به نام خدا

تمرین سری اول - مهدی کاظمی ۹۹۲۱۰۷۵۳

1) .۱

• ضریب: $\int_{-\infty}^{+\infty} f_X(x)\,dx = 1 \Rightarrow \cancel{\int_{-\infty}^{0} 0\,dx} + \int_{0}^{1} cx\,dx + \int_{1}^{2} c(2-x)\,dx + \cancel{\int_{2}^{+\infty} 0\,dx} = 1$

$$\int_0^1 cx\,dx = \frac{c}{2}x^2 + c_1 \Big|_0^1 = \frac{c}{2} + c_1 - (0 + c_1) = \frac{c}{2}$$

$$\int_1^2 c(2-x)\,dx = -\frac{c}{2}(2-x)^2 \Big|_1^2 = -\frac{c}{2}\left[0-1\right] = \frac{c}{2}$$

$$\Rightarrow \frac{c}{2} + \frac{c}{2} = 1 \Rightarrow \boxed{c = 1}$$

• CDF

$$CDF_X(x) = \int_{-\infty}^{x} f_T(t)\,dt \quad , \quad f_X(x) = \begin{cases} x & 0 \leq x \leq 1 \\ 2-x & 1 \leq x \leq 2 \\ 0 & \text{o.w.} \end{cases}$$

$$x \leq 0 \rightarrow F_X(x) = \int_{-\infty}^{x} 0\,dt = 0$$

$$0 \leq x \leq 1 \rightarrow F_X(x) = \int_{-\infty}^{x} t\,dt = \int_0^x t\,dt = \frac{t^2}{2}\Big|_0^x = \frac{x^2}{2}$$

$$1 \leq x \leq 2 \rightarrow F_X(x) = \int_{-\infty}^{x} f(t)\,dt = \underbrace{\int_0^1 t\,dt}_{0 \leq x \leq 1} + \underbrace{\int_1^x (2-t)\,dt}_{1 \leq x \leq 2}$$

$$= \frac{t^2}{2}\Big|_0^1 + \left(-\frac{1}{2}(2-t)^2\right)\Big|_1^x = \frac{1}{2} + \frac{1}{2} - 2 - \frac{x^2}{2} + 2x = -\frac{x^2}{2} + 2x - 1$$

$$x \geq 2 \rightarrow F_X(x) = \int_{-\infty}^{x} f(t)\,dt = 1$$

s.a.m

$$CDF(x) = \begin{cases} 0 & x \leqslant 0 \\ \frac{x^2}{2} & 0 \leqslant x \leqslant 1 \\ -\frac{x^2}{2} + 2x - 1 & 1 \leqslant x \leqslant 2 \\ 1 & x \geqslant 2 \end{cases}$$

- Expected value

$$EX = \int_{-\infty}^{+\infty} x f_X(x)\, dx = \int_0^1 x \times x\, dx + \int_1^2 x(2-x)\, dx = \int_0^1 x^2\, dx + \int_1^2 2x - \int_1^2 x^2\, dx$$

$$= \frac{1}{3}x^3\Big|_0^1 + x^2\Big|_1^2 - \frac{1}{3}x^3\Big|_1^2 = \frac{1}{3} + (4-1) - \frac{1}{3}[8-1] = \frac{1}{3} + \frac{9}{3} - \frac{7}{3} = \frac{3}{3} = 1$$

$$E_{x\sim p}[f(x)] = 1$$

1.2

$$p(X=k) = \theta(1-\theta)^{k-1} \quad k \in \{1, 2, \dots\}$$

$$L(\theta; x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n | \theta) \xrightarrow[\text{random vars}]{\text{independent}} = f(x_1|\theta) \times f(x_2|\theta) \cdots \times f(x_n|\theta)$$

$$\theta_{ML} = \underset{\theta}{\arg\max}\, L(\theta; x_1, \dots, x_n) = \underset{\theta}{\arg\max}\, \log(L(\theta; x_1, \dots, x_n))$$

