Sreenidhi · Reddy Kalakonda
700771953
21331

Machine learning

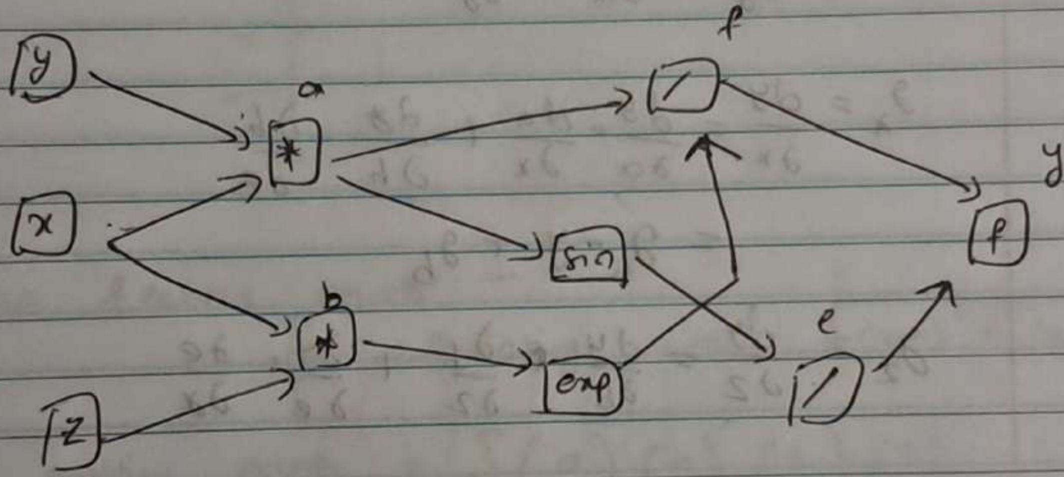
Assignment - 3

1) For the following function f, create the computation graph:

$$f(x,y,z) = \frac{\sin(xz)}{z} + \frac{e^{xz}}{xy}$$

Let,

$a = x \cdot y$
$b = x \cdot z$
$c = \sin(a)$
$d = e^b$
$f = d/a$
$e = c/z$
$y = f + e$



2) Based on computation graph and chain rule, show the derivative $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$

please note that derivative $\frac{\partial \sin(x)}{\partial x} = \cos(x)$, derivative $\frac{\partial(1/x)}{\partial(x)} = -1/x^2$,

the derivative of $\frac{\partial e^x}{\partial x} = e^x$

$$g_y = \frac{\partial f}{\partial y} = 1$$

$$g_e = g_f = 1$$

$$g_c = \frac{\partial y}{\partial c} = \frac{\partial y}{\partial c} \cdot \frac{\partial e}{\partial c} = g_e \cdot \tfrac{1}{2} = \tfrac{1}{2}$$

$$g_d = \frac{\partial y}{\partial d} = \frac{\partial y}{\partial f} \cdot \frac{\partial f}{\partial d} = g_f \cdot \tfrac{1}{a} = \tfrac{1}{a}$$

$$g_a = \frac{\partial y}{\partial a} = \frac{\partial y}{\partial c} \cdot \frac{\partial c}{\partial a} + \frac{\partial y}{\partial f} \cdot \frac{\partial f}{\partial a} = g_e$$

$$\cos a + g_f \cdot \frac{-d}{a^2}$$

$$g_a = \frac{\cos a}{2} - \frac{d}{a^2}$$

$$g_b = \frac{\partial y}{\partial b} = \frac{\partial y}{\partial d} \cdot \frac{\partial d}{\partial b} = g_b \cdot e^b = \frac{e^b}{a}$$

$$g_y = \frac{\partial y}{\partial y} = \frac{\partial y}{\partial a} \cdot \frac{\partial a}{\partial y} = g_a \cdot x$$

$$g_x = \frac{\partial y}{\partial x} = \frac{\partial y}{\partial a} \cdot \frac{\partial a}{\partial x} + \frac{\partial y}{\partial b} \cdot \frac{\partial b}{\partial x}$$

$$= g_a \cdot y + g_b$$

$$g_z = \frac{\partial y}{\partial z} = \frac{\partial y}{\partial b} \cdot \frac{\partial b}{\partial z} + \frac{\partial y}{\partial e} \cdot \frac{\partial e}{\partial x}$$

