Date:- 02/08/19

CSE 6363 - Machine Learning

Homework-1 Spring 2019

SUNDESH RAJ
1001633297

Q1 :- MLE and MAP

Given:- for a poisson process probability for the first event to occur is described by an exponential distribution

$$\boxed{P_\lambda(x) = \lambda e^{-\lambda x}}$$

(a) MLE for λ of the model given $D = \{k_1 \dots k_n\}$

(i) likelihood function :-

$$L(\lambda : k_1, \dots k_n) = \lambda^n e^{\left(-\lambda \sum_{i=1}^{n} k_i\right)}$$

Proof:- $L(\lambda : k_1, \dots k_n) = \prod_{i=1}^{n} f(k_i : \lambda)$

$$= \prod_{i=1}^{n} \lambda e^{(-\lambda k_i)}$$

$$= \lambda^n e^{\left(-\lambda \sum_{i=1}^{n} k_i\right)}$$

(ii) log likelihood function :-

$$l(\lambda ; k_1 \dots k_n) = \ln\left(L(\lambda ; k_1, \dots k_n)\right)$$

$$= \ln\left(\lambda^n \exp\left(-\lambda \sum_{i=1}^{n} k_i\right)\right)$$

$$\Rightarrow \ln(\lambda^n) + \ln\left(\exp\left(-\lambda \sum_{i=1}^{n} k_i\right)\right)$$

$$= n \ln(\lambda) - \lambda \sum_{i=1}^{n} k_i$$

(iii) Maximum likelihood estimator

$$\hat{\lambda} = \underset{\lambda}{\operatorname{argmax}}\ l(\lambda; k_1, \ldots k_n)$$

$$\Rightarrow \frac{d}{d\lambda} l(\lambda; k_1, \ldots k_n) = 0$$ —— first order condition for a maximum

$$\Rightarrow \frac{d}{d\lambda} l(\lambda; k_1, \ldots k_n) = \frac{d}{d\lambda}\left(n\ln(\lambda) - \lambda \sum_{i=1}^{n} k_i\right)$$

$$= \frac{n}{\lambda} - \sum_{i=1}^{n} k_i$$

By setting the value equal to 0, we get

$$\boxed{\lambda = \frac{n}{\sum_{i=1}^{n} k_i}}$$

(b) given data $D = \{1.5, 3, 2.5, 2.75, 2.9, 3\}$

$$\lambda = \frac{n}{\sum_{i=1}^{n} k_i}$$

$$\Rightarrow \frac{6}{1.5 + 3 + 2.5 + 2.75 + 2.9 + 3}$$

$$\Rightarrow \underline{\underline{0.38338}}$$

1(c) $\hat{\lambda}_{map} = \text{argmax } P(D|\lambda)\, P(\lambda)$

$$= \lambda^{n} e^{-\lambda \Sigma x_i} \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$$

$$= \frac{\beta^{\alpha}}{\Gamma(\alpha)} \cdot \lambda^{n+\alpha-1} e^{-\lambda(\sum_i x_i + \beta)}$$

$$= \text{Gama}\left(\alpha+n, \sum_i x_i + \beta\right)$$

log posterior,

$$\log P(\lambda|D) = (n+\alpha+1)\log\lambda - \lambda\left(\sum_i x_i + \beta\right)$$

$$\frac{\partial}{\partial\lambda} \log(P(\lambda|D)) = \frac{(n+\alpha-1)}{\lambda} - \sum_i x_i - \beta = 0$$

$$\Rightarrow \lambda = \frac{n+\alpha-1}{\sum_i x_i + \beta}$$

Given, $\alpha = 5$, $\beta = 10$, $n = 6$

$$\sum_i x_i = \frac{1.5+3+2.5+2.75+2.9+3}{6} = \underline{2.60833}$$

$$\lambda = \frac{6+5-1}{2.60833+10} \Rightarrow \underline{\underline{0.79312}}$$

