Assignment - 3

Ervin Debbarma

16+4

1)

Instances	Distance
1	$\sqrt{(162-158)^2+(61-58)^2} = 5$
2	$\sqrt{(162-158)^2+(61-63)^2} = 4.47$
3	$\sqrt{(162-160)^2+(61-64)^2} = \sqrt{4+9} = \sqrt{13} = 3.6$
4	$\sqrt{(162-163)^2+(61-64)^2} = 3.16$
5	$\sqrt{(162-163)^2+(61-61)^2} = 1$
6	$\sqrt{(162-163)^2+(61-60)^2} = 1.414$
7	$\sqrt{(162-165)^2+(61-65)^2} = \sqrt{9+16} = \sqrt{25} = 5$
8	$\sqrt{(162-165)^2+(61-62)^2} = \sqrt{9+1} = 3.16$
9	$\sqrt{(162-168)^2+(61-63)^2} = \sqrt{20} = 4.47$

162	61	2

Test - 1

$K = 3$

Instances	Class
5	A
6	A
8	B
4	B
3	B
9	B
2	A
1	A
7	B

∴ Class label = A

Test data-2 : | 170 | 60 | ? |

Instances	Distance
1	$\sqrt{(170-158)^2+(60-58)^2} = 12.16$
2	$\sqrt{(170-158)^2+(60-63)^2} = 12.36$
3	$\sqrt{(170-160)^2+(60-64)^2} = 10.77$
4	$\sqrt{(170-163)^2+(60-64)^2} = 8.06$
5	$\sqrt{(170-163)^2+(60-61)^2} = 7.07$
6	$\sqrt{(170-163)^2+(60-60)^2} = 7$
7	$\sqrt{(170-165)^2+(60-65)^2} = 7.07$
8	$\sqrt{(170-165)^2+(60-62)^2} = 5.38$
9	$\sqrt{(170-168)^2+(60-63)^2} = 3.60$

∴

Instances	Class	
9	B	$K=3$
8	B	
6	A	
7	B	
5	A	
4	B	
3	B	
1	A	
2	A	

∴ Class label = B

2) Evade = yes
= 4
Evade = No = 6

$P(\text{Evade} = \text{yes}) = \frac{4}{10} = 0.4$, $P(\text{Evade} = \text{No}) = \frac{6}{10} = 0.6$

For Evade: Yes

i) $P(\text{Refund} = \text{Yes} \mid \text{Evade} = \text{yes}) = \frac{2}{4} = 0.5$

T_3 : Refund = No, Marital = Single, Income = Low
T_5 : Refund = Yes, Marital = Single, Income = Low
T_8 : Refund = Yes, Marital = Divorced, Income = Medium
T_9 : Refund = No, Marital = married, Income = Medium

ii) $P(\text{Marital} = \text{single} \mid \text{Evade} = \text{yes}) = \frac{2}{4} = 0.5$

iii) $P(\text{TaxableIncome} = \text{Low} \mid \text{Evade} = \text{Yes}) = \frac{2}{4} = 0.5$

Total: $= 0.4 \times 0.5 \times 0.5 \times 0.5$
$= 0.005$

For Evade: No

T_1 (Refund = Yes, Marital = Single, Inc = High)
T_2 (Refund = No, Marital = married, Inc = Med)
T_4 (Refund = No, Marital = Married, Inc = Low)
T_6 (Refund = No, marital = divorced, inc = med)
T_7 (refund = No, marital = married, inc = high)
T_{10} (refund = yes, marital = divorced, inc = low)

i) $P(\text{refund} = \text{yes} \mid \text{Evade} = \text{No}) = \frac{2}{6} = 0.33$
ii) $P(\text{marital} = \text{single} \mid \text{Evade} = \text{No}) = \frac{1}{6} = 0.166$
iii) $P(\text{taxableincome} = \text{low} \mid \text{evade} = \text{No}) = \frac{2}{6} = 0.33$
total = $0.6 \times 0.33 \times 0.16 \times 0.33 = 0.0111$

∴ 0.05 > 0.0111
Yes No

∴ Evade = Yes . for (Yes, Single, Low)

3) X = (Refund = Yes, Marital Status = Single, Income = medium)

Evade = Yes = 4
= No = 6

P(Evade = yes) = $\frac{4}{10}$ = 0.4

P(Evade = No) = $\frac{6}{10}$ = 0.6

For Evade : Yes

From T_3, T_5, T_7, T_9 from table.

∴ P(Refund = Yes | Yes) = $\frac{4}{4}$ = 1

P(Marital = single | Yes) = $\frac{1}{4}$ = 0.25

P(Income = Medium | Yes) = $\frac{2}{4}$ = 0.5

∴ Total = 0.4 × 1 × 0.25 × 0.5
= 0.05

For Evade : No $T_1, T_2, T_4, T_6, T_8, T_{10}$.

