# تمرین سری 5 درس ML - رادین خیام - 99101579

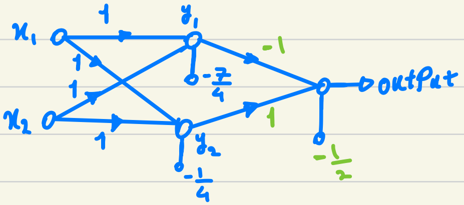
## سوال 1-

الف) باید حداکثر 3 رأس مربع استفاده شده باشد که بتوان به ازای هر حالتی با یک نورون خروجی جدا بشوند.

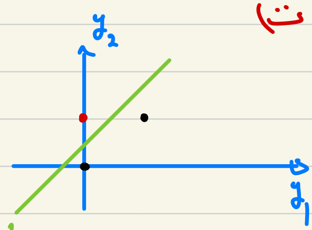
ب) باید حداکثر 3 رأس مکعب استفاده شده باشد چون اگر مثلاً 4 رأس استفاده شده باشند مثل اون حالت XOR می‌شود که هر 4 تا روی یک وجه مکعب باشند و خب دیگه با یک صفحه نمی‌شود جدا کردشون.

پ) بستگی به داده‌های داده شده دارد اگر خطی جدایی پذیر باشند می‌شود اگر نباشند خیر، مثلاً داده‌های XOR را نمی‌توان.

ت)




| $x_1$ | $x_2$ | $y_1$ | $y_2$ | output |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 1 | 1 |
| 1 | 0 | 0 | 1 | 1 |
| 1 | 1 | 1 | 1 | 0 |



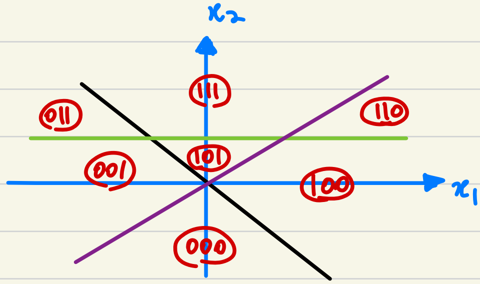

$L: y_2 - y_1 - \frac{1}{2} = 0$

سؤال 2 -

$x_1+x_2=0$

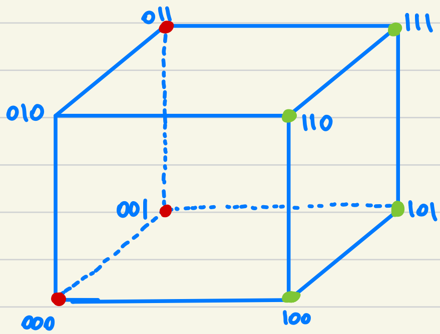
$x_2=\frac{1}{4}$

$x_1-x_2=0$



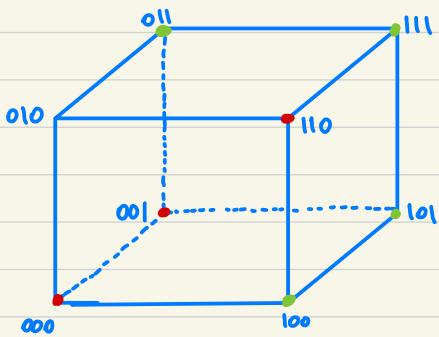

* برای اینکه دولایه کافی باشد :

(با یک صفحه باید بتوان جدا کرد)



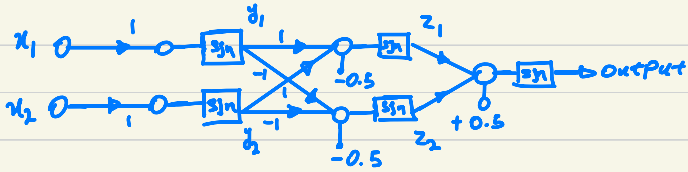

* برای اینکه سه لایه نیاز باشد :

(با یک صفحه جدا نشوند)