$$\Rightarrow L(\theta; x_1, \dots, x_n) = \prod_{i=1}^{N} \theta(1-\theta)^{x_i - 1} \Rightarrow l = \log(L(\theta; x_1, \dots, x_n)) = \log\left(\prod_{i=1}^{N} \theta(1-\theta)^{x_i-1}\right)$$

$$= \sum_{i=1}^{N} \log(\theta(1-\theta)^{x_i-1}) = \sum_{i=1}^{N} [\log\theta + \log(1-\theta)^{x_i-1}]$$

$$= \sum_{i=1}^{N} \log\theta + \sum_{i=1}^{N} [(x_i-1)\log(1-\theta)] = N\log\theta + \sum_{i=1}^{N} x_i \log(1-\theta) - \sum_{i=1}^{N} \log(1-\theta)$$

$$= N\log\theta + N\bar{x}\log(1-\theta) - N\log(1-\theta)$$

$$\frac{\partial l}{\partial \theta} = \frac{\partial(N\log\theta + N\bar{X}\log(1-\theta) - N\log(1-\theta))}{\partial\theta} = N\frac{1}{\theta} + N\bar{X}\frac{-1}{(1-\theta)} + N\frac{1}{(1-\theta)}$$

$$\Rightarrow \frac{N}{\theta} - \frac{N\bar{X}}{(1-\theta)} + \frac{N}{(1-\theta)} = 0 \Rightarrow N(1-\theta) - N\bar{X}\theta + N\theta = 0$$

$$= N - N\theta - N\bar{X}\theta + N\theta = 0 \Rightarrow N - N\bar{X}\theta = 0 \Rightarrow N\bar{X}\theta = N \Rightarrow \theta = \frac{1}{\bar{X}}$$

$$\theta_{ML} = \frac{1}{\bar{X}}$$

2.1

- $$\frac{da^Tx}{dx} = \frac{d[a_1, a_2, \dots, a_n]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}{dx} = \frac{d\overbrace{(a_1x_1 + a_2x_2 + \dots + a_nx_n)}^{P}}{d[x_1, x_2, \dots, x_n]^T}$$

$$= \left[\frac{\partial P}{\partial x_1}, \frac{\partial P}{\partial x_2}, \dots, \frac{\partial P}{\partial x_n}\right] = [a_1, a_2, \dots, a_n] = a^T$$

$$\frac{dx^Ta}{dx} = \frac{d[x_1, x_2, \dots, x_n]\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}}{dx} = \frac{d\overbrace{(x_1a_1 + x_2a_2 + \dots + x_na_n)}^{P}}{d[x_1, x_2, \dots, x_n]^T}$$

$$= \left[\frac{\partial P}{\partial x_1}, \frac{\partial P}{\partial x_2}, \dots, \frac{\partial P}{\partial x_n}\right] = [a_1, a_2, \dots, a_n] = a^T$$

$$\Rightarrow \frac{da^Tx}{dx} = \frac{dx^Ta}{dx} = a^T$$

- $$\frac{d(x^Ta)^2}{dx} = 2 \times \frac{d(x^Ta)}{dx} \times (x^Ta) = 2 \times a^T \times x^Ta = 2\, x^Ta\, a^T$$

$$\bullet\ \frac{dx^T A x}{dx} = d[x_1, x_2, \dots, x_n] \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & & & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= d\left[x_1 A_{11} + x_2 A_{21} + \dots + x_n A_{n1}, \dots, \sum_{i=1}^{n} x_i A_{in}\right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} / dx$$

$$= d\left[x_1 \sum_{i=1}^{n} x_i A_{i1} + x_2 \sum_{i=1}^{n} x_i A_{i2} + \dots + x_n \sum_{i=1}^{n} x_i A_{in}\right] / dx$$

$$= d \underbrace{\sum_{j=1}^{n} x_j \sum_{i=1}^{n} x_i A_{ij}}_{S} / dx = \frac{dS}{dx} = \left[\frac{\partial S}{\partial x_1}, \frac{\partial S}{\partial x_2}, \dots, \frac{\partial S}{\partial x_n}\right]$$