P(refund = Yes | No) = $\frac{1}{6}$ = 0.1667

P(Marital = single | No) = $\frac{3}{6}$ = 0.5

P(income = medium | No) = $\frac{2}{6}$ = 0.3333

Total

$= 0.6 \times 0.1667 \times 0.5 \times 0.33$

$= 0.01667$

$Evade_{yes} > Evade_{No}$

$0.05 > 0.01667$

4) $X = (Age = 40, Income = 80, student = No)$

Class = yes $\quad \therefore P(Yes) = \frac{4}{10} = 0.4$

= 4 $\quad P(No) = \frac{6}{10} = 0.6$

Class = No = 6

Records with buys = yes $\therefore$ Sl: 1, 3, 7, 9

Age: Yes $\quad f(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$\mu_{Age|yes} : (35 + 40 + 28 + 50)/4 = 38.25$

$\sigma^2 : (35 - 38.25)^2 = 10.5625$

$(40 - 38.25)^2 = 3.0625$

$(28 - 38.25)^2 = 105.0625$

$(50 - 38.25)^2 = 138.0625$

Sum = 256.75

$\sigma^2 = 256.75/4 = 64.1875$

$\sigma_{Age|yes} = 8.0167$

Income: Yes

Values: 95, 85, 50, 90

Mean: $\mu_{Income|yes}$

$= (95+85+50+90)/4$

$= 80$

Variance:

$(95-80)^2 = 15^2 = 225$

$(85-80)^2 = 25$

$(50-80)^2 = 900$

$(90-80)^2 = 100$

Sum $= 225+25+900+100 = 1250$

$\sigma^2 = 1250/4 = 312.5$

$\sigma_{Income|yes} = 17.678$

Records with buys = No : SL: 2, 4, 5, 6, 8, 10

Age (No)

Values: 30, 45, 50, 25, 29, 36

Mean, $\mu_{age|No}$: $(30+45+50+25+29+36)/6$

$= \frac{215}{6} = 35.833$

⇒ Variance:

$(30-35.833)^2 = 34.0278$

$(45-35.833)^2 = 84.0278$

$(50-35.833)^2 = 200.6944$

$(25-35.833)^2 = 117.3611$

$(29-35.833)^2 = 46.6944$

$(36-35.833)^2 = 0.0278$

Sum $= 482.8333$

$\sigma^2 = 482.8333/6 = 80.4722$

$\sigma_{age|No} = 8.976$

Income(No)

$\mu_{Income|No}$: $(55+60+70+60+50+82)/6$

$= 377/6$

$= 62.8333$

$\therefore (55-62.8333)^2 = 61.3889$

$(60-62.8333)^2 = 8.0278$

$(70-62.8333)^2 = 51.3611$

$(60-62.8333)^2 = 8.0278$

$(50-62.8333)^2 = 164.3611$

$(82-62.8333)^2 = 367.3611$

Sum $= 660.5278$

$\sigma^2 = 660.5278/6$

$= 110.088$

$\sigma_{Income|No} = 10.495$

Student Dist. Class = Yes : sl = 1, 3, 7, 9

$P(student = Yes | Yes) = \frac{3}{4} = 0.75$

$P(student = No | Yes) = \frac{1}{4} = 0.25$

Student dist. class = "No" sl: 2, 4, 5, 6, 8, 10

$P(student = Yes | No) = \frac{1}{6}$ $P(student = No/No) = \frac{5}{6}$

$= 0.1667$ $= 0.8333$

$P(Age = 40, Income = 80, Student = No \mid class)$

Likelihood (class = Yes)

$f(Age = 40 \mid \mu = 38.25, \sigma = 8.0167)$

$diff = 40 - 38.25 = 1.75$

$(1.75)^2 = 3.0625$,

$2\sigma^2 \simeq 2 \times (8.0167)^2 \simeq 128.534$

$$\exp\left(\frac{-3.0625}{128.534}\right) \simeq \exp(-0.02384) \simeq 0.9765$$

$$\therefore \frac{1}{\sqrt{2\pi}\sigma} = \frac{1}{2.5066 \times 8.0167} \simeq 0.0498$$

$\therefore f(40) = 0.0498 \times 0.9765$

$= 0.0486$

$f(Income = 80 \mid \mu = 80, \sigma = 17.678)$

When $x = \mu$,

$$f(80) = \frac{1}{\sqrt{2\pi}\,17.678} = \frac{1}{2.5066 \times 17.678}$$

$$= \frac{1}{44.3} = 0.0226$$

$\therefore P(student = No \mid Yes) = 0.25$

Total $= 0.0488 \times 0.0226 \times 0.25$

$= 0.0002745.$

Class (Yes) $= 0.4.$

$\therefore 0.0002745 \times 0.4 = 0.0001098$

Likelihood Class = No

$f(Age = 40 \mid \mu = 35.8333, \sigma \simeq 8.976)$

$diff = 40 - 35.8333$

$= 4.1667$

$(4.1667)^2 = 17.3611$

$2\sigma^2 \simeq 2 \times (8.976)^2 = 2 \times 80.64$

$= 161.28$

$\exp\left(-\frac{17.3611}{161.28}\right) \simeq \exp(-0.1077) \simeq 0.898.$

$\therefore \frac{1}{\sqrt{2\pi}\sigma} = \frac{1}{2.5066 \times 8.976} = \frac{1}{22.51} = 0.0444.$

$f(40) = 0.0444 \times 0.898 = 0.0399$

$f(Income = 80 \mid \mu = 62.8333, \sigma \simeq 0.0399)$

$diff = 80 - 62.8333 = 17.1667.$

$(17.1667)^2 = 294.7778$

$2\sigma^2 = 2 \times (10.495)^2 = 2 \times 110.40 = 220.80$

$\exp\left(-\frac{294.7778}{220.80}\right) = \exp(-1.335) = 0.262$

$\therefore \frac{1}{2.5066 \times 10.495} = \frac{1}{26.28} = 0.038$

$\therefore f(80) \simeq 0.038 \times 0.262$

$= 0.00996$

$\therefore P(Student = No / No) = \frac{5}{6} = 0.8333$

total

$\rightarrow = 0.0399 \times 0.00996 \times 0.8333$

$= 0.0003307$

Now, with class $P(No) = 0.6$

$\therefore No = 0.0003307 \times 0.6 = 0.0001984.$

Buys Computer No ~~Yes~~

$0.0001984 > 0.0001098$

$\therefore$ No