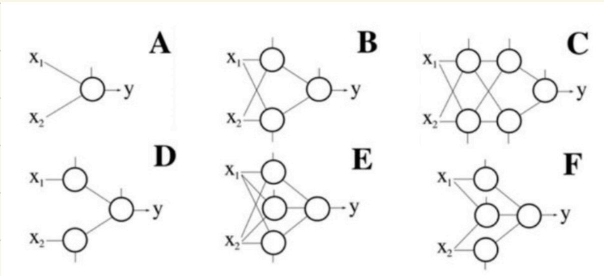
## سؤال ۳ -

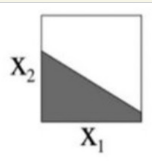
* با دو لایه نمی‌شود پس با سه لایه امتحان می‌کنیم:



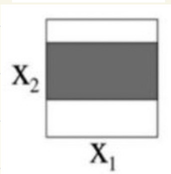

| | $x_1$ | $x_2$ | $y_1$ | $y_2$ | $z_1$ | $z_2$ | output |
|---|---|---|---|---|---|---|---|
| ربع اول | + | + | +1 | +1 | +1 | -1 | +1 |
| ربع چهارم | + | − | +1 | -1 | -1 | -1 | -1 |
| ربع دوم | − | + | -1 | +1 | -1 | -1 | -1 |
| ربع سوم | − | − | -1 | -1 | -1 | +1 | +1 |

## سؤال 4 -

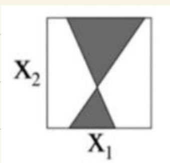


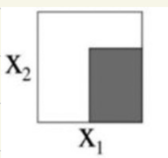

→ A → چون یک نرون تنها یک معادله خط می‌سازد برای ما.



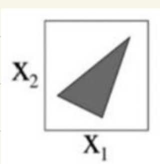

→ B → دو خط داریم که با دو نرون در لایه پنهان امکان پذیر است.



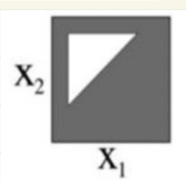

→ C → در لایه پنهان اول دو خط کشیده می‌شود و فضا به ۴ ناحیه تقسیم می‌شود با دو نرون لایه پنهان دوم هم دو تا از این چهار تا ناحیه انتخاب می‌شوند.




→ D → چون خط‌ها یکی شون باید فقط تابع $x_1$ و دیگری فقط تابع $x_2$ باشد.

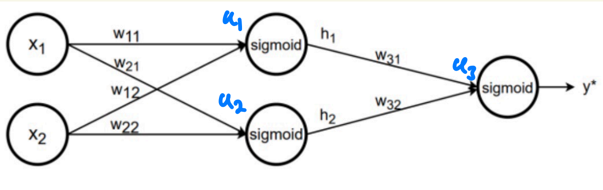



→ E → چون سه تا خط کشیدیم با سه نرون داشته باشیم با F هم نمی‌شود.




→ F → چون دو تا از خط‌ها یکی باید فقط تابع $x_1$ باشد دیگری هم فقط تابع $x_2$.

سؤال 5 -




$$H = -\left(y \log y^* + (1-y)\log(1-y^*)\right)$$

$$f(x) = \text{sigmoid} = \frac{1}{1+e^{-ax}} \longrightarrow f'(x) = a f(x)(1-f(x))$$

$$u_1 = w_{11}x_1 + w_{12}x_2$$

$$u_2 = w_{21}x_1 + w_{22}x_2$$

$$h_1 = f(u_1)$$

$$h_2 = f(u_2)$$

$$u_3 = h_1 w_{31} + h_2 w_{32} = f(u_1)w_{31} + f(u_2)w_{32} = f(w_{11}x_1 + w_{12}x_2)w_{31} + f(w_{21}x_1 + w_{22}x_2)w_{32}$$