$$= g_b\, x - \frac{b}{z^2}$$

$$g_x = \frac{y \cos xy}{2} + \frac{e^{xz} \cdot z}{xy} - \frac{e^{xz}}{x^2 y}$$

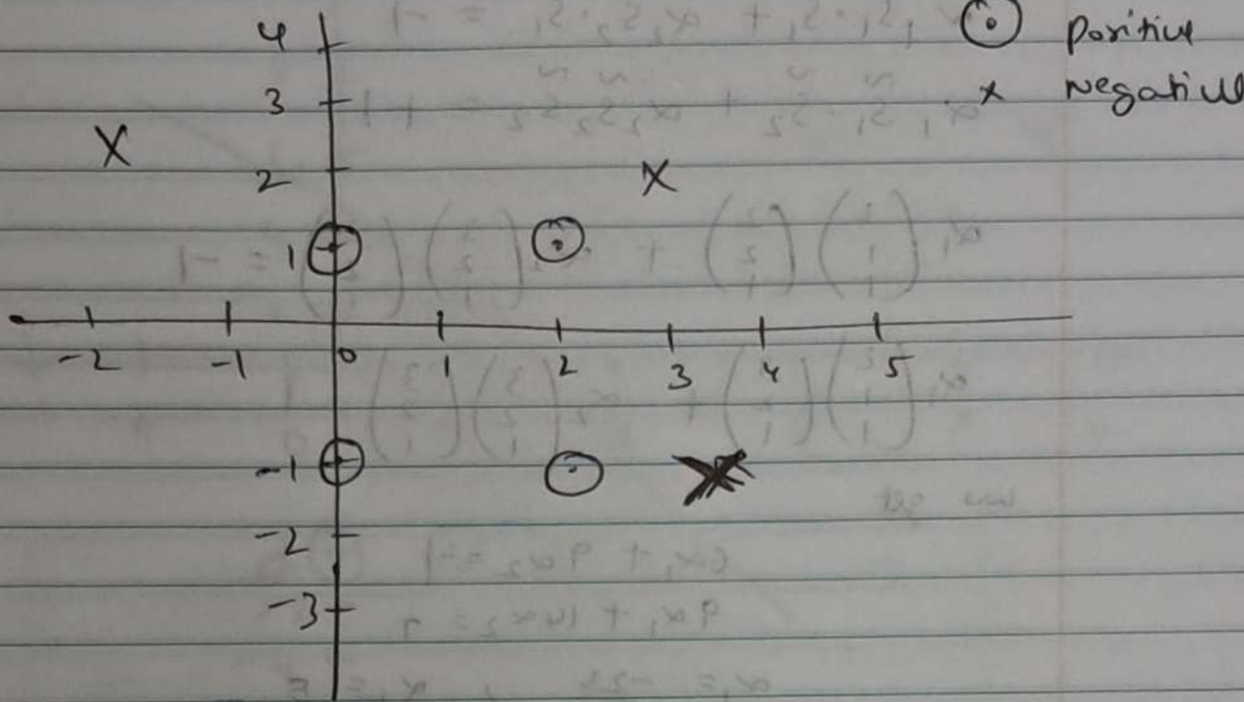
$$g_y = \frac{\cos xy}{2} \cdot x - \frac{e^{xz}}{y^2}$$

$$g_2 = \frac{-\sin xy}{z^2} + \frac{e^{xz}}{y}$$

SVM

Positive Points $(0,-1)$ $(0,1)$ $(2,-1)$ $(2,1)$

Negative point $(-1,2)$ $(-1,-2)$ $(3,2)$ $(3,-1)$



1) New feature points :—

after calculating with $\sqrt{(x_1-1)^2 + (x_2)^2} > 2$

positive points $= \left\{ \begin{pmatrix} 0 \\ -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right.$

negative points $= \left\{ \begin{pmatrix} 5 \\ 8 \end{pmatrix} \begin{pmatrix} 7 \\ 6 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} \begin{pmatrix} 9 \\ 5 \end{pmatrix} \right.$

2) Support vectors =

$$\left\{ s_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad s_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\}$$

each arg 1 with αl as bias i/p

$$\tilde{s}_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \quad \tilde{s}_2 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

$$\alpha_1 \tilde{s}_1 \cdot \tilde{s}_1 + \alpha_2 \tilde{s}_2 \cdot \tilde{s}_1 = -1$$

$$\alpha_1 \tilde{s}_1 \cdot \tilde{s}_2 + \alpha_2 \tilde{s}_2 \tilde{s}_2 = +1$$