② given

D = { ((170, 57, 32), W),
((192, 95, 28), M),
((150, 45, 30), W),
((170, 65, 29), M),
((175, 78, 35), M),
((185, 90, 32), M),
((170, 65, 28), W),
((155, 48, 31), W),
((160, 55, 30), W),
((182, 80, 30), M),
((175, 69, 28), W),
((180, 80, 27), M),
((160, 50, 31), W),
((175, 72, 30), M), }

(a)

(i) Input → (155, 40, 35)

Height	Weight	Age		Distance	
170	57	32	W	22.869	5
192	95	28	M	66.655	
150	45	30	W	8.660	1
170	65	29	M	29.765	
175	78	35	M	42.941	
185	90	32	M	58.386	
170	65	28	W	29.983	
155	48	31	W	8.944	2
160	55	30	W	16.583	4

182	80	30	M	48.518	
175	69	28	W	35.916	
180	80	27	M	47.843	
160	50	31	W	11.874	3
175	72	30	M	38.065	

If $K = 1$

distance $= \sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}$

where x = height, y = weight and z = age

distance = 8.660 Gender = Woman

If $K = 3$

distance $= \sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}$

$$\left.\begin{array}{lll} \text{distance } 1 = 8.660 & \text{Gender} = \text{Woman} \\ 2 = 8.944 & \text{Gender} = \text{Woman} \\ 3 = 11.874 & \text{Gender} = \text{Woman} \end{array}\right\} \text{Woman}$$

If $K = 5$

distance $= \sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}$

distance 1 = 8.660 Gender = Woman
2 = 8.944 Gender = Woman
3 = 11.874 Gender = Woman
4 = 16.583 Gender = Woman
5 = 22.869 Gender = Woman

$\therefore$ (155, 40, 35) $\Rightarrow$ Woman

(ii) Input = (170, 70, 32)

Height	Weight	Age		Distance	
170	57	32	W	13	
192	95	28	M	33.541	
150	45	30	W	32.078	
170	65	29	M	5.830	2
175	78	35	M	9.899	5
185	90	32	M	25	
170	65	28	W	6.403	3
155	48	31	W	26.645	
160	55	30	W	18.138	
182	80	30	M	15.748	
175	69	28	W	6.480	4
180	80	27	M	15	
160	50	31	W	22.383	
175	72	30	M	5.744	1

If k=1, distance = 5.744, Gender = Man

If k=3, distance -> 1. 5.744, Gender = Man
2. 5.830, Gender = Man
3. 6.403, Gender = Woman

Man

If K=5, distance -> 1. 5.744, Gender = Man
2. 5.830, Gender = Man
3. 6.403, Gender = Woman
4. 6.480, Gender = Woman
5. 9.899, Gender = Man

Man

(iii) Input = (175, 70, 35)

Height	Weight	Age		Distance	
170	57	32	W	14.247	
192	95	28	M	31.032	
150	45	30	W	35.707	
170	65	29	M	9.273	4
175	78	35	M	8	3
185	90	32	M	22.561	
170	65	28	W	9.949	5
155	48	31	W	30	
160	55	30	W	21.794	
182	80	30	M	13.190	
175	69	28	W	7.071	2
180	80	27	M	13.747	
160	50	31	W	25.317	
175	72	30	M	5.385	1

If K=1, distance = 5.385, Gender = Male

If K=3, distance 1 = 5.385 Gender = Male
2 = 7.071 Gender = Woman
3 : 8 Gender : male

Man

If K=5 distance 1 = 5.385 Gender = Man
2 : 7.071 Gender : Woman
3 : 8 Gender : Man
4 : 9.273 Gender = Man
5 : 9.949 Gender : Woman

Man

Scanned by CamScanner

(iv) Input :- (180, 90, 20)