$$y^* = f(u_3)$$

$$\frac{\partial H}{\partial w_{12}} = \frac{\partial H}{\partial u_3}\frac{\partial u_3}{\partial w_{12}} = \frac{\partial H}{\partial u_3}\left(\frac{\partial u_3}{\partial h_1}\underbrace{\frac{\partial h_1}{\partial w_{12}}}_{\frac{\partial h_1}{\partial u_1}\frac{\partial u_1}{\partial w_{12}}} + \frac{\partial u_3}{\partial h_2}\overset{0}{\cancel{\frac{\partial h_2}{\partial w_{12}}}}\right) = \frac{\partial H}{\partial u_3}\frac{\partial u_3}{\partial h_1}\frac{\partial h_1}{\partial u_1}\frac{\partial u_1}{\partial w_{12}}$$

$$\frac{\partial u_1}{\partial w_{12}} = x_2 \ , \ \frac{\partial h_1}{\partial u_1} = f'(u_1) \ , \ \frac{\partial u_3}{\partial h_1} = w_{31}$$

$$\frac{\partial H}{\partial u_3} = \frac{\partial H}{\partial y^*}\frac{\partial y^*}{\partial u_3} = \left(\frac{-y}{y^*} + \frac{1-y}{1-y^*}\right)f'(u_3) = a f(u_3)(1-f(u_3))\left(\frac{-y}{f(u_3)} + \frac{1-y}{1-f(u_3)}\right)$$

$$= a\left(y(f(u_3)-1) + (1-y)f(u_3)\right) = a\left(f(u_3) - y\right)$$

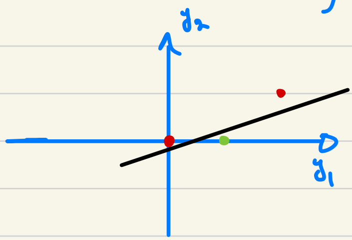
$$\longrightarrow \frac{\partial H}{\partial w_{12}} = a(f(u_3)-y)\, w_{31}\, a f(u_1)(1-f(u_1))\, x_2$$

$$\longrightarrow \boxed{\frac{\partial H}{\partial w_{12}} = a^2(y^*-y)\, w_{31}\, x_2\, h_1(1-h_1)}$$

سؤال 6 -

الف)

$$y=\begin{bmatrix} x_1+x_2 \\ x_1x_2 \end{bmatrix} \to \left.\begin{aligned} X=\begin{bmatrix}0\\0\end{bmatrix} &\to y=\begin{bmatrix}0\\0\end{bmatrix} \\ X=\begin{bmatrix}0\\1\end{bmatrix} &\to y=\begin{bmatrix}1\\0\end{bmatrix} \\ X=\begin{bmatrix}1\\0\end{bmatrix} &\to y=\begin{bmatrix}1\\0\end{bmatrix} \\ X=\begin{bmatrix}1\\1\end{bmatrix} &\to y=\begin{bmatrix}2\\1\end{bmatrix} \end{aligned}\right\} \to$$ جدای پذیر است




$$g(y)=y_2-\frac{1}{2}y_1+\frac{1}{3}=0$$

$$y=\begin{bmatrix} x_1^2 \\ x_2^2 \end{bmatrix} \to \left.\begin{aligned} X=\begin{bmatrix}0\\0\end{bmatrix} &\to y=\begin{bmatrix}0\\0\end{bmatrix} \\ X=\begin{bmatrix}0\\1\end{bmatrix} &\to y=\begin{bmatrix}0\\1\end{bmatrix} \\ X=\begin{bmatrix}1\\0\end{bmatrix} &\to y=\begin{bmatrix}1\\0\end{bmatrix} \\ X=\begin{bmatrix}1\\1\end{bmatrix} &\to y=\begin{bmatrix}1\\1\end{bmatrix} \end{aligned}\right\} \to$$ جدای پذیر نیست

$$y=\begin{bmatrix} x_1^2+x_2^2 \\ x_2 \end{bmatrix} \to \left.\begin{aligned} X=\begin{bmatrix}0\\0\end{bmatrix} &\to y=\begin{bmatrix}0\\0\end{bmatrix} \\ X=\begin{bmatrix}0\\1\end{bmatrix} &\to y=\begin{bmatrix}1\\1\end{bmatrix} \\ X=\begin{bmatrix}1\\0\end{bmatrix} &\to y=\begin{bmatrix}1\\0\end{bmatrix} \\ X=\begin{bmatrix}1\\1\end{bmatrix} &\to y=\begin{bmatrix}2\\1\end{bmatrix} \end{aligned}\right\} \to$$ جدای پذیر نیست

