Probability

Probability mass function:

$P_x(x_k) = P(X = x_k)$, for $k = 1, 2, 3, \ldots$

$P_x(x_0) = \sum_1 x f(x)$

Binomial Distribution:

$P_x(k) = {}^nC_k\, p^k (1-p)^{n-k}$

mean = np

varience = npq

Poission Distribution:

$P_x(k) = \frac{e^{-\lambda} \lambda^k}{k!}$

$\lambda = np$ = mean = varience

Cummulative Distribution function (CDF)

$P(a < x \leq b) = F(b) - F(a) = \int_a^b f(x)\,dx$

Probability Density function:

$f(x) = F'(x) = \frac{dF(x)}{dx}$

$F(x)$ = Cummulative distributive function

$\int_{-\infty}^{\infty} f(x)\,dx = 1$

$f(x) \geq 0$

Mean:

$E(x) = \mu = \int_{-\infty}^{\infty} x f(x)\,dx$

Varience:

$Var(x) = \sigma^2 = \left(E(x^2) - (E(x))^2\right)$

Baye's Theorem

$P(h|D) = \frac{P(D|h)\,P(h)}{P(D)}$

Principle Component Analysis

	Example 1	2	3	4
x1	4	8	13	7
x2	11	4	5	14

Step 1: Calculate Mean

$\bar{x_1} = (4+8+13+7)/4 = 8$

$\bar{x_2} = (11+4+5+14)/4 = 8.5$

Step 2: Calculation of the covariance matrix

$$covar(x1,x1) = \frac{1}{N-1}\sum_{k=1}^{N}(x_{1k}-\bar{x}_1)^2$$

$$= \frac{1}{3}\left((4-8)^2+(8-8)^2+(13-8)^2+(7-8)^2\right)$$

$$= +14$$

$$covar(x1,x2) = \frac{1}{N-1}\sum_{k=1}^{N}(x_{1k}-\bar{x}_1)(x_{2k}-\bar{x}_2)$$

$$= \frac{1}{3}\left((4-8)(11-8.5)+(8-8)(4-8.5)+(13-8)(5-8.5)+(7-8)(14-8.5)\right)$$

$$= -11$$

$$covar(x_2,x_1) = covar(x_1,x_2) = -11$$

$$covar(x_2,x_2) = \frac{1}{N-1}\sum_{k=1}^{N}(x_{2k}-\bar{x}_2)^2$$

$$= \frac{1}{3}\left((11-8.5)^2+(4-8.5)^2+(5-8.5)^2+(14-8.5)^2\right) = 23$$

Step: $S = \begin{bmatrix} cov(a,a) & cov(a,b) \\ cov(b,a) & cov(b,b) \end{bmatrix} = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$

Step-4: calculate eigin values, eign vectors, Normalized eign vectors.

$det(S-\lambda I) = 0$

$$det\left(\begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = 0 \Rightarrow \begin{bmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{bmatrix} = 0$$

$(14-\lambda)(23-\lambda) - (-11)(-11) = 0$

$\Rightarrow \lambda^2 - 37\lambda + 201 = 0$

$\therefore \lambda_1 = 30.38 \quad \lambda_2 = 6.62 \quad (\text{max of } \lambda_1, \lambda_2)$

vectors:

$$\begin{bmatrix} (14-\lambda_1) & -11 \\ -11 & (23-\lambda_1) \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$(14-\lambda_1) U_1 + \cancel{(23-\lambda_1) U_2} - 11 U_2 = 0 \quad -(a)$

$-11 U_1 + (23-\lambda_1) U_2 = 0 \quad -(b)$

from (a)

$(14-\lambda_1) U_1 + (-11) U_2 = 0$

$(14-\lambda_1) U_1 = 11 U_2 \Rightarrow \frac{U_1}{11} = \frac{U_2}{14-\lambda_1}$

let $\frac{U_1}{11} = \frac{U_2}{14-\lambda_1} = t = 1 \Rightarrow U_1 = 11 \;;\; U_2 = 14-\lambda_1 = -16.38$

