Q1 Abhishek Mod - 1001558833

M.L

a) Probability of the first event to occur at time x after restart $= P(S)$ is the performance measure

$$P(D \mid P(S)) = \lambda e^{-\lambda x}$$ ..... described by exponential distribution

~~for~~ Analytic MLE optimization

- Convert to log likelihood

$$\hat{u} = \ln \sum P(D \mid P(S)) = \sum_{i=1}^{n} \ln\left(\lambda e^{-\lambda x_i}\right)$$

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

Taking derivative w.r.t λ

$$\frac{\partial}{\partial \lambda}\left(n \ln \lambda - \lambda \sum_{i=1}^{n} x_i\right) = 0$$

$$\frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow \hat{\lambda} = \frac{n}{\sum x_i}$$

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

$\Rightarrow n = 6$

$\sum x_i = 1.5 + 3 + 2.5 + 2.75 + 2.9 + 3$

$= 15.65$

$$\hat{\lambda} = \frac{n}{\sum x_i} = \frac{6}{15.65} = 0.3833$$

Q1 c

$$P(\lambda|x) \propto P(x|\lambda)P(\lambda)$$
$$\propto \lambda^n e^{-\lambda \Sigma x_i} \lambda^{\alpha-1} e^{-\beta\lambda}$$
$$\propto e^{-\lambda(\Sigma x_i - \beta)} \lambda^{n+\alpha-1}$$

$$\Rightarrow P(\lambda|x) \propto Gamma(\alpha+n, \Sigma x_i + \beta)$$

Take log

$$\log P(\lambda|x) \propto -\lambda(\sum_i x_i + \beta) + (n+\alpha-1)\log\lambda$$

Take derivate w.r.t λ & set it $=0$.

$$\therefore -\Sigma x_i - \beta + \frac{n+\alpha-1}{\lambda} = 0$$

$$\therefore \lambda = \frac{n+\alpha-1}{\Sigma_i x_i + \beta}$$

$\alpha = 5$ & $\beta = 10$

$$\Rightarrow \lambda = \frac{6+5-1}{15.65+10} = 0.3898$$

Q2a)

	Height	Weight	Age	Cartesian Distance for Point 1	Point 2	Point 3	Point 4
W	170	57	32	22.869	13.0	14.247	36.510
M	192	95	28	66.655	33.541	31.032	15.264
W	150	45	30	8.66	32.078	35.707	55.0
M	170	65	29	29.765	5.830	9.273	28.390
M	175	78	35	42.941	9.899	8.0	19.849
M	185	90	32	58.386	25.0	22.561	13.0
W	170	65	28	29.983	6.403	9.949	28.089
W	155	48	31	8.944	26.645	30.0	50.099
W	160	55	30	16.583	18.138	21.794	41.533
M	182	80	30	48.518	15.748	13.190	14.282
W	175	69	28	35.916	6.480	7.071	23.021
M	180	80	27	47.843	15.0	13.747	12.206
W	160	50	31	11.874	22.383	25.317	46.054
M	175	72	30	38.065	5.744	5.385	21.189

Cartesian Dist^n formula $= \sqrt{(\text{Height} - \text{Test.Height})^2 + (\text{Weight} - \text{TestWeight})^2 + (\text{Age} - \text{TestAge})^2}$

for K=1

Select the data which has the minimum cartesian distance

Point 1 → 8.66 → [(150, 45, 30), 'W']

Point 2 → 5.744 → [(175, 72, 30), 'M']

Point 3 → 5.385 → [(175, 72, 30), 'M']

Point 4 → 12.206 → [(180, 80, 27), 'M']

Return Majority label

∴ Point 1 → W

Point 2 → M

Point 3 → M

Point 4 → M

for K = 3

point 1 → 3 nearest → [(150, 45, 30), 'W'], [(155, 48, 31), 'W'], [(150, 60, 31), 'W']

point 2 → 3 nearest → [(175, 72, 30), 'M'], [(170, 65, 29), 'M'], [(175, 69, 28), 'W']

Similarly for Test point 3 & 4

- Return Majority label for points

∴ Test point 1 → W

Test point 2 → M

Similarly for Test point 3 & Test point 4 → M

Repeat same (steps) for K = 5

Q3a

$P(\text{height}|W) = N(\text{height}|\mu,\sigma) = 0.03022$
$P(\text{weight}|W) = N(\text{weight}|\mu,\sigma) = 0.00962$
$P(\text{Age}|W) = N(\text{Age}|\mu,\sigma) = 0.0012$

$P(\text{height}|M) = N(\text{height}|\mu,\sigma) = 0.00017$
$P(\text{weight}|M) = N(\text{weight}|\mu,\sigma) = 0.000001$
$P(\text{Age}|M) = N(\text{Age}|\mu,\sigma) = 0.0286$

$$N(z|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\text{Mean} = \hat{\mu}_{i,j} = E[x_i|c_j] = \frac{\sum_{k:y^{(k)}=c_j} x_i^{(k)}}{\#(x^{(k)},y^{(k)}): y^{(k)}=c_j}$$

$$\hat{\sigma}^2_{i,j} = \frac{\sum_{k:y^{(k)}=c_j} (x_i^{(k)}-\mu_{i,j})^2}{(\#(x^{(k)},y^{(k)}): y^{(k)}=c_j) - 1}$$

$$\hat{c} = \operatorname{argmax} P(c) \prod_i P(x_i|c)$$

Scanned by CamScanner

	Column1.	Column2	Column3.1	Colum		Column1	Column2	Column3	Colum
	170	57	32	W		192	95	28	M
	150	45	30	W		170	65	29	M
	170	65	28	W		175	78	35	M
	155	48	31	W		185	90	32	M
	160	55	30	W		182	80	30	M
	160	50	31	W		180	80	27	M
	175	69	28	W		175	72	30	M
mean	**162.85714**	**55.57142857**	**30**			**179.85714**	**80**	**30.142857**	
stdev	9.0632697	8.866738274	1.527525232			7.3354975	10.148892	2.6726124	

Q3d

- I think KNN is better than Gaussian Naive Bayes classifier, because Gaussian Naive Bayes makes one crucial assumption -> that the features of the input data are independent from each other; which can be a huge drawback in some cases.
- Also KNN is quite simple to implement

Scanned by CamScanner