Height	Weight	Age		Distance	
170	57	32	W	36.510	
192	95	28	M	15.264	4
150	45	30	W	55	
170	65	29	M	28.390	
175	78	35	M	19.849	5
185	90	32	M	13	2
170	65	28	W	28.089	
155	48	31	W	50.099	
160	55	30	W	41.533	
182	80	30	M	14.282	3
175	69	28	W	23.021	
180	80	27	M	12.206	1
160	50	31	W	46.054	
175	72	30	M	21.1897	

If K=1, distance = 12.206, Gender = Man

If K=3, distance 1 = 12.206, Gender = Man
2 = 13, Gender = Man
3 = 14.282, Gender = Man
} Man

If K=5, distance 1 = 12.206, Gender = Man
2 = 13, Gender = Man
3 = 14.282, Gender = Man
4 = 15.264, Gender = Man
5 = 19.849, Gender = Man
} Man

2C) Repeat KNN prediction removing age data.

Given :-

(i) Input = (155, 40)

Height	weight		Distance	
170	57	W	22.671	5
192	95	M	66.287	
150	45	W	7.071	1
170	65	M	29.154	
175	78	M	42.941	
185	90	M	58.309	
170	65	W	29.154	
155	48	W	8	2
160	55	W	15.811	4
182	80	M	48.259	
175	69	W	35.227	
180	80	M	47.169	
160	50	W	11.180	3
175	72	M	37.735	

If K=1, distance = 7.071, Gender = <u>woman</u>

If K=3, distance → 1) 7.071, Gender = Woman
2) 8, Gender = woman
3) 11.180, Gender = woman

} <u>Woman</u>

If K=5, distance → 1) 7.071, Gender = woman
2) 8, Gender = woman
3) 11.180, Gender = woman
4) 15.811, Gender = woman
5) 22.671, Gender = woman

} <u>Woman</u>

(ii) Input :- (170, 70)

Height	Weight		Distance	
170	57	W	13	5
192	95	M	33.301	
150	45	W	32.015	
170	65	M	5	1
175	78	M	9.433	4
185	90	M	25	
170	65	W	5	21
155	48	W	26.627	
160	55	W	18.027	
182	80	M	15.620	
175	69	W	5.099	2
180	80	M	14.142	
160	50	W	22.360	
175	72	M	5.385	3

If K=1, distance = 5, Gender = Male/woman

If K=3, distance → 1) 5, Gender = Man
2) 5, Gender = woman
3) 5.099, Gender = woman

$\Big\}$ Woman

If K=5, distance = 1) 5, Gender = Man
2) 5, Gender = woman
3) 5.099, Gender = woman
4) 5.385, Gender = Man
5) 9.433, Gender = Man

$\Big\}$ Man

(iii) Input → (175,70)

Height	Weight		Distance	
170	57	W	13.928	
192	95	M	30.232	
150	45	W	35.355	
170	65	M	7.071	3
175	78	M	8	5
185	90	M	22.360	
170	65	W	7.071	4
155	48	W	29.732	
160	55	W	21.213	
182	80	M	12.206	
175	69	W	1	1
180	80	M	11.180	
160	50	W	25	
175	72	M	2	2

If k=1, Distance = 1 , gender = woman

If k=3, Distance = 1) 1 , Gender = woman
2) 2 , Gender = man
3) 7.071 , Gender = Man/woman

Man/woman

If k=5, distance, 1) 1 , Gender = woman
2) 2 , Gender = Man
3) 7.071 , Gender = Man
4) 7.071 , Gender = woman
5) 8 , Gender = Man

Man

(iv) Input → (180, 90)

Height	Weight		Distance	
170	57	W	34.481	
192	95	M	13	5
150	45	W	54.083	
170	65	M	26.925	
175	78	M	13	4
185	90	M	5	1
170	65	W	26.925	
155	48	W	48.877	
160	55	W	40.311	
182 ✓	80	M	10.198	3
175	69	W	21.587	
180	80	M	10	2
160	50	W	44.721	
175	72	M	18.681	

If K=1, distance = 5, Gender = Man

If k=3, distance 1) 5, Gender = Man
2) 10, Gender = Man
3) 10.198, Gender = Man
} Man

If k=5, distance, 1) 5, Gender = Man
2) 10, Gender = Man
3) 10.198, Gender = Man
4) 13, Gender = Man
5) 13, Gender = Man
} Man

Question 2

Conclusion:-

Form the calculations and predictions seen in 2a and 2c we can clearly see that 2a provides a better prediction. With the age data present the prediction is more accurate.

