سوال ۱۵ در این حالت داریم :

$$h(z') = g(\theta (x_1)^2 + \theta (x_2)^2 + \dots z + \alpha)$$

۱) در صورتی که $\theta = 0$ باشد ، یک خط را جدا خواهد کرد

۲) در صورتی که $\theta \neq 0$ باشد ، یک دایره را جدا خواهد کرد

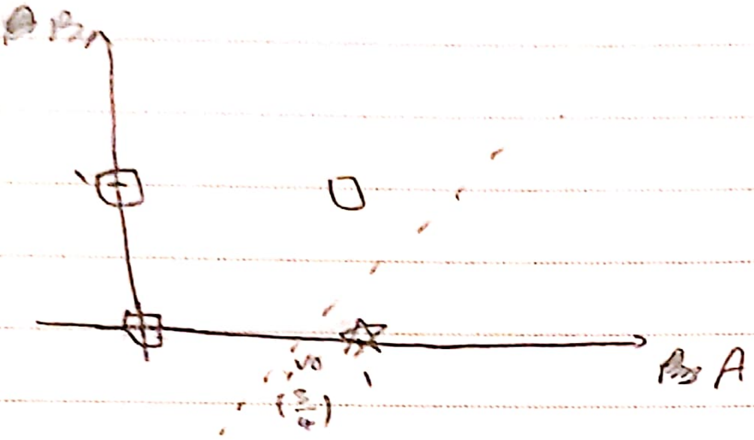
۳) از آن جایی که ضرایب $(x_1)^2$ و $(x_2)^2$ یکسان هستند بیضی نمی‌تواند باشد.

۴) از آن جایی که درجه‌ی معادله تنها تا ۲ درجه می‌باشد پس امکان به وجود آوردن منحنی مانند ک را نخواهد داشت.

سوال ۲

برچسب صفر ← □

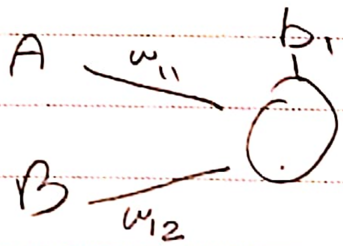
برچسب یک ← ☆



از شکل مشخص است که می‌توان یک خط آن را جدا کرد پس می‌توان با یک گره شبکه عصبی حل کرد

معادله خط $+35A - 40B - 10$

A	B	Z	a
0	0	-10	0
0	1	-50	0
1	0	25	1
1	1	-15	0



$b_1 = -10$

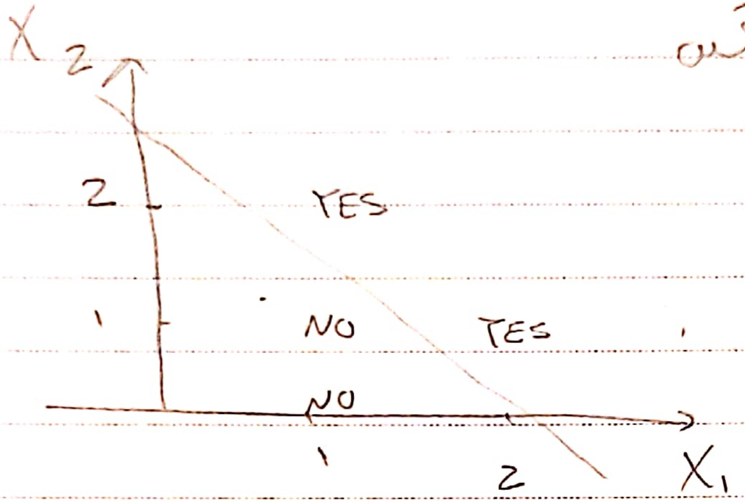
$w_{11} = 35$

$w_{12} = -40$

Subject:
Date

سوال ۸۳ ب با توجه به شکل

(الف)



برای مثال خط زیر می‌توانند جدا کننده =

x_1	x_2	z	a
۱	۰	-۲۰	۰
۱	۱	-۵	۰
۲	۱	۲۰	۱
۱	۲	۱۰	۱

۲۵x_1 + ۱۵x_2 - ۴۵ = ۰

$x_1 + x_2 + b < 0$

-۳۰ + ۱ + ۱۰

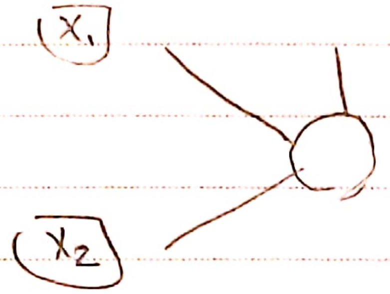
۲۵x_1 + ۱۵x_2 - ۴۵ = ۰

-۲۵x_1 + ۵x_2

PAPCO

Subject:
Date

سوال ۳ پ:



$J(\theta)_2$

تابع خطا $-(Y \log \hat{Y} + (1-Y)\log(1-\hat{Y}))$ cross entropy

$\hat{Y} = \frac{1}{1+e^{-z}} = \sigma(z)$ تابع Activation

$z = x w^T$

$X_A = (-۱, ۱, ۲)$

$w = (۰, -۱, ۱)$

$z = x w^T = (-۱, ۱, ۲)\begin{pmatrix} ۰ \\ -۱ \\ ۱ \end{pmatrix} = -۱ + ۲ = \boxed{1}$

$\hat{Y} = \sigma(z) = \sigma(1) = \frac{1}{1+\frac{1}{e}} = ۰,۷۳$

$J(\theta) = -\text{Ln}(۰,۷۳) = ۰,۳۱$

$w = w - \alpha \frac{dJ(\theta)}{dw}$

$\frac{dJ(\theta)}{dw} = \frac{dJ(\theta)}{d\hat{Y}} \times \frac{d\hat{Y}}{dz} \times \frac{dz}{dw}$

$= -\left(\frac{Y}{\hat{Y}} + (Y-1) \times \left(\frac{1}{1-\hat{Y}}\right)\right) \times \hat{Y}(1-\hat{Y}) \times X$

$(-Y(1-\hat{Y}) - (Y-1)\hat{Y})\, X =$

$-Y + Y\hat{Y} - Y\hat{Y} + \hat{Y} * X = (\hat{Y} - Y) * X$

PAPCO

Subject:
Date

$$\frac{dJ(\theta)}{dw} = (\hat{y} - y) * X$$

$X_A = (-1, 1, 2)$ $\quad y_A = 1$ $\quad \hat{y}_A = 0.73$

$$\frac{dJ(\theta)}{dw} = (0.73 - 1)(-1, 1, 2) = (+0.27, -0.27, -0.54)$$

-0.27

$$w = (0, -1, 1) - 0.1\,(+0.27, -0.27, -0.54)$$

$$= (-0.027, -0.973, 1.054)$$

برای نقطه B

$X_B = (-1, 2, 1)$ $\quad Z = X_B w^T =$

$$(-1, 2, 1) * \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = \boxed{-1}$$

$$\sigma(z) = \hat{y} = \sigma(-1) = \frac{1}{1+e} = 0.27$$

$$J(\theta) = -(y \log \hat{y} + (1-y)\log(1-\hat{y})) = -\log(0.27) = 1.30$$

$$\frac{dJ(\theta)}{dw} = (0.27 - 1)(-1, 2, 1) = (0.73, -1.46, -0.73)$$

$$w = (0, -1, 1) - 0.1\,(0.73, -1.46, -0.73)$$

$$= (-0.073, 0.854, 1.073)$$

PAPCO

Subject:
Date

سوال ۳ :

$\boxed{\text{نقطه } A}$

برای linear از تابع خطای MSE استفاده می‌کنیم

$$J = \frac{1}{2}(\hat{y} - y)^2$$

$$X_A = (-1, 1, 2) \qquad y = 1 \qquad w = (0, -1, 1)$$

$$\hat{y} = X w^T = (-1, 1, 2)\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = -1 + 2 = \boxed{1}$$

$$J = 0$$

$$\frac{dJ}{dw} = (\hat{y} - y) \times X = (1-1) \times X = 0$$

ضرایب تغییری نمی‌کنند.

نقطه B :

$$X = (-1, 2, 1) \qquad w = (0, -1, 1)$$

$$\hat{y} = X w^T = (-1, 2, 1)\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = -1$$

$$J = \frac{1}{2}(1 - (-1))^2 = \boxed{2}$$

$$\frac{dJ}{dw} = (-1 - 1) \times (X) = (2, -4, -2)$$

$$w = w - \alpha \frac{dJ}{dw} = (0, -1, 1) - 0.1\,(2, -4, -2)$$

$$= \boxed{(-0.2, -0.6, 1.2)}$$

Subject:
Date

سوال ۴

$$z_1 = W_1 \, x^{(i)} + b_1$$
$(D_{a_1}\times 1) \quad (D_x \times 1)$

الف: ابعاد $\boxed{W_1}$ برابر $\boxed{D_{a_1} \times D_x}$ می‌باشد

ابعاد $\boxed{b_1}$ برابر تعداد گره‌های لایه پنهان می‌باشد پس $\boxed{D_{a_1} * 1}$

← این هم از فرمول بالا راحت به جمع (و خروجی) است می‌آید.

ابعاد W_2 برابر ?

ابعاد a برابر $D_{a_1}\times 1$ است (ابعاد a برابر a که آن برابر
ابعاد z_1 است (خروجی $Relu(z_1)$ برابر یک دیگر با ابعاد z_1 است.))

$$\hat{y}^{(i)} = \sigma(W_2 a + b_2) \Rightarrow 1\times 1$$ یک عدد

$$a \rightarrow D_{a_1}\times 1 \Rightarrow \boxed{W_2 \rightarrow 1\times D_{a_1}}$$

پس $\boxed{b_2 \text{ نیز } 1\times1}$ است چون می‌خواهیم جمع کنیم با

$W_2 a$ که 1×1 است و باید خروجی‌های (1×1) شود.

