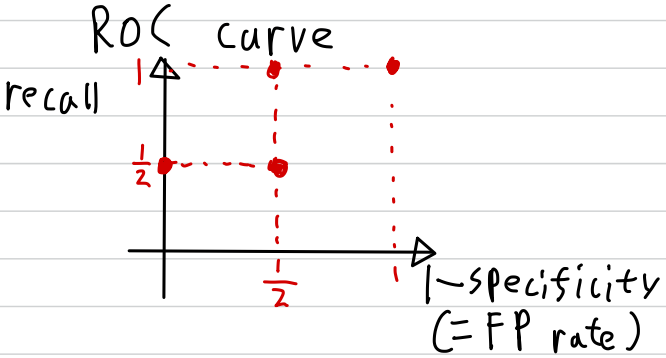
$f(x) > 5,\ S = \frac{1}{1+e^{-(1)}} = 0.73106$

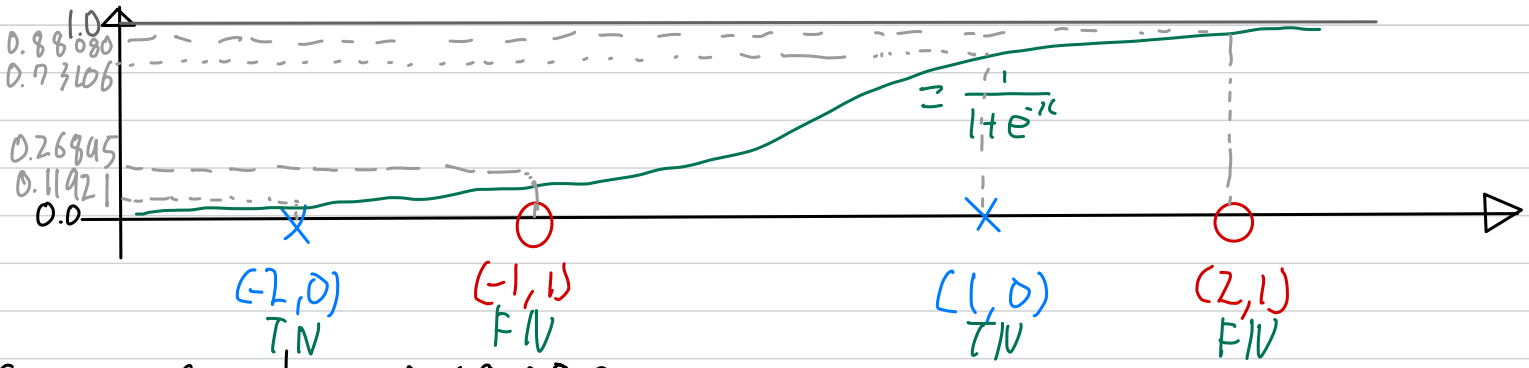
	Predicted as Positive	Predicted as Negative
Ground Truth Positive	1 = TP	1 = FN
Grounf Truth Negative	0 = FP	2 = TN

$$\text{recall} = \frac{TP}{TP+FN} = \frac{0}{2+0} = 0$$

$$\text{FP rate} = \frac{FP}{TN+FP} = \frac{1}{1+1} = \frac{1}{2}$$

$\therefore (\text{FP rate}, \text{recall}) = (0, \frac{1}{2})$





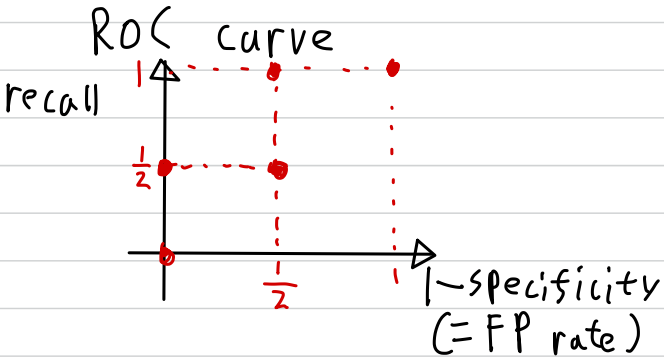
$f(x) > 5,\ S = \frac{1}{1+e^{-(2)}} = 0.88080$