$$= [x_1 A_{11} + x_2 A_{21} + \dots + x_n A_{n1} + x_1 A_{11} + x_2 A_{12} + \dots x_n A_{1n},$$

$$\vdots$$

$$x_1 A_{1n} + x_2 A_{2n} + \dots + x_n A_{nn} + x_1 A_{n1} + x_2 A_{n2} + \dots + x_n A_{nn}]$$

$$= \left[\sum_{i=1}^{n} x_i (A_{1i} + A_{i1}), \sum_{i=1}^{n} x_i (A_{2i} + A_{i2}), \dots, \sum_{i=1}^{n} x_i (A_{ni} + A_{in})\right]$$

$$= x^T (A^T + A)$$

2.2

$$H = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \Rightarrow HH^T = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix} = A$$

$$A\vec{v} = \lambda \vec{v} \Rightarrow (A - \lambda I_2)\vec{v} = 0 \Rightarrow \det(A - \lambda I_2) = 0$$

$$A - \lambda I_2 = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 5-\lambda & 2 \\ 2 & 2-\lambda \end{bmatrix} \Rightarrow \det(A - \lambda I_2) = (5-\lambda)(2-\lambda) - 4$$

$\Rightarrow$ Characteristic Polynom. $= (5-\lambda)(2-\lambda) - 4 = 0 \Rightarrow \lambda = 1, 6 \leftarrow$ eigen values

s.a.m

$\lambda = \begin{cases} 1 \\ 6 \end{cases} \qquad (\lambda I_2 - A)\vec{v} = 0$

$\lambda = 1 \Rightarrow \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}\right)\vec{v} = \begin{bmatrix} -4 & -2 \\ -2 & -1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} -4 & -2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\Rightarrow -4v_1 - 2v_2 = 0 \Rightarrow -4v_1 = 2v_2 \Rightarrow -2v_1 = v_2 \Rightarrow \begin{pmatrix} 1 \\ -2 \end{pmatrix} t$ = eig-vecs-1

$\lambda = 6 \Rightarrow \left(\begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix} - \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}\right)\vec{v} = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}\vec{v} = 0 \Rightarrow \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\Rightarrow v_1 - 2v_2 = 0 \Rightarrow v_1 = 2v_2 \Rightarrow \begin{pmatrix} 2 \\ 1 \end{pmatrix} t$ = eig-vecs-2

* پاسخ کلی بردار ویژه‌ها برای $\lambda=1$ به صورت $\begin{pmatrix} t \\ -2t \end{pmatrix}$ و برای $\lambda=6$ به صورت $\begin{pmatrix} 2t \\ t \end{pmatrix}$ است و ما می‌خواهیم

نرمال که اندازه‌ای برابر ۱ داشته باشد محاسبه کنیم به صورت زیر عمل می‌کنیم:

$\left\|\begin{pmatrix} t \\ -2t \end{pmatrix}\right\| = 1 \Rightarrow \sqrt{t^2 + 4t^2} = \sqrt{5t^2} = t\sqrt{5} = 1 \Rightarrow t = \frac{1}{\sqrt{5}} \Rightarrow ev1 = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{pmatrix}$

$\left\|\begin{pmatrix} 2t \\ t \end{pmatrix}\right\| = 1 \Rightarrow \sqrt{4t^2 + t^2} = \sqrt{5t^2} = t\sqrt{5} = 1 \Rightarrow t = 1/\sqrt{5} \Rightarrow ev2 = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix}$

3.1

$J(w) = \sum_{i=1}^{n} (y^{(i)} - w^T x^{(i)})^2$ dims $\Rightarrow$ $X : (n \times (d+1))$ $w : ((d+1) \times 1)$ $y : (n \times 1)$

$J(w) = \|y - Xw\|^2 = (y - Xw)^T (y - Xw) = (y^T - w^T X^T)(y - Xw)$

$= (y^T y - y^T X w - w^T X^T y + w^T X^T X w) \xrightarrow{\dim(w^T X^T y) = 1} = (y^T y - y^T X w - (y^T X w) + w^T X^T X w$

