

## سوال ۱

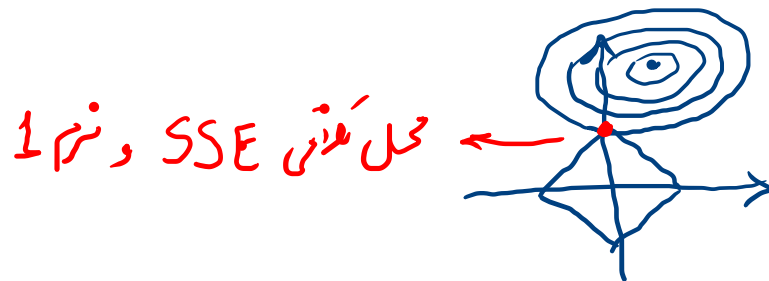
الف) درست، در مدل های غیر پارامتریک برخلاف مدل های پارامتریک فرض محدود کننده ای برای نحوه تولید داده ها در نظر گرفته نمی شود.

ب) درست،  
$$\Phi(x) = \begin{bmatrix} x_1^2 & x_2 \\ x_3 \end{bmatrix} \Rightarrow y = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}^T \Phi(x)$$

ج) درست، زیرا اثر یک ترین نمونه به هر نمونه آماری، خودش می شود.

د) غلط، زیرا مدل داده شده نمی توان به صورت خطی نسبت به  $\alpha_1$ ،  $\alpha_2$  نوشت.

ه) درست،



و) غلط، زیرا می توان توزیع  $class_{conditional}$  را با استفاده از روش بازین کمین از سپس از صبه بند  
بیراسته نمود.

سوال ٢

$$\text{الف) } p(y|x) = \mathcal{N}(y | wx, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-wx)^2}$$

$$\log \text{likelihood} = \log P(D|w) = \sum_{i=1}^n \log p(y_i | x_i, w) = \sum_{i=1}^n -\log \sqrt{2\pi} - \frac{1}{2}(y_i - wx_i)^2$$

$$= \text{const} - \frac{1}{2} \sum_{i=1}^n (y_i - wx_i)^2$$

SSE

$$\Rightarrow \arg \max_w \log P(D|w) = \arg \min_w \text{SSE}$$

ب)

$$\log P(w|D) = \log \frac{P(D|w) P(w)}{P(D)} = \log P(D|w) + \log P(w) - \underbrace{\log P(D)}_{\text{const w.r.t. } w}$$

$$= \underbrace{n \log \frac{1}{\sqrt{2\pi}} - \frac{1}{2} \sum_{i=1}^n (y_i - wx_i)^2 + \log \frac{1}{\sqrt{2\pi}\sigma} - \frac{1}{2\sigma^2} w^2}_{\text{طبق نسبت الف}} - \underbrace{\log P(D)}_{\text{const}}$$

$$\Rightarrow \arg \max_w \log P(w|D) = \arg \min_w \sum_{i=1}^n (y_i - wx_i)^2 + \frac{1}{\sigma^2} w^2$$

$$\Rightarrow \boxed{\lambda = 1/\sigma^2}$$

$$P(y=0) = \frac{3}{7} \quad P(y=1) = \frac{4}{7}$$

سوال ۳

$$P(A=0 | y=0) = \frac{2}{3} \quad P(A=0 | y=1) = \frac{1}{4}$$

$$P(B=0 | y=0) = \frac{1}{3} \quad P(B=0 | y=1) = \frac{1}{2}$$

$$P(C=0 | y=0) = \frac{2}{3} \quad P(C=0 | y=1) = \frac{1}{2}$$

$$P(0,0,1 | y=0) = \frac{2}{3} \times \frac{1}{3} \times \frac{1}{3}$$

$$P(0,0,1 | y=1) = \frac{1}{4} \times \frac{1}{2} \times \frac{1}{2}$$

$$P(1,1,1 | y=0) = \frac{1}{3} \times \frac{2}{3} \times \frac{1}{3}$$

$$P(1,1,1 | y=1) = \frac{3}{4} \times \frac{1}{2} \times \frac{1}{2}$$

$$\frac{P(y=0 | 0,0,1)}{P(y=1 | 0,0,1)} = \frac{\frac{2}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{3}{7}}{\frac{1}{4} \times \frac{1}{2} \times \frac{1}{2} \times \frac{4}{7}} = \frac{\frac{2}{9}}{\frac{1}{4}} = \frac{8}{9} < 1 \Rightarrow y^* = 1$$