$\therefore$ eigen vector for $\lambda_1 = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = \begin{bmatrix} 11 \\ -16.38 \end{bmatrix}$

normalized eign vector:

$$n_1 = \begin{bmatrix} 11/\sqrt{11^2+16.38^2} \\ -16.38/\sqrt{11^2+16.38^2} \end{bmatrix} = \begin{bmatrix} 11/19.73 \\ -16.38/19.73 \end{bmatrix} =$$

$$= \begin{bmatrix} 0.5575 \\ -0.8302 \end{bmatrix}$$

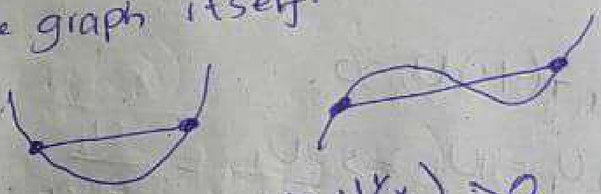
step-5 $\quad PC1 = n_1^T \begin{bmatrix} x_1 - \bar{x}_1 \\ x_2 - \bar{x}_2 \end{bmatrix}$

$\therefore PC1 = -4.305 \quad PC2 = 3.735 \quad PC_3 = 5.694$

$PC4 = -5.123$

Convex optimization

* Convex optimization is a powerful tool for solving optimization problems in various fields such as finance, engineering and machine learning.
* In a convex optimization problem, the goal is to find a point that maximizes the objective function.
* Linear functions are convex, so linear programming problems are convex problems.
* A convex function is a function whose graph is always curved upwards, which mean the the line segment connecting any two points of graph is always above or on the graph itself.



$f(x) \Rightarrow f'(x) = 0 \quad \Rightarrow f'(x) > 0$

Gradient-Based optimization

for $f(x)$ is Not given

$$x_n = x_{n-1} - \alpha * d(f(x_{n-1})) \qquad \alpha - \text{Learning rate}$$

$$D_m = \frac{-2}{n} \cdot \sum x_i (y_i - \bar{y})$$

$$D_c = -\frac{2}{n} \sum (y_i - \bar{y})$$

$m_0 = 0 \quad c_0 = 0 \quad \bar{y} = 0 \quad \alpha = 0.01$ (default)

x 1 2 3 4

y 2 3 4 5

x	y	$y-\bar{y}$	$x(y-\bar{y})$
1	2	2	2
2	3	3	6
3	4	4	12
4	5	5	20

$m_1 = m_0 - \alpha D_{m_0}$

$m_1 = 0 - (0.01) * \left(\frac{-2}{4}(40)\right) = \cancel{20}0.2$

$c_1 = 0 - (0.01) * \left(\frac{-2}{4} \times 14\right) = 0.07$

$\therefore \bar{y} = m\bar{x} + c$

$\bar{y} = 0.2x + 0.07$

x	y	$\bar{y}$	$y-\bar{y}$	$x(y-\bar{y})$
1	2	0.27	1.73	1.73
2	3	0.47	2.53	5.06
3	4	0.67	3.33	9.99
4	5	0.87	4.13	16.52
			11.72	33.3

1st itration

$y = 0.3665x + 0.1286$

$m_2 = 0.2 - (0.01) * \left(\frac{-2}{4} * 33.3\right)$

$= 0.2 + 0.1665$

$= 0.3665$

$c_2 = c_1 - (0.01) * \left(\frac{-2}{4} * 11.72\right)$

$= 0.07 +$

$= 0.1286$

x	y	$\bar{y}$	$y-\bar{y}$	$x(y-\bar{y})$

2nd itration.

Cor is given

$f(x) = x^2 - 5x + 15$

$x_0 = 3 \quad \alpha = 0.25$

$\therefore x_n = x_{n-1} - \alpha\,(d(f(x_{n-1})))$

$d(f(x_0)) = 2x - 5 = 2(3) - 5 = 1$

$\therefore x_1 = 3 - (0.25)(1)$

$= 3 - 0.25$

$= 2.75$

$\therefore x_2 = x_1 - \alpha\,(df(x_1))$

$= 2.75 - 0.25\,(2.75 \times 2 - 5)$

$= 2.75 - 0.25 \times 0.5$

$x_2 = 2.625$

$\therefore x_3 = x_2 - \alpha\,(df(x_2))$

$= 2.625 - 0.25 \times 0.25$

$= 2.625 - 0.625$

$x_3 = 2$

$df(x_2)$

$2(2.625) - 5$

$= 0.25$

logistic regression

$s(x) = \dfrac{1}{1+e^{-x}}$

Calculate the probability of pass for the student who studied 33 hours.