ب)

$$y_1 = f_1(\underline{x}) = |x_1 x_2 + x_1 - x_2|$$

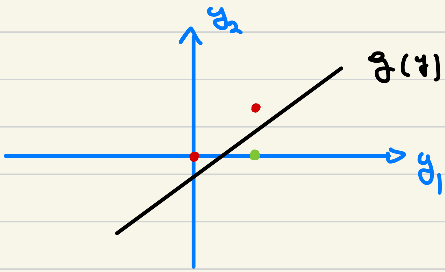
$$y_2 = f_2(\underline{x}) = x_1 x_2$$

$$X = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \to y = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$X = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \to y = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \to y = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

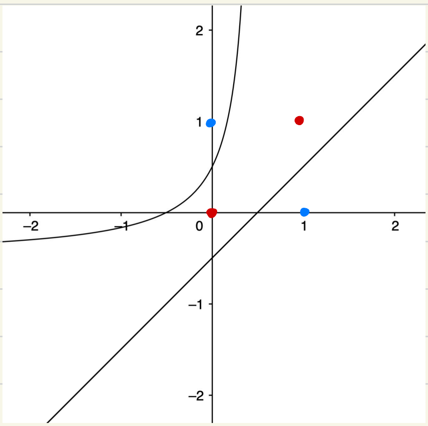
$$X = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \to y = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$




$$g(y) = y_2 - y_1 + \frac{1}{2} = 0$$

$$g(x) = x_1 x_2 - |x_1 x_2 + x_1 - x_2| + \frac{1}{2} = 0$$

$g(x) \to$




Made with Goodnotes

سؤال 7 -

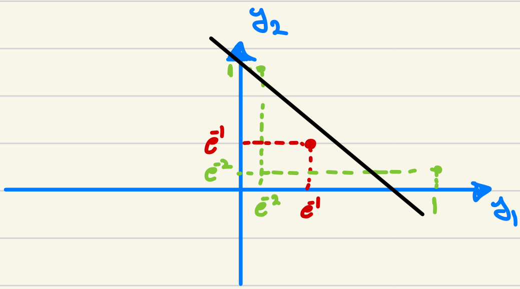
$$c_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad c_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \longrightarrow y = \begin{bmatrix} e^{-\|x - c_1\|_2^2} \\ e^{-\|x - c_2\|_2^2} \end{bmatrix}$$

$$X = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \longrightarrow y = \begin{bmatrix} e^{-1} \\ e^{-1} \end{bmatrix}$$

$$X = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \longrightarrow y = \begin{bmatrix} 1 \\ e^{-2} \end{bmatrix}$$

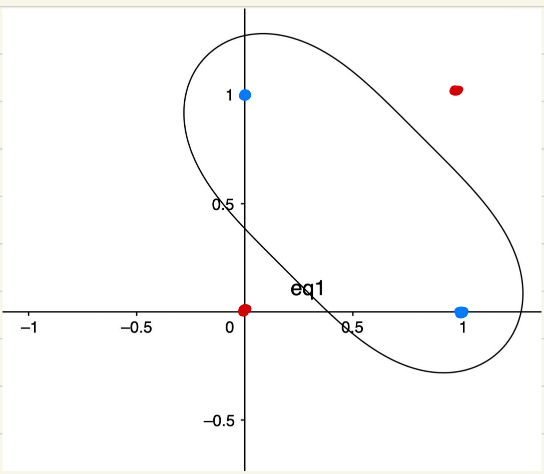
$$X = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \longrightarrow y = \begin{bmatrix} e^{-2} \\ 1 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \longrightarrow y = \begin{bmatrix} e^{-1} \\ e^{-1} \end{bmatrix}$$




$$g(y) = y_2 + y_1 - 1 = 0$$

$$\longrightarrow g(x) = e^{-(\|x - c_1\|_2^2)} + e^{-(\|x - c_2\|_2^2)} - 1 = 0$$