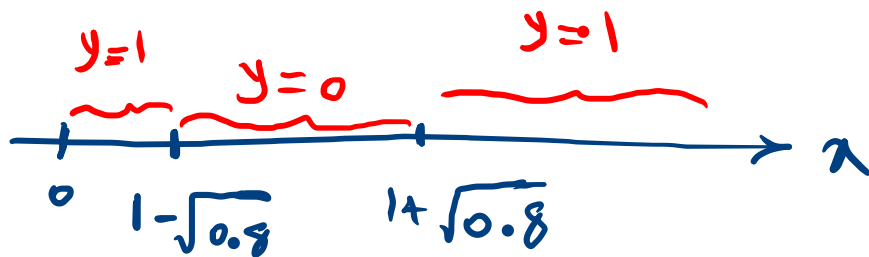
$$\frac{P(y=0 | 1,1,1)}{P(y=1 | 1,1,1)} = \frac{\frac{1}{3} \times \frac{2}{3} \times \frac{1}{3} \times \frac{3}{7}}{\frac{3}{4} \times \frac{1}{2} \times \frac{1}{2} \times \frac{4}{7}} = \frac{\frac{2}{9}}{\frac{3}{4}} = \frac{8}{27} < 1 \Rightarrow y^* = 1$$

سوال ۴

$$g(x) = \log \frac{p(x|y=0) \cancel{p(y=0)}}{p(x|y=1) \cancel{p(y=1)}} = 0$$

$$\Rightarrow \log \frac{2}{\sqrt{2\pi}} - \frac{x^2}{2} + x = 0 \Rightarrow x^2 - 2x + 0.2 = 0$$

$$x_1, x_2 = \frac{1 \pm \sqrt{1 - 0.2}}{1} = 1 \pm \sqrt{0.8}$$

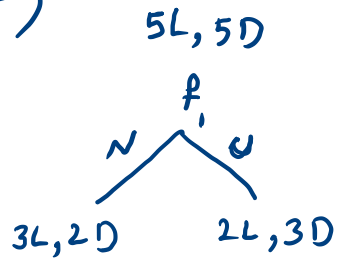


## سوال 5

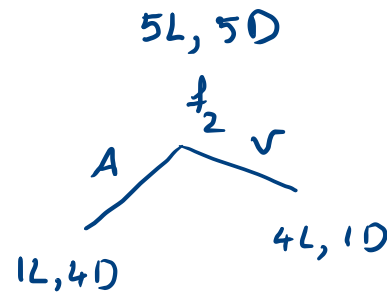
الف)  $H(f_2 | f_1 = N) = ?$

$$P(f_2 = A | f_1 = N) = \frac{1}{5} \Rightarrow H(f_2 | f_1 = N) = -\frac{1}{5} \log \frac{1}{5} - \frac{4}{5} \log \frac{4}{5}$$

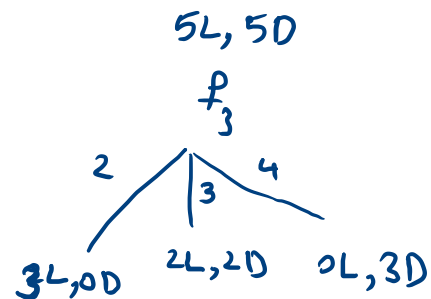
ب)



$$H(y | f_1) = \frac{1}{2} \left( -\frac{2}{5} \log \frac{2}{5} - \frac{3}{5} \log \frac{3}{5} \right) + \frac{1}{2} \left( -\frac{2}{5} \log \frac{2}{5} - \frac{3}{5} \log \frac{3}{5} \right) \cong 2.14$$



$$H(y | f_2) = -\frac{1}{5} \log \frac{1}{5} - \frac{4}{5} \log \frac{4}{5} \cong 0.7$$



$$H(y | f_3) = \frac{4}{10} \left( -\frac{1}{2} \log \frac{1}{2} - \frac{1}{2} \log \frac{1}{2} \right) + \frac{3}{10} \times 0 + \frac{3}{10} \times 0 = 0.4$$

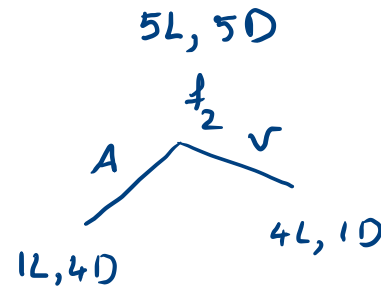
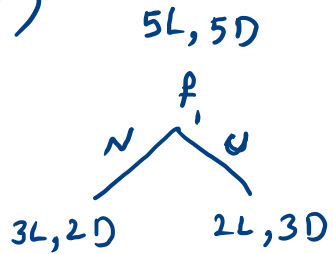
لذا  $f_3$  به عنوان رتبه اتی ب می شود.