$= (y^T y - 2 y^T X w + w^T X^T X w)$

$\nabla_w J(w) = 0 \Rightarrow \nabla_w (y^T y - 2 y^T X w + w^T X^T X w) = 0$

s.a.m

$$\Rightarrow \nabla_w J(w) = -2(y^T X)^T + 2X^T X w = -2X^T y + 2X^T X w = 0$$

$$\Rightarrow 2X^T X w = 2X^T y \Rightarrow w = (X^T X)^{-1} X^T y$$

(3.2)

* یکی از مشکلاتی که در استفاده از رابطه به دست آمده وجود دارد، این است که در این رابطه نیاز است که معکوس ماتریس X^TX محاسبه شود در حالیکه ممکن است این ماتریس معکوس پذیر نباشد. می توان برای محاسبه مقادیر بهینه برای وزن ها از روش های iterative مثل gradient descent استفاده کنیم.

(3.3) $$J(w) = \sum_{i=1}^{n} (y^{(i)} - w^T x^{(i)})^2 + \|w\|^2 = \|y - Xw\|^2 + \|w\|^2$$

$$= (y - Xw)^T (y - Xw) + w^T w = (y^T y - 2y^T X w + w^T X^T X w) + w^T w$$

$$= y^T y - 2y^T X w + (w^T X^T X + w^T) w = y^T y - 2y^T X w + [w^T (X^T X + I_{d+1})] w$$

to find $w^* = \arg\min_w J(w) \Rightarrow \nabla_w J(w) = 0 \Rightarrow \nabla_w (y^T y - 2y^T X w + [w^T (X^T X + I_{d+1})] w) = 0$

$$= -2(y^T X)^T + \cancel{2}(X^T X + I) w = 0 \Rightarrow (X^T X + I) w = X^T y$$

$$\Rightarrow w = (X^T X + I)^{-1} X^T y$$

s.a.m

3.4

$$J(w) = \sum_{i=1}^{n} f_i \left(y^{(i)} - w^T x^{(i)}\right)^2$$

اگر مقادیر تفاوت بین $\hat{y}$ و y را در یک بردار ستونی مثل r بریزیم در اینصورت برای محاسبه تابع هزینه با وزن های مختص هر داده می توانیم به صورت زیر عمل کنیم:

$$r = \begin{bmatrix} y^{(1)} - w^T x^{(1)} \\ \vdots \\ y^{(n)} - w^T x^{(n)} \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}$$

$$\Rightarrow J(w) = [r_1, r_2, \dots, r_n] \begin{bmatrix} f_1 & 0 & \cdots & 0 \\ 0 & f_2 & \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & & f_n \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} = r^T F r$$

$$\Rightarrow J(w) = (y - Xw)^T F (y - Xw) = (y^T - w^T X^T) F (y - Xw)$$

$$= y^T F y - y^T F X w - \underbrace{w^T X^T F y}_{1\times 1} + w^T X^T F X w$$

$$= y^T F y - 2 y^T F X w + w^T X^T F X w$$

$$w^* = \arg\min J(w) \Rightarrow \nabla_w J(w) = 0 \Rightarrow \nabla_w \left(y^T F y - 2 y^T F X w + w^T X^T F X w\right)$$

$$= -2 (y^T F X)^T + 2 X^T F X w = 0 \Rightarrow X^T F X w = X^T F^T y$$

$$\Rightarrow W^* = (X^T F X)^{-1} X^T F y$$

s.a.m

داده‌های آموزش:

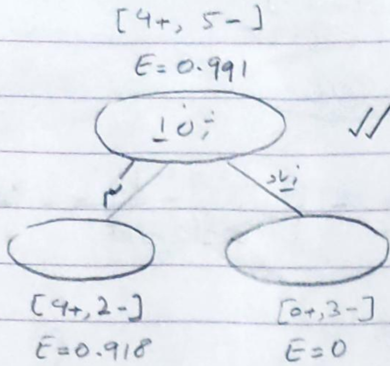
نام شخص	ژن ۱	ژن ۲	ژن ۳	ژن ۴	ژن ۵	بیمار بودن
X1	کم	کم	زیاد	کم	زیاد	بله
X2	زیاد	کم	متوسط	کم	کم	خیر
X3	کم	زیاد	زیاد	کم	زیاد	بله
X4	زیاد	کم	کم	کم	کم	خیر
X5	کم	زیاد	متوسط	زیاد	کم	بله
X6	کم	کم	متوسط	زیاد	زیاد	خیر
X7	کم	زیاد	زیاد	زیاد	زیاد	خیر
X8	کم	کم	کم	کم	کم	بله
X9	زیاد	زیاد	متوسط	زیاد	کم	خیر
X10	کم	زیاد	کم	زیاد	کم	خیر

داده‌های تست:

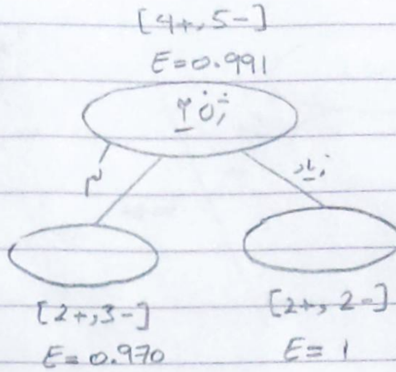
نام شخص	ژن ۱	ژن ۲	ژن ۳	ژن ۴	ژن ۵	بیمار بودن
Y_1	کم	کم	زیاد	کم	زیاد	بله
Y_2	زیاد	زیاد	کم	زیاد	زیاد	خیر
Y_3	کم	کم	زیاد	کم	زیاد	بله
Y_4	کم	کم	متوسط	کم	زیاد	خیر
Y_5	کم	زیاد	متوسط	کم	زیاد	بله
Y_6	کم	کم	زیاد	کم	کم	خیر

$$E = -\sum_{i} P(x_i) \log_2 P(x_i)$$

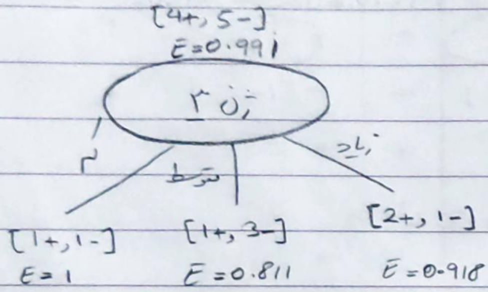
(۴-۱)



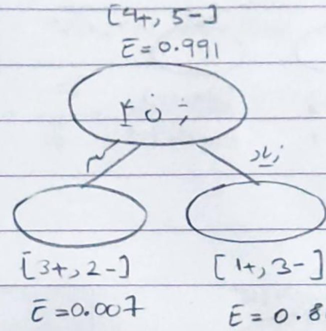
Information Gain $= 0.991 - \frac{6}{9} \times 0.918 - \frac{3}{9} \times 0 = 0.379$



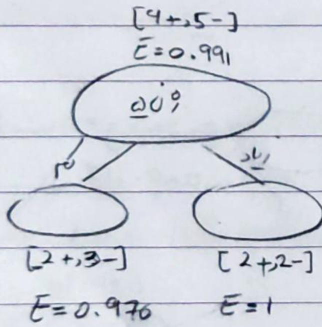
$IG = 0.991 - \frac{5}{9} \times 0.970 - \frac{4}{9} \times 1 = 0.007$



$IG = 0.991 - \frac{2}{9} \times 1 - \frac{4}{9} \times 0.811 - \frac{3}{9} \times 0.918 = 0.102$