Hours study	Pass(1) / Fail(0)
29	0
15	0
33	1
28	1
39	1

$\log(\text{odds}) = z = -64 + 2 * \text{hours}$

(i) $z = -64 + 2 * 33 = -64 + 66 = 2$

$$p = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-2}} = 0.88$$

if the studies 33 hours, then there is 88% chance.

(ii) At least how many hours students should study that make he will pass the course with the p more than 95%...

$$p = \frac{1}{1+e^{-z}} = 0.95$$

$$\ln(e^{-z}) = \ln\left(\frac{1}{0.95} - 1\right) \Rightarrow z = 2.94$$

$\log(\text{odds}) = z = -64 + 2 * \text{hours}$

$z = 2.94 \Rightarrow$

$$\frac{z+64}{2} = \text{hours} = 33.47 \text{ Hours}$$

---

x is the $\rightarrow z = mx + c \Rightarrow p = \dfrac{1}{1+e^{-z}}$ (remember "$-$")

p is the $\Rightarrow p = \dfrac{1}{1+e^{-z}} \Rightarrow$ find z $\Rightarrow z = mx + c$ find x

KNN

~~Step-1~~ Gives one table dataset, and asks to find the value of a last coloumn based on Given other values.

Step-1: Find Distance

$\text{Distance} = \sqrt{(x-a)^2 + (y-b)^2}$

a = Given data sets values
b = Given data sets values
x = finding values
y = finding values.

find all distances from Given values.

Step-2 Find Rank

Rank the ~~val~~ distance by least is 1. (Sort in ascending)

Step-3 find the nearest neighbors.

$K = \sqrt{\text{total No. of tuple}}$

if K = even
then +1 or -1
should be done

~~Decision~~
LDA

Step-1: Compute the class Means of dependent variables

$\mu_1 = \frac{1}{N_1} \sum_{x \in \omega_1} x \quad (\mu_2)$

Step-2 Derive the Covarience matrix of the class variable.

$S_1 = \sum_{x \in \omega_1} \frac{(x-\mu_1)(x-\mu_1)^T}{N-1} \quad (S_2)$

Step-3 Compute the within class - scatter matrix (S_W) $(S_1 + S_2)$

Step-4 Compute the between Class scatter matrix

$S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$

es.

ues

kers

ding

one

ies.

ble.

$+S_2)$

steps
Compute the Eigen values and eigen vectors from the within class and between class slatter matrix

$$S_W^{-1} S_B W = \lambda W$$

step6 sort the values of eigen values and select the top K values.

step7 Find the eigen vectors corresponding to the top K eigen values.

$$\left[S_W^{-1} S_B - \lambda I\right] \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0$$

step8 obtain the LDA by taking the dot product of eigen vectors & original data.

Q) $X_1 = \{(4,2), (2,4), (2,3), (3,6), (4,4)\}$

$X_2 = \{(9,10), (6,8), (9,5), (8,7), (10,8)\}$

sol)

$$\mu_1 = \frac{1}{N_1} \sum_{x \in \omega_1} x = \frac{1}{5}\left[\begin{pmatrix}4\\2\end{pmatrix} + \begin{pmatrix}2\\4\end{pmatrix} + \begin{pmatrix}2\\3\end{pmatrix} + \begin{pmatrix}3\\6\end{pmatrix} + \begin{pmatrix}4\\4\end{pmatrix}\right]$$

$$= \begin{pmatrix}3\\3.8\end{pmatrix}$$

$$\mu_2 = \frac{1}{N_2} \sum_{x \in \omega_2} x = \frac{1}{5}\left[\begin{pmatrix}9\\10\end{pmatrix} + \begin{pmatrix}6\\8\end{pmatrix} + \begin{pmatrix}9\\5\end{pmatrix} + \begin{pmatrix}8\\7\end{pmatrix} + \begin{pmatrix}10\\8\end{pmatrix}\right] = \begin{pmatrix}8.4\\7.6\end{pmatrix}$$

step-2

S_1

(i) Covariance: $(x - \mu_1) = \begin{bmatrix} 1 & -1 & -1 & 0 & 1 \\ -1.8 & 0.4 & -0.6 & 2.4 & 0.4 \end{bmatrix}$