X که ورودی ما می‌باشد دارای D_x ویژگی است و تعدادش نیز برابر m است

$$D_x \left\{ \underbrace{\begin{pmatrix} \vdots & & \vdots \\ x_{11} & \cdots & x_{m1} \\ \vdots & & \vdots \end{pmatrix}}_{m} \right.$$

ابعاد X برابر $D_x * m$

Y نیز خروجی دیگر عدد است پس ابعاد $1\times m$ خواهد داشت.

PAPCO

ب 8

$$\frac{\partial J}{\partial z_3} = \frac{\partial J}{\partial L} \times \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial z_3}$$

$$\frac{\partial J}{\partial L} = -\frac{1}{m}$$

$$\frac{\partial L}{\partial \hat{y}} = \frac{y}{\hat{y}} + (1-y)\frac{-1}{1-\hat{y}} = \frac{y - y\hat{y} - \hat{y} + y\hat{y}}{(1-\hat{y})\hat{y}}$$

$$= \frac{y-\hat{y}}{(1-\hat{y})\hat{y}}$$

$$\frac{\partial \hat{y}}{\partial z_3} = \sigma(z_3) \times (1-\sigma(z_3))$$

$$= \frac{\hat{y}-y}{\hat{y}(1-\hat{y})} \times (\hat{y}) \times (1-\hat{y}) = \boxed{\frac{1}{m}(\hat{y}-y)}$$

$\cancel{\frac{\partial J}{\partial}}$ $$\frac{\partial a}{\partial z_2} = \cancel{\frac{\partial a}{\partial z_2}} + \frac{\partial a}{\partial a_2} \times \frac{\partial a_2}{\partial z_1}$$

0

صفر

$$\frac{\partial a_2}{\partial z_2} = \begin{cases} 1 & z_2 > 0 \\ 0 & z_2 < 0 \end{cases}$$

در $z_2 = 0$ تعریف نشده است.

$$= \begin{cases} -1 & z_2 > 0 \\ 0 & z_2 < 0 \end{cases}$$

$$\frac{\partial J}{\partial W_1} = \frac{\partial J}{\partial z_3} \times \frac{\partial z_3}{\partial a} \times \left(\frac{\partial a}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial w_1} + \frac{\partial a}{\partial a_2} \times \frac{\partial a_2}{\partial z_2} \times \frac{\partial z_2}{\partial w_2} \right)$$

$$\frac{\hat{y} - y}{m} \times W_r \times \left(\left(1 \times \underbrace{(0,1)}_{\begin{cases} 1 & z_1 > 0 \\ 0 & z_1 < 0 \end{cases}} \times X \right) + \left((-1) \times \underbrace{(0,1)}_{\begin{cases} 1 & z_2 > 0 \\ 0 & z_2 < 0 \end{cases}} \times X \times x' \right) \right)$$

$$W_1 = w_1 - \alpha \frac{\partial J}{\partial w_1}$$

$$= w_1 - \alpha \left(\frac{\hat{y} - y}{m} \times w_r \times \left(\left(\underbrace{0,1\ X}_{\begin{cases} x & z_1 > 0 \\ 0 & z_1 < 0 \end{cases}} \right) + \begin{cases} -x' & z_r > 0 \\ 0 & z_r < 0 \end{cases} \right) \right)$$

$$B_1 = B_1 - \alpha \frac{\partial J}{\partial B_1} \qquad \frac{\partial J}{\partial B_1} = \frac{\partial J}{\partial a} \left(\frac{\partial a}{\partial z_1} \times \frac{\partial z_1}{\partial B_1} + \frac{\partial a}{\partial z_r} \times \frac{\partial z_r}{\partial B_2} \right)$$

$$= B_1 - \alpha \left(\frac{\hat{y} - y}{m} \times w \times \left(\begin{cases} 1 & z_1 > 0 \\ 0 & z_1 < 0 \end{cases} + \begin{cases} -1 & z_r > 0 \\ 0 & z_r < 0 \end{cases} \right) \right)$$

$$w_2 = w_2 - \alpha \frac{\partial J}{\partial w_2} \qquad \frac{\partial J}{\partial w_2} = \frac{\partial J}{\partial z_3} \times \frac{\partial z}{\partial w_r}$$

$$= w_r - \alpha \left(\frac{\hat{y} - y}{m} \times a \right)$$

Subject:
Date

$$B_2 = B_2 - \alpha \frac{\partial J}{\partial B_2}$$

$$\frac{\partial J}{\partial B_2} = \frac{\partial J}{\partial z_3} \times \frac{\partial z_3}{\partial b_2} = \frac{\partial J}{\partial z_3} = \frac{\hat{y} - y}{m}$$

$$B_2 = B_2 - \alpha \frac{\hat{y} - y}{m}$$