## سوال 5

الف)  $H(f_2 | f_1 = N) = ?$

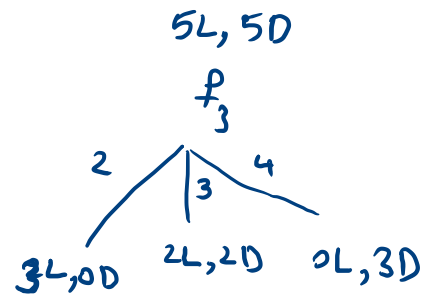
$$P(f_2 = A | f_1 = N) = \frac{2}{5} \Rightarrow H(f_2 | f_1 = N) = -\frac{2}{5} \log \frac{2}{5} - \frac{3}{5} \log \frac{3}{5}$$

ب)



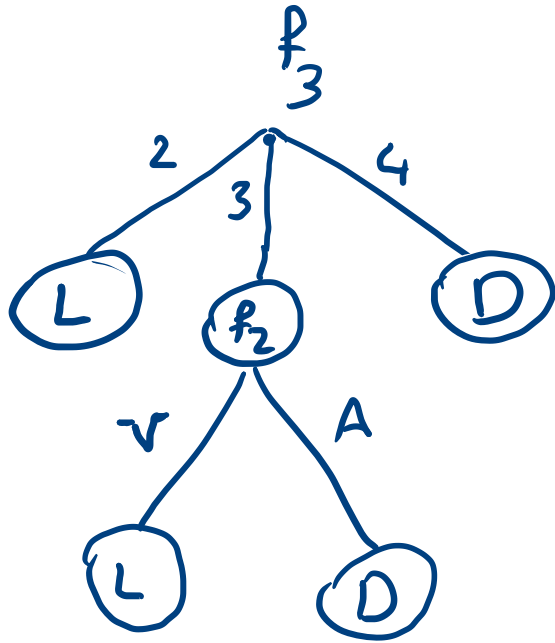
$$\begin{aligned} H(y | f_1) &= \frac{1}{2} \left( -\frac{2}{5} \log \frac{2}{5} - \frac{3}{5} \log \frac{3}{5} \right) \\ &+ \frac{1}{2} \left( -\frac{2}{5} \log \frac{2}{5} - \frac{3}{5} \log \frac{3}{5} \right) \\ &\cong 2.14 \end{aligned}$$

$$\begin{aligned} H(y | f_2) &= -\frac{1}{5} \log \frac{1}{5} - \frac{4}{5} \log \frac{4}{5} \\ &\cong 0.7 \end{aligned}$$



$$\begin{aligned} H(y | f_3) &= \frac{4}{10} \left( -\frac{1}{2} \log \frac{1}{2} - \frac{1}{2} \log \frac{1}{2} \right) \\ &+ \frac{3}{10} \times 0 + \frac{3}{10} \times 0 = 0.4 \end{aligned}$$

لذا  $f_3$  به عنوان ریشه انتخاب می‌شود.



(ج) در قسمت قبل ریشه دخت  $P_3$  بدتر آمد.  
 به لزی  $P_3 = 3$  واضح است که  $P_2$  بهترین  
 نماینده بین دو کلاس ایجاد می کند.

$$Acc = 100\%$$

$$Recall = 100\%$$

(د) با توجه به دخت داریم ،



الف) حرره نوردن  $\mathcal{L}$  بابر  $\mathcal{L}$  قرار ديم، در اين صورت داريم،

$$Y = c\omega_5 (c\omega_1 x_1 + c\omega_3 x_2) + c\omega_6 (c\omega_2 x_1 + c\omega_4 x_2)$$

$$= \underbrace{c^2(\omega_1 \omega_5 + \omega_6 \omega_2)}_{\beta_1} x_1 + \underbrace{c^2(\omega_3 \omega_5 + \omega_6 \omega_4)}_{\beta_2} x_2$$

ب) نورون ۱ و ۲ را با بیدار قرار دهیم و نورون ۳ و ۴ را خاموش قرار دهیم. در این صورت داریم،

$$y = \text{sign} \left( \frac{1}{1 + \exp(\omega_5(\omega_1 x_1 + c \omega_3 x_2) + \omega_6(c \omega_2 x_1 + c \omega_4 x_2))} - \frac{1}{2} \right)$$

$$= \text{sign} \left( \frac{1}{1 + \exp(\underbrace{c(\omega_1 \omega_5 + \omega_2 \omega_6)}_{-\beta_1} x_1 + \underbrace{c(\omega_3 \omega_5 + \omega_4 \omega_6)}_{-\beta_2} x_2)} - \frac{1}{2} \right)$$

چنین رابطه‌ای به ازای هر فرضی شبکه ۱+ سود، یعنی  $P(y=+1 | x) > \frac{1}{2}$  است و اگر خودی شبکه ۱- شود یعنی  $P(y=+1 | x) < \frac{1}{2}$  است و این دقیقاً معادل Logistic Regression است.

سوال ۷

به طور کلی خطای یک مدل (اگرسیون یا طبقه بندی) را می توان به صورت حاصل جمع دو  
نرم فزشت: **تخرم بایس** و **نرم داریانس**

حرفی مدلل سیدہ ترابند، خطای واریانس آن اترائیں می باید دی خطای بایس آن کا حصہ می باید

و ~ ~ سا دتر ~ ~ کا حصہ ~ ~ اترائیں ~