سؤال 8 -

الف)

$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad , \quad \underline{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$

$K(\underline{x}, \underline{z}) = (\underline{x}^T \underline{z})^3 = (x_1 z_1 + x_2 z_2)^3$

$= x_1^3 z_1^3 + x_2^3 z_2^3 + 3 x_1^2 z_1^2 x_2 z_2 + 3 x_1 z_1 x_2^2 z_2^2 = \langle \varphi(\underline{x}), \varphi(\underline{z}) \rangle$

$\longrightarrow \varphi(\underline{x}) = \begin{bmatrix} x_1^3 \\ x_2^3 \\ \sqrt{3}\, x_1^2 x_2 \\ \sqrt{3}\, x_1 x_2^2 \end{bmatrix}$

$x_1 - z_1$

$x_2 - z_2$

$-2x_1 z_1 - 2x_2 z_2$

ب)

$\exp\left(-\frac{1}{2} \|\underline{x} - \underline{z}\|_2^2\right) = \exp\left(-\frac{1}{2}\left(\|\underline{x}\|_2^2 + \|\underline{z}\|_2^2 - 2\underline{x}^T \underline{z}\right)\right)$

$= \exp(x^T z) \exp\left(-\frac{1}{2}\|\underline{x}\|_2^2\right) \exp\left(-\frac{1}{2}\|\underline{z}\|_2^2\right)$

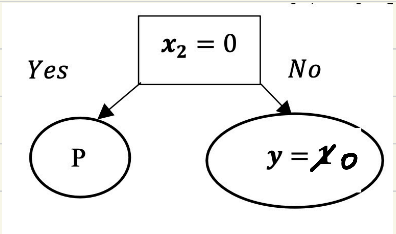
$\xrightarrow{\text{تیلور}} = \sum_{j=0}^{\infty} \frac{(x^T z)^j}{j!} \exp\left(-\frac{1}{2}\|x\|_2^2\right) \exp\left(-\frac{1}{2}\|\underline{z}\|_2^2\right)$

$= \sum_{j=0}^{\infty} \sum_{\sum_{k=0}^{K} n_k = j} \exp\left(-\frac{1}{2}\|\underline{x}\|_2^2\right) \frac{x_1^{n_1} \cdots\cdots x_K^{n_K}}{\sqrt{n_1! \cdots n_K!}} \exp\left(-\frac{1}{2}\|\underline{z}\|_2^2\right) \frac{z_1^{n_1} \cdots\cdots z_K^{n_K}}{\sqrt{n_1! \cdots n_K!}}$

$\longrightarrow \varphi(x) = \exp\left(-\frac{1}{2}\|\underline{x}\|_2^2\right) \begin{bmatrix} a_{l_0}^{(0)} \\ a_1^{(1)} \\ \vdots \\ a_{l_1}^{(1)} \\ \vdots \\ a_1^{(j)} \\ \vdots \\ a_{l_j}^{(j)} \end{bmatrix} \quad , \quad l_j = \binom{K+j-1}{j} \quad a_l^{(j)} = \frac{x_1^{n_1} \cdots x_K^{n_K}}{\sqrt{n_1! \cdots n_K!}}$

$n_1 + n_2 + \cdots + n_K = j \quad . \quad 1 \leq l \leq l_j$

سؤال و-




| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $y$ |
|---|---|---|---|---|
| 0 | 1 | 0 | 1 | **0** |
| 1 | 0 | 1 | 0 | **1** |
| 1 | 0 | 0 | 0 | **0** |
| 0 | 0 | 1 | 0 | **1** |
| 0 | 1 | 1 | 0 | **0** |

تو حالت $x_2=0$ خواهیم داشت:

| $x_1$ | $x_3$ | $x_4$ | $y$ |
|---|---|---|---|
| 1 | 1 | 0 | 1 |
| 1 | 0 | 0 | 0 |
| 0 | 1 | 0 | 1 |

۱) یک برگ با برچسب ۰ ✗ ← کلاً برگ نداریم چون y مشخص نیست.

