

## Assignment 3

**Make the calculations in the following questions by hand and show the solution.**

1. Find the Eigenvalues and Eigenvectors for the below matrix : (40 points)

$$\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$$

2. Create the decision tree with the **ID3 algorithm (use Information Gain as criteria)** with the SATISFIED class being the target attribute based on the following observations. (30 points)

CUSTOMER	INCOME	EDUCATION LEVEL	SECTOR	SATISFIED
1	NORMAL	HIGH SCHOOL	INFORMATION TECH	YES
2	HIGH	MIDDLE SCHOOL	INFORMATION TECH	YES
3	LOW	MIDDLE SCHOOL	CONSTRUCTION	YES
4	HIGH	HIGH SCHOOL	CONSTRUCTION	YES
5	LOW	HIGH SCHOOL	CONSTRUCTION	YES
6	HIGH	BACHELOR	CONSTRUCTION	YES
7	LOW	BACHELOR	CONSTRUCTION	YES
8	HIGH	HIGH SCHOOL	INFORMATION TECH	NO
9	LOW	HIGH SCHOOL	INFORMATION TECH	NO
10	HIGH	BACHELOR	INFORMATION TECH	NO
11	LOW	BACHELOR	INFORMATION TECH	NO

3. In the dataset below, (X1, X2) indicates the observation feature coordinates and the DECISION is the target feature. Based on this information, determine which class the observation (8,8) belongs to using the nearest k-neighborhood algorithm. (30 points)

**(k=3, using Euclidean distance and weighted voting)**

Observation	X1	X2	DECISION
1	9	1	YES
2	3	5	YES
3	1	9	NO
4	4	9	NO
5	4	6	YES
6	2	1	NO
7	2	10	YES
8	5	5	YES

① Eigenvalues  $Cv = \lambda v \Rightarrow |C - \lambda I| = 0$

$$C = \begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$$

$$\left| \begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right| = 0$$

$$\left| \begin{bmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{bmatrix} \right| = 0$$

1st row expansion

$$(8-\lambda) \cdot \underline{M_{1,1}} - (-6) \cdot \underline{M_{1,2}} + 2 \cdot \underline{M_{1,3}} = 0$$

$$(8-\lambda) \cdot \underline{\begin{vmatrix} 7-\lambda & -4 \\ -4 & 3-\lambda \end{vmatrix}} + 6 \cdot \underline{\begin{vmatrix} -6 & -4 \\ 2 & 3-\lambda \end{vmatrix}} + 2 \cdot \underline{\begin{vmatrix} -6 & 7-\lambda \\ 2 & -4 \end{vmatrix}} = 0$$

$$\begin{aligned} M_{1,1} \begin{vmatrix} 7-\lambda & -4 \\ -4 & 3-\lambda \end{vmatrix} &= (7-\lambda)(3-\lambda) - (-4)(-4) \\ &= 21 - 7\lambda - 3\lambda + \lambda^2 - 16 \\ &= \lambda^2 - 10\lambda + 5 \end{aligned}$$

$$\begin{aligned} M_{1,2} \begin{vmatrix} -6 & -4 \\ 2 & 3-\lambda \end{vmatrix} &= (-6)(3-\lambda) - (-4)(2) \\ &= 6\lambda - 18 + 8 \\ &= 6\lambda - 10 \end{aligned}$$

$$\begin{aligned} M_{1,3} \begin{vmatrix} -6 & 7-\lambda \\ 2 & -4 \end{vmatrix} &= (-6)(-4) - (2)(7-\lambda) \\ &= 24 - (14 - 2\lambda) \\ &= 2\lambda + 10 \end{aligned}$$

$$(8-\lambda)(\lambda^2-10\lambda+5) + (6)(6\lambda-10) + 2(2\lambda+10) = 0$$

$$8\lambda^2 - 80\lambda + 40 - \lambda^3 + 10\lambda^2 - 5\lambda + 36\lambda - 60 + 4\lambda + 20 = 0$$

$$-\lambda^3 + 18\lambda^2 - 45\lambda = 0$$

$$\lambda(-\lambda^2 + 18\lambda - 45) = 0$$

$$\lambda = 0 \quad \text{or} \quad \lambda = \frac{-18 \pm \sqrt{324 - 180}}{-2}$$

$$\lambda = \frac{-18 \pm 12}{-2}$$

$$\lambda = 3, 15$$

$$\lambda_1 = 0$$

$$\lambda_2 = 3$$

$$\lambda_3 = 15$$

Eigen vectors

$$\left| \begin{bmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{bmatrix} \right| = 0$$

$$\begin{array}{l} \lambda_1 = 0 \\ \lambda_2 = 3 \\ \lambda_3 = 15 \end{array}$$