$$\alpha_1 \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} = -1$$

$$\alpha_1 \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = 1$$

we get

$$6\alpha_1 + 9\alpha_2 = -1$$

$$9\alpha_1 + 14\alpha_2 = 7$$

$$\alpha_1 = \frac{-23}{3}, \quad \alpha_2 = 5$$

$$\bar{w} = \Sigma a_i \tilde{s}_i$$

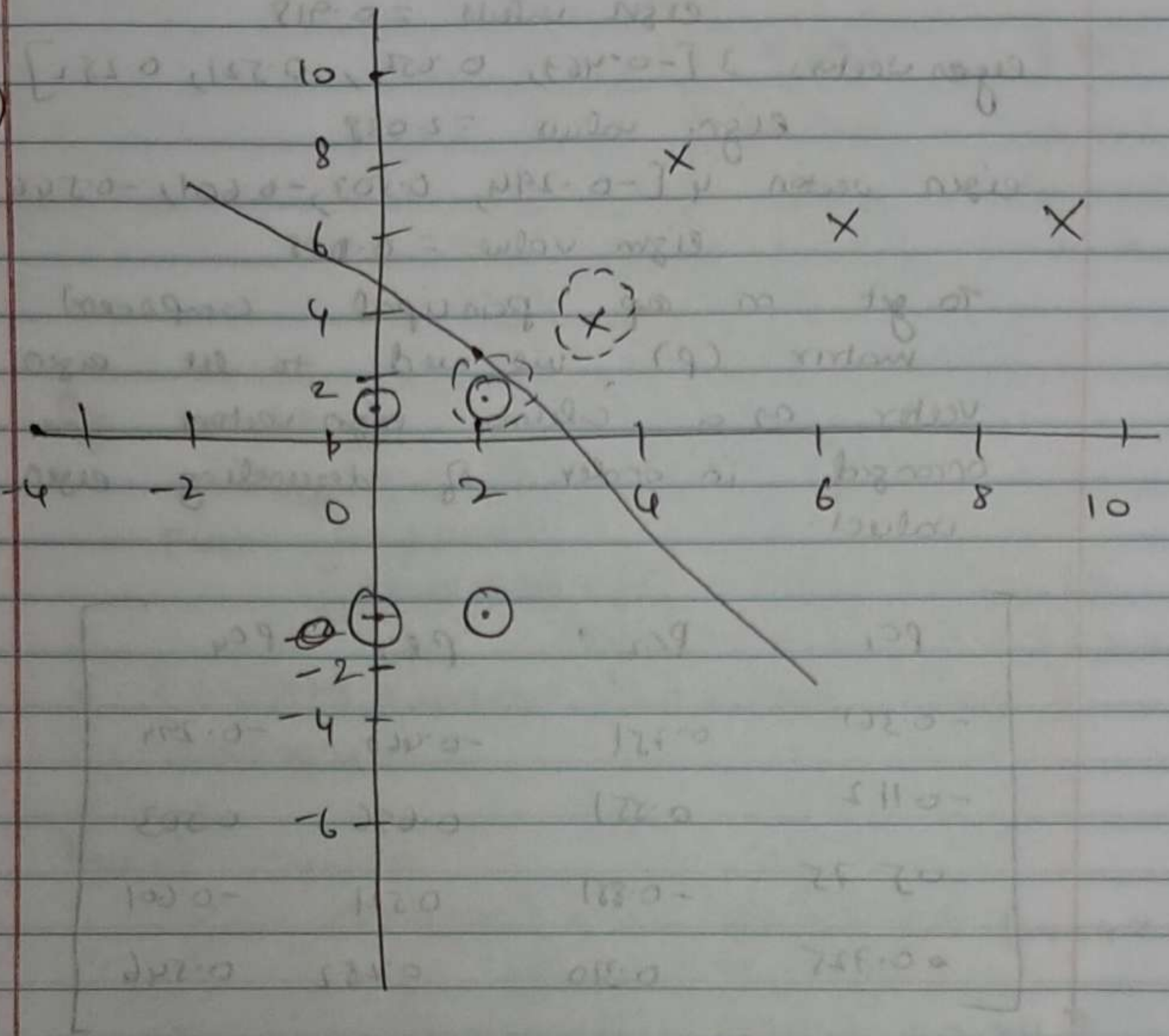
$$\frac{-23}{3} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + 5 \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} -46/3 \\ -23/3 \\ 3 \end{pmatrix} + \begin{pmatrix} 15 \\ 10 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} -0.33 \\ 2.33 \\ 8 \end{pmatrix}$$

3) $\boxed{\omega = \begin{pmatrix} -0.33 \\ 2.33 \end{pmatrix} \qquad b = 8.}$

4)



(*) PCA

Given eigen vectors & eigen values

1) eigen valectors 1 [-0.361, -0.112, -0.575, 0.725]

eigen value = 1.756

eigen vectors 2 [0.751, 0.551, -0.18, 0.310]

eigen values = 0.918

eigen vectors 3 [-0.467, 0.656, 0.521, 0.282]

eigen value = 2.038

eigen vectors 4 [-0.294, 0.503, -0.601, -0.546]

eigen value = 0.088

To get an ~~ap~~ principal component matrix (P) we need to let eigen vector as a column. eigen vectors are arranged in order of descending eigen values.