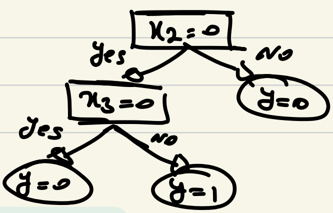
۲) یک برگ با برچسب ۱ ✗ ←

۳) یک گره تصمیم‌گیری بر اساس $x_1$ ✗ ← با تغییر $x_1$، y عوض نمی‌شود.

۴) یک گره تصمیم گیری بر اساس $x_2$ ✗ ← معنایی نمی‌دهد چون $x_2=0$ است

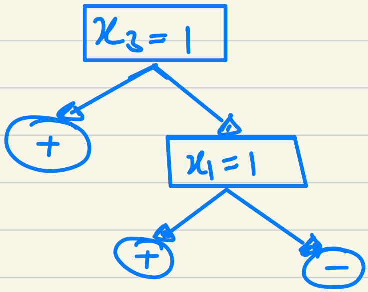
۵) یک گره تصمیم گیری بر اساس $x_3$ ✓ ← این می‌شود چون با تغییر $x_3$, y هم عوض می‌شود

۶) یک گره تصمیم گیری بر اساس $x_4$ ✗ ← $x_4$ کلا صفر است از روی اون نمی‌شود تشخیص داد



سوال 10-

بهترین ویژگی $x_3$ هست چون آنتروپی را به بیشترین میزان کاهش میدهد، سپس وضعیت $x_1$ را بررسی نماییم:




سوال 11-

الف) ویژگی های فضای جدید ←

$$r=\sqrt{x^2+y^2}$$

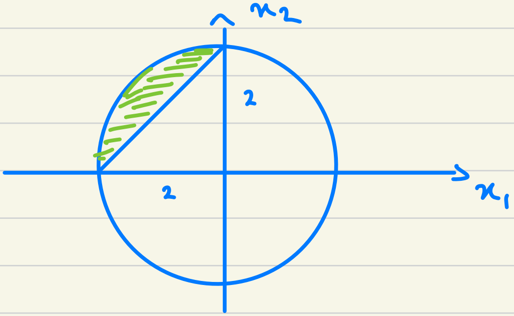
$$\theta = \arctan\left(\frac{y}{x}\right)$$

در فضای جدید داده ها با استفاده از خط قابل جدا سازی هستند.

ب) برای کرنل های SVM می توانیم از کرنل RBF استفاده کنیم چون داده ها بر اساس فاصلشون تا مرکز کلاس بندی شده اند.

سوال 12-

الف)




$\varphi(\underline{x})$

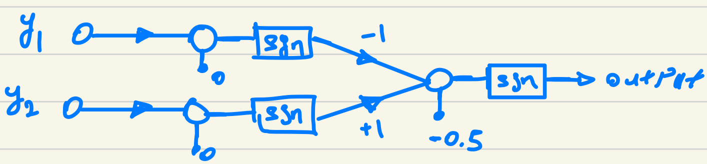
$$\underline{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1^2 + x_2^2 - 4 \\ x_2 - x_1 - 2 \end{bmatrix}$$

$x_2^2 + x_1^2 - 4 < 0$

$x_2 - x_1 - 2 > 0$

هر تمام نقاط داخل مساحت هاشور خورده تحت این نگاشت به داخل ربع دوم مختصات می روند بنابراین با یک طبقه بند خطی می توانیم آن را جدا کنیم.

ب)




| $y_1$ | $y_2$ | output |
|---|---|---|
| + | + | 0 |
| + | − | 0 |
| − | + | 1 |
| − | − | 0 |

Made with Goodnotes