$$S_1 = \frac{\sum_1 (x-\mu_1)(x-\mu_1)^T}{N-1}$$

$$= \left[\begin{pmatrix}4\\2\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]\left[\begin{pmatrix}4\\2\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]^T + \left[\begin{pmatrix}2\\4\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]\left[\begin{pmatrix}2\\4\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]^T$$

$$+ \cdots$$

$$\cancel{N-1}\ 5-1$$

$$= \begin{pmatrix}1 & -0.25\\ -0.25 & 2.2\end{pmatrix}$$

$$\boxed{\begin{aligned} &\text{(i)}\ \begin{pmatrix}4\\2\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix} = \begin{pmatrix}1\\-1.8\end{pmatrix} \\ &\begin{pmatrix}1\\-1.8\end{pmatrix}\begin{pmatrix}1 & -1.8\end{pmatrix} = \begin{bmatrix}1 & -1.8\\ -1.8 & 3.24\end{bmatrix}\end{aligned}}$$

Same for

$$S_2 = \frac{\sum_i (x-\mu_2)(x-\mu_2)^T}{N-1} = \begin{pmatrix}2.3 & -0.05\\ -0.05 & 3.3\end{pmatrix}$$

$$\therefore S_W = S_1 + S_2 = \begin{pmatrix}1 & -0.25\\ -0.25 & 2.2\end{pmatrix} + \begin{pmatrix}2.3 & -0.05\\ -0.05 & 3.3\end{pmatrix}$$

$$= \begin{pmatrix}3.3 & -0.3\\ -0.3 & 5.5\end{pmatrix}$$

$$S_B = (\mu_1-\mu_2)(\mu_1-\mu_2)^T = \left[\begin{pmatrix}3\\3.8\end{pmatrix} - \begin{pmatrix}8.4\\7.6\end{pmatrix}\right]\left[\begin{pmatrix}3\\3.8\end{pmatrix} - \begin{pmatrix}8.4\\7.6\end{pmatrix}\right]^T$$

$$= \begin{pmatrix}-5.4\\-3.8\end{pmatrix}\begin{pmatrix}-5.4 & -3.8\end{pmatrix} = \begin{bmatrix}29.16 & 20.52\\ 20.52 & 14.44\end{bmatrix}$$

find eignvalues.

$S_\omega^{-1} S_B \omega = \lambda\omega$

$|S_\omega^{-1} S_B \omega - \lambda I| = 0$

$$\left| \begin{pmatrix} 3.3 & -0.3 \\ -0.3 & 5.5 \end{pmatrix}^{-1} \begin{pmatrix} 29.16 & 20.52 \\ 20.52 & 14.44 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = 0$$

$$\left| \begin{pmatrix} 0.3045 & 0.0166 \\ 0.0166 & 0.1822 \end{pmatrix} \begin{pmatrix} 29.16 & 20.52 \\ 20.52 & 14.44 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right| = 0$$

$$\left| \begin{pmatrix} 9.2213-\lambda & 6.489 \\ 4.2339 & 2.9794-\lambda \end{pmatrix} \right|$$

$\Rightarrow (9.2213-\lambda)(2.9794-\lambda) - 6.489 \times 4.2339 = 0$

$\Rightarrow \lambda^2 - 12.2007\lambda = 0 \Rightarrow \lambda(\lambda - 12.2007) = 0$

$\therefore \lambda_1 = 0 \quad \lambda_2 = 12.2007$

$$\left(S_\omega^{-1} S_B - \lambda I\right) \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = 0$$

$\omega_1 = \begin{pmatrix} -0.5755 \\ 0.8178 \end{pmatrix}$ $(\lambda = 0)$

$\omega_2 = \begin{pmatrix} 0.9088 \\ 0.4173 \end{pmatrix} = \omega^*$ $(\lambda = 12.2007)$

$$w^* = S_w^{-1}(\mu_1 - \mu_2) = \begin{pmatrix} 3.3 & -0.3 \\ -0.3 & 5.5 \end{pmatrix}^{-1} \begin{bmatrix} -5.4 \\ -3.8 \end{bmatrix}$$

$$= \begin{pmatrix} 0.9088 \\ 0.4173 \end{pmatrix}$$

---

ada boost

step find weight = 1/total no of.