If we remove the age data we can see that there are redundant predictions made which can either result in Man or Woman.

Q3

(a) Given: data from Q2

Probabilistic classifier

$$\text{posterior} = \frac{\text{prior} \times \text{likelihood}}{\text{evidence}}$$

$$P(C_k, x_1, \ldots x_n)$$

$$P(x_i \mid x_{i+1}, \ldots, x_n, C_k) = P(x_i \mid C_k)$$

$$= P(C_k) \prod_{i=1}^{n} P(x_i \mid C_k)$$

$$\therefore \hat{y} = \underset{K \in (1,\ldots K)}{\operatorname{argmax}} P(C_k) \prod_{i=1}^{n} P(x_i \mid C_k)$$

Gaussian Naive Bayes:-

$$P(x = v \mid C_k) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{(\cancel{x}-\mu)^2}{2\sigma^2}\right)}$$

Determine:-

$$\text{posterior(male)} = \frac{P(\text{male})\, p(\text{height} \mid \text{male})\, p(\text{weight} \mid \text{male})\, p(\text{age} \mid \text{male})}{\text{evidence}}$$

$$\text{posterior(woman)} = \frac{P(\text{woman})\, p(\text{height} \mid \text{woman})\, p(\text{weight} \mid \text{woman})\, P(\text{age} \mid \text{woman})}{\text{evidence}}$$

$p(\text{male}) = 0.5$

$P(\text{woman}) = 0.5$

Man :-

Mean(height) = 179.857
Var(height) = 53.809

Mean(Weight) = 80
Var(weight) = 103

Mean(age) = 30.142
Var(age) = 7.142

Woman :-

Mean(height) = 162.857
Var(height) = 82.142

Mean(weight) = 55.571
Var(weight) = 78.619

Mean(age) = 30
Var(age) = 2.333

(i) Input → (155, 40, 35)

$$P(\text{height}|\text{male}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(155-179.865)^2}{2\sigma^2}\right) = 0.00017$$

$$P(\text{weight}|\text{male}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(40-80)^2}{2\sigma^2}\right) = 1.664 \times 10^{-5}$$

$$P(\text{age}|\text{male}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(35-30.142)^2}{2\sigma^2}\right) = 0.02863$$

$$P(\text{height}|\text{woman}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(155-162.857)^2}{2\sigma^2}\right) = 0.03023$$

$$P(\text{weight}|\text{woman}) = \frac{1}{\sqrt{2\pi\sigma^2}} = \exp\left(-\frac{(40-55.57)^2}{2\sigma^2}\right) = 0.00962$$

$$P(\text{age}|\text{woman}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(35-30)^2}{2\sigma^2}\right) = 0.00123$$

posterior(Man) = P(M) P(Hight|m) P(weight|M) P(Age|M)

considering only the numerator

$\Rightarrow (0.5)(0.0017)(1.664 \times 10^{-5})(0.02863)$

$\Rightarrow \underline{4.1602 \times 10^{-11}}$

Posterior(W) = P(W) P(height|W) P(weight|W) P(Age|W)

considering the numerator:-

$\Rightarrow (0.5)(0.0302)(0.0096)(0.00123)$

$\Rightarrow \underline{1.7914 \times 10^{-7}}$

As the posterior(W) is greater than posterior(M) the prediction for (155, 40, 35) ⇒ woman.