$$\begin{bmatrix} PC_1 & PC_2 & PC_3 & PC_4 \\ -0.361 & 0.751 & -0.467 & -0.294 \\ -0.112 & 0.551 & 0.656 & 0.503 \\ -0.575 & -0.881 & 0.521 & -0.601 \\ 0.725 & 0.310 & 0.282 & 0.546 \end{bmatrix}$$

Given,

$$\begin{bmatrix} -0.467 & 0.751 \\ 0.656 & 0.551 \\ 0.521 & -0.188 \\ 0.282 & 0.810 \end{bmatrix}$$

From data,

First sample

$$x_{\text{first sample}} = [7.49, 19.01, 14.64, 11.97]$$

Transformation formulae $x_{PCA} = xP$

$$x_{PCA} = [7.49, 19.01 \quad 14.64 \quad 11.97] \times \begin{bmatrix} -0.467 & 0.751 \\ 0.656 & 0.551 \\ 0.521 & -0.188 \\ 0.282 & 0.310 \end{bmatrix}$$

2)

By computing

$x_{PCA} =$

First component

$$[(7.49 \times -0.467) + (19.01 \times 0.656) + (14.64 \times 0.521) + (11.97 \times 0.2821)]$$

$$= [-3.49983 + 12.47056 + 7.62984 + 3.375]$$

$$= 19.967611$$

Second component

$$= (17.49 \times 0.751) + (14.01 \times 0.551) + (14.64 \times -0.188) + (11.97 * 0.310)]$$

$$= [5.62599 + 10.47451 - 2.75232 + 3.711]$$

$$= 17.05518.$$

PCA transformed, Result of First sampled using third component.

$$x_{PCA} \text{ first sample} = [19.96711 \quad 17.05618]$$

Clustering :-

	P_1	P_2	P_3	P_4	P_5	P_6
P_1	0					
P_2	0.23	0				
P_3	0.22	0.15	0			
P_4	0.37	0.20	0.15	0		
P_5	0.14	0.14	0.28	0.29	0	
P_6	0.23	0.25	0.11	0.22	0.39	0

Given to merge only P_3 & P_6
After merging both P_3 & P_6
we consider both as one entity
for further calculations

1) Single Link (min distance)

To calculate minimum distance b/w the new ~~di~~ cluster (P_3, P_6) and all other points

To distance b/w matrix

$MIN\ (dist\ (P_3, P_1), (P_6, P_1))$

$= min\ ((0.22, 0.23))$

$= 0.22$

To update distance matrix $MIN\ [dist\ (P_3, P_6), P_2]$

$MIN\ (dist\ (P_3, P_2), (P_6, P_2))$

$= min\ ((0.15, 0.25)) = 0.15$.

To update the distance matrix

$MIN\ [dist\ (P_3, P_6), P_4]$

$= min\ ((0.15, 0.22))$

$= 0.15$

To update the distance matrix

$MIN\ [dist\ (P_3, P_6); P_5]$

$= min\ ((0.28, 0.39)) = 0.28$

To update distance matrix for P_3, P_6

	P_1	P_2	P_3,P_6	P_4	P_5
P_1	0				
P_2	0.23	0			
P_3,P_6	0.22	0.15	0		
P_4	0.37	0.20	0.15	0	
P_5	0.34	0.14	0.28	0.24	0

2) Complete Link (max distance)

To update the distance matrix

$= MAX [dist ((P_3, P_6), P_1)]$

$= max [dist (P_3, P_1), (P_6, P_1)]$

$= MAX [(0.22, 0.23)] = 0.23$

To update the distance matrix

$MAX [dist (P_3, P_6), P_2]$

$MAX [(0.15, 0.25)] = 0.25$

To update the distance matrix

$MAX [dist ((P_3, P_6), P_4)]$

$= MAX [dist (P_3, P_4)]$

$= MAX [(0.15, 0.22)] = 0.22$

To update the distance matrix

$= MAX [dist ((P_3, P_6), P_5)]$

$= MAX [dist (P_3, P_5), (P_6, P_5)]$

$= MAX [(0.28, 0.39)] = 0.39$

To update distance matrix for cluster (P_3, P_6)

	P_1	P_2	P_3, P_6	P_4	P_5
P_1	0				
P_2	0.23	0			
P_3, P_6	0.23	0.25	0		
P_4	0.37	0.20	0.22	0	
P_5	0.34	0.19	0.39	0.29	0