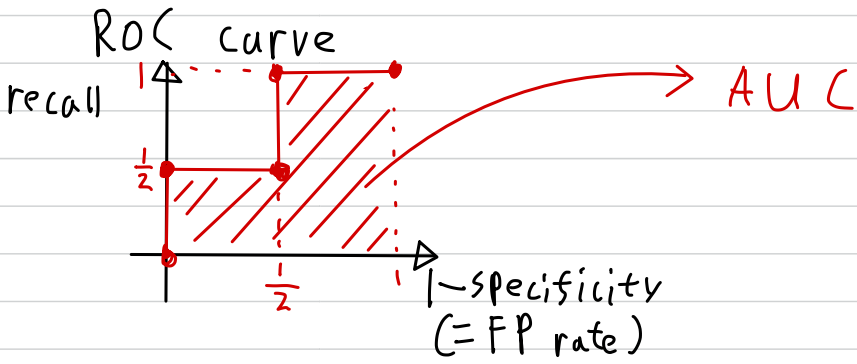
	Predicted as Positive	Predicted as Negative
Ground Truth Positive	0 = TP	2 = FN
Grounf Truth Negative	0 = FP	2 = TN

$$\text{recall} = \frac{TP}{TP+FN} = \frac{0}{2+0} = 0$$

$$\text{FP rate} = \frac{FP}{TN+FP} = \frac{0}{0+2} = 0$$

$\therefore$ (FP rate, recall) = (0,0)





$AUC = \frac{1}{2} \times \frac{1}{2} + 1 \times \frac{1}{2}$
$= 0.75$

6)

$$\text{softmax}(x) = \frac{\exp(x_i)}{\sum_{j=1}^{n} \exp(x_j)} \quad (i = 1, 2, \cdots, n)$$

$$s = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \quad \text{softmax}(s) = \begin{bmatrix} \frac{e^2}{e^2+e^1+e^0} \\ \frac{e^1}{e^2+e^1+e^0} \\ \frac{e^0}{e^2+e^1+e^0} \end{bmatrix} \approx \begin{bmatrix} 0.665241 \\ 0.244728 \\ 0.090031 \end{bmatrix}$$

7)

$$s = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

ground truth label for s is class 2

$$\text{softmax}(s) = \begin{bmatrix} \frac{e^2}{e^2+e^1+e^0} \\ \frac{e^1}{e^2+e^1+e^0} \\ \frac{e^0}{e^2+e^1+e^0} \end{bmatrix} \approx \begin{bmatrix} 0.665241 \\ 0.244728 \\ 0.090031 \end{bmatrix}$$

$L = -\log(0.244728)$

≈ 1.407608

8)

$y = \text{softmax}(W_2 \cdot \text{ReLU}(W_1 x))$

$$W_1 = \begin{bmatrix} -2 & 1 \\ 3 & -4 \end{bmatrix}, \quad W_2 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad x = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

$$W_1 \cdot x = \begin{bmatrix} -2 & 1 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \end{bmatrix}$$

$$\text{ReLU}(W_1 x) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (\because \text{ReLU}(x) = \max(0, x))$$

$$W_2 \cdot \text{ReLU}(W_1 x) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$y = \text{softmax}(W_2 \cdot \text{ReLU}(W_1 x)) = \begin{bmatrix} \frac{e^2}{e^2+e^1} \\ \frac{e^1}{e^2+e^1} \end{bmatrix} = \begin{bmatrix} 0.7311 \\ 0.2689 \end{bmatrix}$$

$(Y_1, Y_2) = (0.7311, 0.2689)$

9)

gradient descent를 사용하면,

$$(x, y) \rightarrow \left(x - \alpha \frac{\partial L(x,y)}{\partial x}, y - \alpha \frac{\partial L(x,y)}{\partial y}\right)$$

$$= \left(x - \alpha \frac{\partial(2x + 3xy)}{\partial x}, y - \alpha \frac{\partial(2x + 3xy)}{\partial y}\right)$$

$$= (x - \alpha(2 + 3y), y - \alpha(3x))$$

$x := x - 2\alpha - 3\alpha y$

$y := y - 3\alpha x$