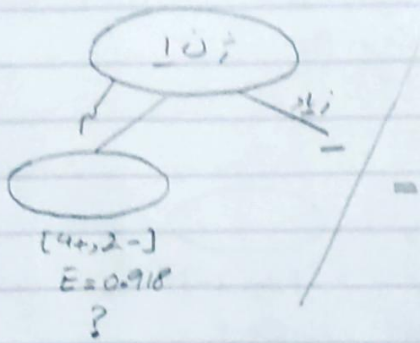


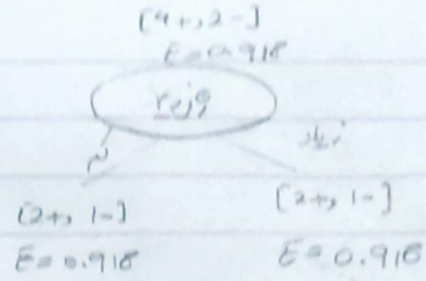
$IG = 0.091$



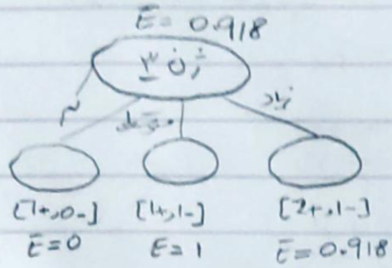
$IG = 0.007$

* با مقایسه IG ها بین ویژگی‌ها، ژن ۱ دارای بیشترین مقدار است و بنابراین این ژن در ریشه قرار می‌گیرد. با انجام محاسباتی مانند این محاسبات درخت را به صورت کامل رسم می‌کنیم که در ادامه آمده است.

