Machine Learning.

1) a. performance function,

prob of event to occur b/w x & $\lambda e^{-\lambda x}$ interval with rate parameter λ.

$P_x(x) = \lambda e^{-\lambda x}$.

$f(x) = \lambda e^{-\lambda x}$, continous random variable.

MLE,

analytic optimization,

(i) $\frac{d f(x)}{dx} = 0$.

$P(\lambda|D) = \overbrace{P(D|\lambda)}^{MLE} P(\lambda)$

$P(D|\lambda) = \prod_{i=1}^{N} \lambda e^{-\lambda x_i}$

$= \lambda \prod_{i=1}^{N} e^{-\lambda x_i}$

$= \lambda e^{-\lambda \prod_{i=1}^{N} x_i}$

to find λ.

for deriving $\frac{\partial(\lambda e^{-\lambda x})}{\partial \lambda}$, $\text{argmax}_\lambda (\lambda e^{-\lambda x})$

Apply Log likelihood

$= \sum_{i=1}^{N} (\log(\lambda e^{-\lambda x_i}))$

$= \sum_{i=1}^{N} -\lambda x_i + \log \lambda$

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

$n \log \lambda - \lambda \sum_i x_i$

$= n * \frac{1}{\lambda} - \sum_i x_i$

Set derivative to '0'

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

$\lambda = \frac{1}{\bar{x}}$ $\bar{x}$, sample mean $= \frac{\sum x}{n}$

i.e mean of dataset.

1b) Data from link,

$\sum x = 2.1193.. + 3.225.. + 0.822.. + 0.94.. + 0.02.. + 0.143$

$\sum x = 7.274204679$, $n = 6$, $\bar{x} = 1.212367447$

$\lambda = \underline{\underline{0.8248324404}}$.

1c) MAP approach using conjugate prior.

$x_i \sim poisson(\lambda)$ with gamma (α, β) prior distribution.
leads to a posterior distribution

$\lambda | D \sim Gamma(\alpha, \beta)$

$P_{\alpha,\beta}(\lambda) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$.

MAP

$P(\mu|D) = \frac{P(D|\mu)P(\mu)}{P(D)}$

$P(\lambda) = \text{argmax}_\lambda \, P(D|\mu)P(\mu)$

$$P(D|\lambda)\, P(\lambda)$$

$$P(D|\lambda) = \prod_{i=1}^{N} \lambda e^{-\lambda x_i}$$

$$= \left(\prod_{i=1}^{N} \lambda e^{-\lambda x_i}\right)\left(\frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}\right)$$

α, β are constants.

$\therefore$ log likelihood,

$$= \log\left(\lambda^{n+\alpha-1} e^{-\lambda(\sum_i x_i - \beta)}\right)$$

$$= -\lambda \sum_i x_i + \beta + (n+\alpha-1)\log\lambda$$

Set it equal to 0,

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

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

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

$\sum x_i = 7.274204679.$

$\alpha = 4$

$\beta = 11$

$\lambda \approx 0.4924$

3a)

$$p(x|c) = \frac{\#(x,y) \in D : y = c \wedge x = a}{\#(x,y) \in D,\ y = c}$$

$p(\text{height} \mid w) =$

$p(w \mid \text{height}, \text{weight}, \text{age})$

$= p(w) \cdot p(w \mid \text{height}) \cdot p(w \mid \text{weight})\, p(w \mid \text{age})$

* mean of each feature.

mean (height) = 1.65714027

mean (weight) = 72.64744

mean (age) = 30.57

= 31.

Std (height) = 0.12365444 84

Std (weight) = 6.40043619

Std (age) = 3.71513093.

prior - men = 7/14 = 0.5

prior - women = 7/14 = 0.5

mean (women_height) = 1.56417186.

mean (women - weight) = 68.1

mean (women - age) = 29

mean (~~wo~~men - height) = 1.75010868 6.

mean (~~wo~~men - weight) = 77.194374

mean (~~wo~~men - age) = 31.

$$P(y|x) = \frac{P(y)\,[P(x_1|y).\,P(x_2|y)\ldots]}{P(y_1).P(x|y_1) + P(y_2)P(x|y_2)}$$

Let x_1 = height $\quad y_1$ = Male | men.

x_2 = weight $\quad y_2$ = women

x_3 = age

$$p(\text{men}) = p(\text{women}) = \Sigma = \underline{1}.$$

$$0.5 + 0.5 = \underline{1}$$

$$p(\text{male}|x) = \frac{p(\text{men})\,[p(h|m).p(W|m)\,p(a|m)]}{p(\text{men})\,[p(h|m).p(w|m).\,p(a|m)] + p(\text{women})\,[p(H|w)\,p(w|w)\,p(a|w)]}$$

posterior

||y for $p(\text{women}|x)$ =

denominator is constant for both

$\therefore$ we only calculate $p(y)\left[\prod_{i=1}^{n} p(x_i|y)\right]$

height and weight are continous, Age is discrete

$\therefore$ we calculat prob Density function.

	Mean(h)	mean(w)	mean(a)	Var(h)	Var(w)	var(a)
Men	1.57	68.1	29	0.011445	14.821	15.619
Women	1.75	77.195	31	0.001516	25.6969	12.5714

$$\text{Sample var} = \frac{\Sigma (x - \bar{x})^2}{n-1}$$

$$P(H|M) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(d-\mu)^2}{2\sigma^2}\right)$$

test_data for Example.

H = 1.796918..

W = 71.1174611..

a = 31

$$P(H|m) = \frac{1}{\sqrt{2(3.14)(0.01145)}} e^{\left(-\frac{(1.796918 - 1.57)^2}{2(0.01145)}\right)} \approx 0.39333$$

$$P(W|m) = \frac{1}{\sqrt{2(3.14)(14.821)}} e^{\left(-\frac{(71.1174611 - 68.1)^2}{2(14.821)}\right)} \approx 0.07624$$

$$P(a|m) = \frac{1}{\sqrt{2(3.14)(15.619)}} e^{\left(-\frac{(31-29)^2}{2(15.619)}\right)} \approx 0.87985$$

$$P(H|w) = \frac{1}{\sqrt{2(3.14)(0.001516)}} e^{\left(-\frac{(1.796918 - 1.75)^2}{2(0.001516)}\right)} \approx 4.9587$$

$$P(W|w) = \frac{1}{\sqrt{2(3.14)(25.6969)}} e^{\left(-\frac{(71.1174611 - 77.195)^2}{2(25.6969)}\right)} \approx 0.03837$$

$$P(w|a) = \frac{1}{\sqrt{2(3.14)(12.5714)}} e^{\left(\frac{-(31-31)^2}{2(12.5714)}\right)} \approx 0.112545$$

∴ Since 2 values are greater for ~~women~~ men than woman it is woman.

= Compare posterior men & women.

$$P(m|x) = P(m)\,[P(H|m) \cdot P(W|m) \cdot P(a|m)]$$

$$\approx 0.0132 \, .$$

$$P(w|x) = P(w)\,[P(H|m)\,P(W|w)\,P(a|w)]$$

$$= 0.0107$$

$$P(m|x) > P(w|x)$$

$\therefore$ The given data is of ~~__~~ man class.

Ily we can calulate for other test data.

3D)

$\therefore$ The knn alg for leave_one_out gives better performance for $k=1$.

It algorithm, knn also gives better result when age is removed from the data.

$\Rightarrow$ height and weight are compact in the plane.

Ily.
The o/p for naive bayes for leave_one_out performs better.

It performs slightly better than knn_alg.