$$\lambda_1 \begin{bmatrix} (8-0) & -6 & 2 \\ -6 & (7-0) & -4 \\ 2 & -4 & (3-0) \end{bmatrix} = 0$$

$$\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

$$8x - 6y + 2z = 0$$

$$-6x + 7y - 4z = 0$$

$$2x - 4y + 3z = 0 \Rightarrow -2x = -4y + 3z \Rightarrow x = \frac{4y - 3z}{2}$$

$$8\left(\frac{4y - 3z}{2}\right) - 6y + 2z = 0 \rightarrow x = \frac{4y - 3z}{2}$$

$$16y - 12z - 6y + 2z = 0$$

$$10y - 10z = 0$$

$$y = z$$

$$x = \frac{4z - 3z}{2}$$

$$x = \frac{z}{2}$$

$$x = \frac{z}{2}, y = z, z = z \quad \text{or} \quad \begin{bmatrix} \frac{1}{2} \\ 1 \\ 1 \end{bmatrix} \text{ vector} \quad \text{or} \quad \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \text{ vector}$$

$$\lambda_2 \begin{bmatrix} (8-3) & -6 & 2 \\ -6 & (7-3) & -4 \\ 2 & -4 & (3-3) \end{bmatrix} = 0$$

$$\begin{bmatrix} 5 & -6 & 2 \\ -6 & 4 & -4 \\ 2 & -4 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

$$5x - 6y + 2z = 0$$

$$-6x + 4y - 4z = 0$$

$$2x - 4y = 0 \Rightarrow x = 2y$$

$$5(2y) - 6y + 2z = 0$$

$$10y - 6y + 2z = 0$$

$$4y + 2z = 0$$

$$z = -2y$$

$$x = 2y, y = y, z = -2y \quad \text{or} \quad \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} \text{ vector}$$

$$\lambda_3 \begin{bmatrix} (8-15) - 6 & 2 \\ -6 & (7-15) - 4 \\ 2 & -4 & (3-15) \end{bmatrix} = 0$$

$$\begin{bmatrix} -7 & -6 & 2 \\ -6 & -8 & -4 \\ 2 & -4 & -12 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

$$-7x - 6y + 2z = 0$$

$$-6x - 8y - 4z = 0$$

$$2x - 4y - 12z = 0 \Rightarrow -2x = -4y - 12z \Rightarrow x = 2y + 6z$$

$$-6(2y + 6z) - 8y - 4z = 0$$

$$-12y - 36z - 8y - 4z = 0$$

$$-20y - 40z = 0$$

$$y + 2z = 0$$

$$y = -2z$$

$$x = 2(-2z) + 6z$$

$$x = -4z + 6z$$

$$x = 2z$$

$$x = 2z, y = -2z, z = z \text{ or } \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} \text{ vector}$$

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, v_2 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, v_3 = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$$

② 103

$$H(S) = - \sum_{x \in X} p(x) \log_2 p(x)$$

Satisfied  $P(\text{Yes}) = 7/11$   $P(\text{No}) = 4/11$

$$\text{Entropy} = - (7/11 \log_2 7/11 + 4/11 \log_2 4/11) \approx 0.9457$$

Income

Low

Yes = 3/5

No = 2/5

$$\text{Entropy} = - (3/5 \log_2 (3/5) + 2/5 \log_2 (2/5)) \approx 0.9710$$

Normal

Yes = 1/1

Entropy = 0

High

Yes = 3/5

No = 2/5

$$\text{Entropy} = - (3/5 \log_2 (3/5) + 2/5 \log_2 (2/5)) \approx 0.9710$$

$$\text{Info Gain} = 0.9457 - \left[ 5/11 (0.9710) + \underbrace{1/11 (0)}_{0.8827} + 5/11 (0.9710) \right] = 0.0630$$

Educb.ion

Middle

Yes = 2/2

Entropy = 0

High

Yes = 3/5

No = 2/5

$$\text{Entropy} = - (3/5 \log_2 (3/5) + 2/5 \log_2 (2/5)) \approx 0.9710$$

Bachelors

Yes = 2/4

No = 2/4

Entropy = 1

0.805

$$\text{Info Gain} = 0.9457 - \left[ 2/11 (0) + 5/11 (0.9710) + \underbrace{4/11 (1)}_{0.805} \right] = 0.1407$$