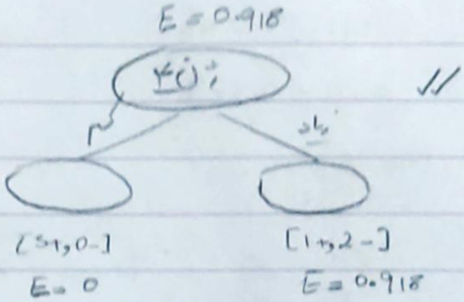




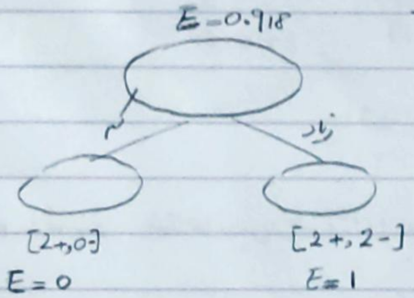
IG = 0



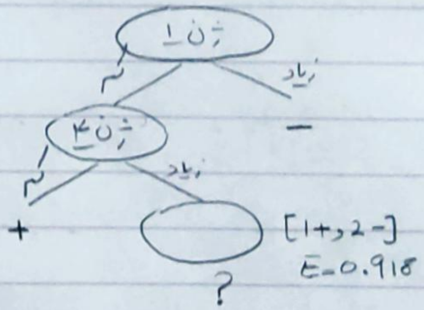
IG = 0.126

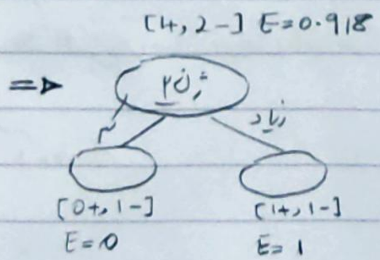


IG = 0.459

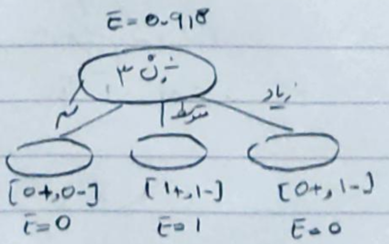


IG = 0.252

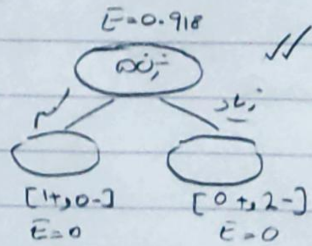




IG = 0.252



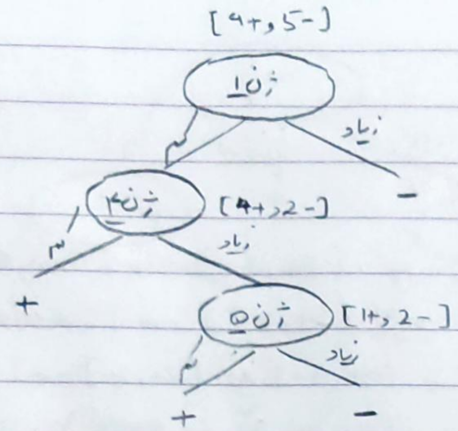
IG = 0.252



IG = 0.918

s.a.m

Decision Tree



جدول عملکرد درخت روی داده های تست

Name	Y	$\hat{Y}$
Y_1	+	+
Y_2	−	−
Y_3	+	+
Y_4	−	+
Y_5	+	+
Y_6	−	+

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

$$Accuracy = \frac{4}{6} = 0.66$$

4.2

* در صورتیکه ژن شماره ۲ و یا ژن شماره ۳ در داده های ما وجود نداشتند، باز هم درخت همین صورت بالا تشکیل می شد و تغییری در آن ایجاد نمی شد؛ زیرا که در هیچ یک از مراحل ساخت درخت و انتخاب attribute، این ژن ها انتخاب نشده اند و در صورت عدم وجودشان باز هم Information Gain که معیار ماکزیمم دارد از بین سایر ژن ها انتخاب می شود و درخت حاصل تغییری نخواهد کرد.

4.3

X5 : سالم => X5 : −

انتخاب ریشه :

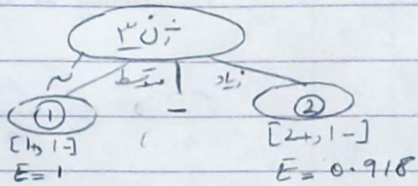
ژن۱ : کم : [3+, 3-] , زیاد : [0+, 3-] => IG = 0.324

ژن۲ : کم : [2+, 3-] , زیاد : [1+, 3-] => IG = 0.091

ژن۳ : کم : [1+, 1-] , متوسط : [0+, 4-] , زیاد : [2+, 1-] => IG = 0.463 //

ژن۴ : کم : [3+, 2-] , زیاد : [0+, 4-] => IG = 0.452

ژن۵ : کم : [1+, 4-] , زیاد : [2+, 2-] => IG = 0.146



انتخاب برای ۱ :

ژن۱ : کم : [1+, 0-] , زیاد : [0+, 1-] => IG = 1 //

ژن۲ : کم : [1+, 1-] , زیاد : [0+, 0-] => IG = 0

ژن۴ : کم : [1+, 1-] , زیاد : [0+, 0-] => IG = 0

ژن۵ : کم : [1+, 1-] , زیاد : [0+, 0-] => IG = 0

انتخاب برای ۲ :

ژن۱ : کم : [2+, 1-] , زیاد : [0+, 0-] => IG = 0.082

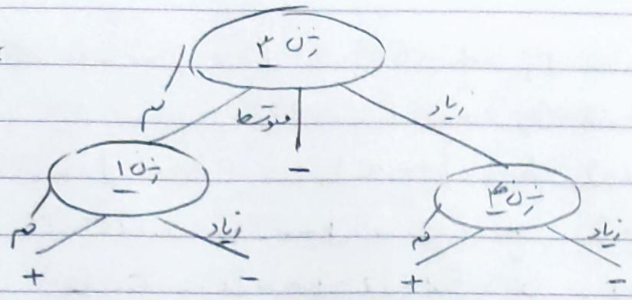
ژن۲ : کم : [1+, 0-] , زیاد : [1+, 1-] => IG = 0.333

ژن۴ : کم : [2+, 0-] , زیاد : [0, 1-] => IG = 1 //

ژن۵ : کم : [0, 0] , زیاد : [2+, 1-] => IG = 0.082

s.a.m

Decision Tree:



جدول عملکرد درخت روی داده های تست :

Name	Y	$\hat{Y}$
y_1	+	+
y_2	−	−
y_3	+	+
y_4	−	−
y_5	+	−
y_6	−	+

$$Accuracy = \frac{4}{6} = 0.66$$

(4.4) * اگر داده اضافه شده X10 شامل ژن های دیگری علاوه بر ۹ ژن قبلی باشد، اطلاعات حاصل از این ژن ها را دور بریزید و با اطلاعات همان ۹ ژن درخت را ایجاد می کنیم؛ زیرا که سایر افراد برای این ژن ها مقداری ندارند و نمی توان برای ساخت درخت، فقط از اطلاعات یک بیمار استفاده کرد و اگر اینکار را کنیم مقدار IG برای تمام این ژن ها ماکزیمم می شوند و باید انتخاب شوند که این کار درستی نیست و باید حذف شوند، حال مراحل ساخت درخت را با داده اضافه شده، بررسی می کنیم.