(ii) input:- (170, 70, 32)

$\rightarrow P(\text{height}|M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(170-179.86)^2}{2\sigma^2}\right) = \underline{0.02205}$

$\rightarrow P(\text{weight}|M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(70-80)^2}{2\sigma^2}\right) = \underline{0.0241}$

$\rightarrow P(\text{age}|M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(32-30.14)^2}{2\sigma^2}\right) = \underline{0.11725}$

$\rightarrow P(\text{height}|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(170-162.857)^2}{2\sigma^2}\right) = \underline{0.03227}$

$\rightarrow P(\text{weight}|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(70-55.57)^2}{2\sigma^2}\right) = \underline{0.01197}$

$\rightarrow P(\text{age}|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(32-30)^2}{2\sigma^2}\right) = \underline{0.11083}$

Posterior (man) = $(0.5)(0.02205)(0.0241)(0.11725)$

$\Rightarrow \underline{0.00003127}$

Posterior (woman) = $(0.5)(0.032)(0.0119)(0.1108)$

$\Rightarrow \underline{0.000021406}$

since posterior(M) is greater than posterior(W)

given input (170, 70, 32) $\rightarrow$ Man

(iii) Input → (175, 70, 35)

$\rightarrow P(\text{height}|M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(175-179.86)^2}{2\sigma^2}\right) = \underline{0.04368}$

$\rightarrow P(\text{weight}|M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(70-80)^2}{2\sigma^2}\right) = \underline{0.0241}$

$\rightarrow P(\text{age}|M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(35-30.14)^2}{2\sigma^2}\right) = \underline{0.02863}$

$\rightarrow P(\text{height}|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(175-162.857)^2}{2\sigma^2}\right) = \underline{0.0179}$

$\rightarrow P(\text{weight}|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(70-55.57)^2}{2\sigma^2}\right) = \underline{0.01197}$

$\rightarrow P(\text{age}|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(35-30)^2}{2\sigma^2}\right) = \underline{0.00123}$

Posterior(man) = $0.5(0.0436)(0.0241)(0.0286)$

$\Rightarrow 1.512 \times 10^{-5} \Rightarrow \underline{0.00001512}$

Posterior(W) = $0.5(0.0179)(0.0119)(0.00123)$

$\Rightarrow 1.322 \times 10^{-7} \Rightarrow \underline{0.0000001322}$

Since posterior(w) < posterior(M)

(175, 70, 35) ⟶ man

Scanned by CamScanner

(iv) Input $\rightarrow$ (180, 90, 20)

$$\rightarrow P(height|M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(180-179.86)^2}{2\sigma^2}\right) = \underline{0.0543}$$

$$\rightarrow P(weight|M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(90-80)^2}{2\sigma^2}\right) = \underline{0.02419}$$

$$\rightarrow P(age|M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(20-30.14)^2}{2\sigma^2}\right) = \underline{0.00011}$$

$$\rightarrow P(height|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(180-162.857)^2}{2\sigma^2}\right) = \underline{0.00736}$$

$$\rightarrow P(weight|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(90-55.57)^2}{2\sigma^2}\right) = \underline{2.39 \times 10^{-5}}$$

$$\rightarrow P(age|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(20-30)^2}{2\sigma^2}\right) = \underline{1.3 \times 10^{-10}}$$

Posterior(man) = $0.5(0.0543)(0.0241)(0.00011)$

$\Rightarrow$ $\underline{7.319 \times 10^{-0}}$

Posterior(W) = $0.5(0.00736)(2.39 \times 10^{-5})(1.3 \times 10^{-10})$

$\Rightarrow$ $\underline{1.136 \times 10^{-11}}$

Since posterior(M) > posterior(W)

(180, 90, 20) $\rightarrow$ $\underline{\text{Man}}$

Q3

(c) Removing the age parameter.