## Sector

Construction

Yes = 5/5

Entropy = 0

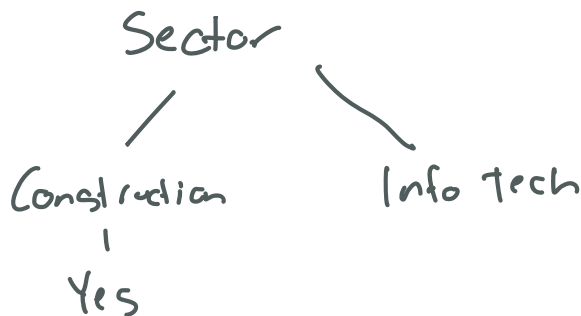
Info Tech

Yes = 2/6

No = 4/6

Entropy =  $-(2/6 \log_2(2/6) + 4/6 \log_2(4/6)) \approx 0.9183$  use in next split

Info Gain =  $0.9457 - [5/11(0) + \overset{0.5009}{6/11}(0.9183)] = 0.4448$  Highest info gain



## Income

Low

No = 2/2

Entropy = 0

Normal

Yes = 1/1

Entropy = 0

High

Yes = 1/3

No = 2/3

Entropy =  $-(1/3 \log_2(1/3) + 2/3 \log_2(2/3)) \approx 0.9183$

Info Gain =  $0.9183 - [3/6(0.9183) + \overset{0.45915}{1/6}(0) + 2/6(0)] = 0.4592$

## Education

Middle

Yes = 1/1

Entropy = 0

High

Yes = 1/3

No = 2/3

Entropy =  $-(1/3 \log_2(1/3) + 2/3 \log_2(2/3)) \approx 0.9183$

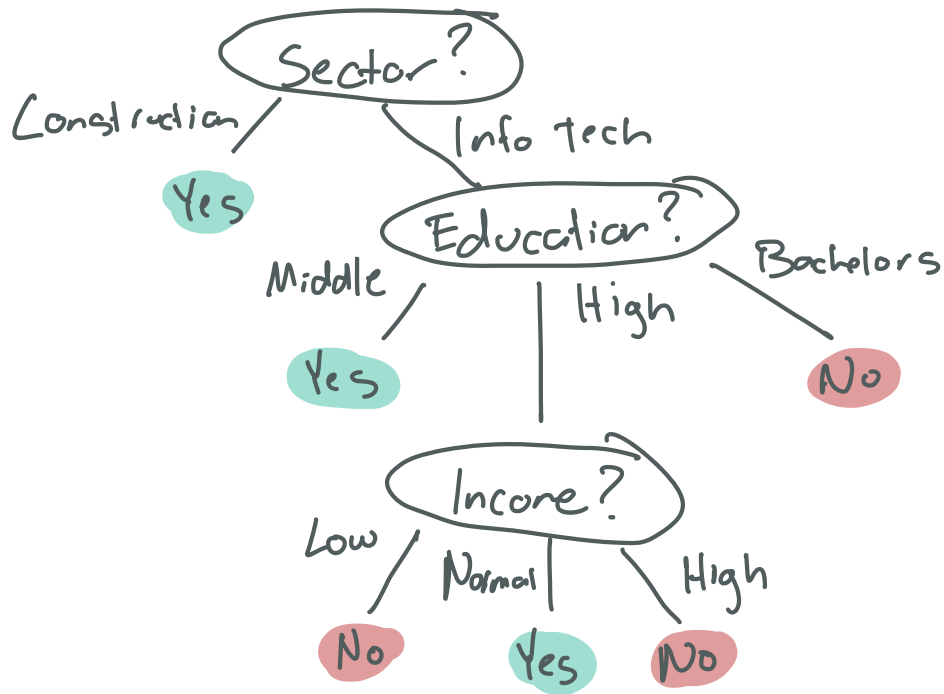
Bachelors  
No = 2/2 Entropy = 0

0.45915

$$\text{Info Gain} = 0.9183 - \left[ \frac{1}{6} (0) + \frac{3}{6} (0.9710) + \frac{2}{6} (0) \right] = 0.4592$$

Tie! Split on Education.

FINAL  
TREE





③ Weighted =  $1/\text{distance}$   $k=3$

Dist (8,8) per obs  $d = \sqrt{(x1_i - 8)^2 + (x2_i - 8)^2}$

1 = 7.0711

2 = 5.8310

3 = 7.0711

4 = 4.1231

5 = 4.4721

6 = 9.2195

7 = 6.3246

8 = 4.2426

4, 5, 8 are nearest 3 neighbors  $\Rightarrow$  5, 8 = Yes 4 = No

Weights

4 =  $1/4.1231 = 0.2425$

5 =  $1/4.4721 = 0.2236$

8 =  $1/4.2426 = 0.2357$

Yes =  $0.2236 + 0.2357 = \underline{0.4593}$

No = 0.2425

(8,8) classified as YES