انتخاب ریشه :

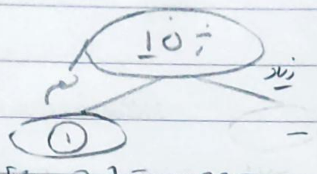
ژن۱ : کم : [4+, 3-] ، زیاد : [0+, 3-] => IG = 0.281 //

ژن۲ : کم : [2+, 8-] ، زیاد : [2+, 3-] => IG = 0

ژن۳ : کم : [1+, 2-] ، متوسط : [1+, 3-] ، زیاد : [2+, 1-] => IG = 0.095

ژن۴ : کم : [3+, 2-] ، زیاد : [1+, 4-] => IG = 0.125

ژن۵ : کم : [2+, 4-] ، زیاد : [2+, 2-] => IG = 0.02



[4+, 3-] E = 0.985

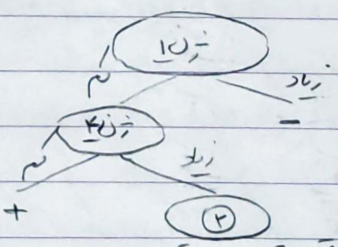
انتخاب برای ① :

ژن۲ : کم : [2+, 1-] ، زیاد : [2+, 2-] => IG = 0.020

ژن۳ : کم : [1+, 1-] ، متوسط : [1+, 1-] ، زیاد : [2+, 1-] => IG = 0.20

ژن۴ : کم : [3+, 0-] ، زیاد : [1+, 3-] => IG = 0.522 //

ژن۵ : کم : [2+, 1-] ، زیاد : [2+, 2-] => IG = 0.020



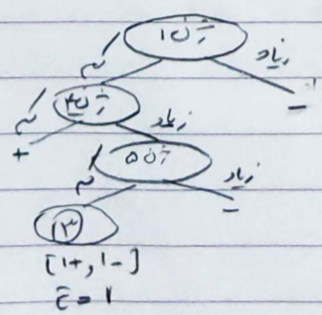
[1+, 3-] E = 0.811

انتخاب برای ②

ژن۲ : کم : [0+, 1-] ، زیاد : [1+, 2-] => IG = 0.123

ژن۳ : کم : [0+, 1-] ، متوسط : [1+, 1-] ، زیاد : [0+, 1-] => IG = 0.311 //

ژن۵ : کم : [1+, 1-] ، زیاد : [0+, 2-] => IG = 0.311 //



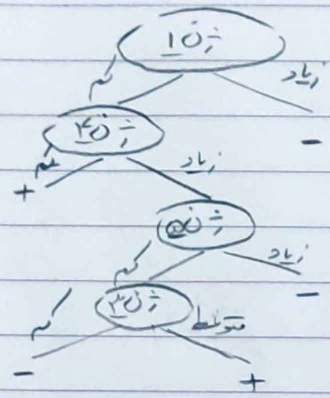
[1+, 1-]
E = 1

انتخاب برای ④

ت۲۵ : کم [0+, 0-], زیاد : [1+, 1-] ⇒ IG = 0

ت۳۵ : کم [0+, 1-], متوسط : [1+, 0-], زیاد : [0+, 0-] ⇒ IG = 1 //

Decision Tree



جدول عملکرد درخت روی داده تست

Name	ی	ی̂
y_1	+	+
y_2	-	-
y_3	+	+
y_4	-	+
y_5	+	+
y_6	-	+

$$Accuracy = \frac{4}{6} = 0.66$$

s.a.m