(i) Input -> (155, 40)

$$P(\text{height}|m) = \frac{1}{\sqrt{(2\pi\sigma^2)}} \exp\left(-\frac{(155 - 179.86)^2}{2\sigma^2}\right) = 0.00017$$

$$P(\text{weight}|m) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(40 - 80)^2}{2\sigma^2}\right) = 1.664 \times 10^{-5}$$

$$P(\text{height}|w) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(155 - 162.857)^2}{2\sigma^2}\right) = 0.03023$$

$$P(\text{weight}|w) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(40 - 55.57)^2}{2\sigma^2}\right) = 0.00962$$

Posterior(m) = $0.5(0.00017)(1.664 \times 10^{-5})$

$\Rightarrow 1.4533 \times 10^{-9}$

Posterior(w) = $0.5(0.03023)(0.00962)$

$\Rightarrow 0.000145$

Since posterior(w) > posterior(M)
(155, 40) → Woman

(ii) Input → (170, 70)

$$P(\text{height}|M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(170-179.86)^2}{2\sigma^2}\right) = \underline{0.02205}$$

$$P(\text{weight}|M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(70-80)^2}{2\sigma^2}\right) = \underline{0.02419}$$

$$P(\text{height}|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(170-162.857)^2}{2\sigma^2}\right) = \underline{0.03227}$$

$$P(\text{weight}|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(70-55.57)^2}{2\sigma^2}\right) = \underline{0.01197}$$

Posterior(M) = 0.5(0.02205)(0.02419)

⇒ $\underline{0.0002667}$

Posterior(W) = 0.5(0.03227)(0.01197)

⇒ $\underline{0.000193}$

Since posterior(m) > posterior(W)

(170, 70) → <u>Man</u>

(iii) input → (175, 70)

$$P(\text{height}|\text{man}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(175-179.86)^2}{2\sigma^2}\right) = 0.04368$$

$$P(\text{weight}|m) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(70-80)^2}{2\sigma^2}\right) = 0.02419$$

$$P(\text{height}|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(175-162.857)^2}{2\sigma^2}\right) = 0.01794$$

$$P(\text{weight}|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(70-55.57)^2}{2\sigma^2}\right) = 0.01197$$

$$\text{Posterior}(m) = 0.5(0.04368)(0.02419)$$

$$\Rightarrow 5.2 \times 10^{-4}$$

$$\Rightarrow 0.000528$$

$$\text{Posterior}(W) = 0.5(0.0179)(0.01197)$$

$$\Rightarrow 1.07 \times 10^{-4}$$

$$\Rightarrow 0.000107$$

Since posterior(m) > posterior(w)

(175, 70) → Man

(iv) Input $\rightarrow$ (180, 90)

$$P(\text{height}|M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(180-179.86)^2}{2\sigma^2}\right) = \underline{0.05437}$$

$$P(\text{weight}|M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(90-80)^2}{2\sigma^2}\right) = \underline{\underline{0.02419}}$$

$$P(\text{height}|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(180-162.857)^2}{2\sigma^2}\right) = \underline{0.00736}$$

$$P(\text{weight}|W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(90-55.57)^2}{2\sigma^2}\right) = \underline{2.39 \times 10^{-5}}$$

$$\text{Posterior}(M) = 0.5(0.05437)(0.02419)$$

$$\Rightarrow 6.5 \times 10^{-4}$$

$$\Rightarrow \underline{\underline{0.000657}}$$

$$\text{Posterior}(W) = 0.5(0.00736)(2.39 \times 10^{-5})$$

$$\Rightarrow \underline{\underline{8.809 \times 10^{-8}}}$$

Since posterior(M) > posterior(W)

(180, 90) $\rightarrow$ <u>Man</u>

Scanned by CamScanner

Q3d)

Conclusions:-

Based on the 2 classifier test performed in 2b,c and 3b,c we can see that there are both pro's an cons in both the classifiers.
in KNN the age parameter removed make redundant predictions which are not justifiable. Similarly in Naive Bayes classifiers removing the age parameter causes changes in the probabilistic values and will result in inconclusive predictions.
KNN classifiers assign category based on selected nearest neighbours.
The Naive Bayes classifier chooses the probability based on the prior knowledge of the value & its probability.
Hence Naive Bayes are comparitively better than KNN considering large data sets.