(1) select one attri

(a) $error = \sum_{j=1}^{N} H_1(d_j) \; wt(d_j)$

d_j = correct = 0

d_j = incrc = 1 $\boxed{\text{error} = \text{No. of Incoret} \times \text{weigh}}$

(b) $\underset{\alpha}{\text{weight (new)}} = \frac{1}{2} \log \frac{(1 - \varepsilon_{capa})}{\varepsilon_{capa}}$

(c) Normalizing factor Z = correct weigh * No. of correct $\times e^{-\alpha}$ + Incor. weigh * No. of incrc * $e^{+\alpha}$

(d) updating new weights

$$\text{new wt}(1) = \frac{\text{old correct weight} * e^{-\alpha}}{Z}$$

$$= \frac{\text{old incort weigh} * e^{\alpha}}{Z}$$

SVM

linear & Non-linear.
used to solve classification & Regression

$f(\vec{x}) = \vec{w} \cdot \vec{x} - b$

$\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$

$\sum_{i=1}^{N} \alpha_i y_i = -\alpha_1 + \alpha_2 + \alpha_3 = 0$

$\phi(\vec{\alpha}) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1, j=1}^{N} \alpha_i \alpha_j y_i y_j (\vec{x_i} \cdot \vec{x_j})$

=

Introduction to ANN (Artificial neuro network)

1) Introduction
2) Basic Architecture
3) Single perception
4) m
5) Build

$x_1 = 1$ x_2 x_3 $x_4 = 1$

Bias = 1

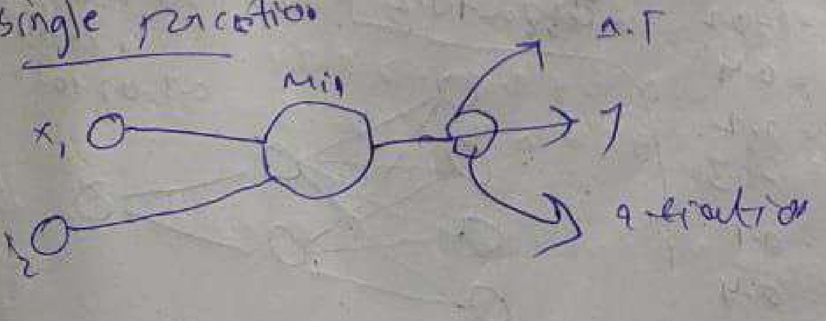
target = 1

$w_1 = w_2 = w_3 = w_4 = 1$

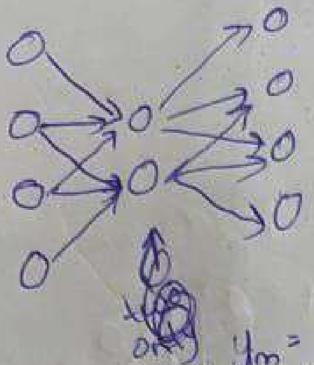
Calc

$y = x_1 \; x_2 +$

6) single perception



Multi-layer



① calculate y_{in}
② ~~calculate~~ apply activation fn
③ check y ~~o~~ ≠ P
④ upgrade.

$$y_{in} = \sum_{i=1}^{n} x_i w_{ij} + b_j$$

$$A.F = \frac{1}{1+e^{-y_{in}}}$$

$T = 0.5$

$x_1 = 0.35$

$x_2 = 0.9$

bias $= 0$

$0.1 = w_{11}$
$0.4 = w_{12}$
$0.8 = w_{21}$
$0.6 = w_{22}$
$0.3 = w_{31}$
$0.9 = w_{32}$

$y_{in} = 0.35 \times 0.1 + 0.9 \times 0.8$

$0.035 + 0.72$

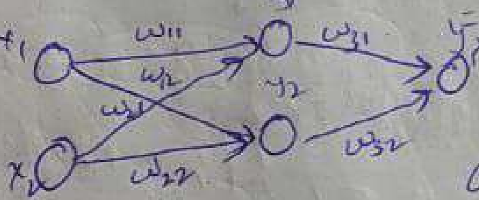
0.755

$y_1 = 0.68$

$y_{in} = 0.35 \times 0.4 + 0.9 \times 0.6$

0.68

$= 0.663$



~~0.755 × 0.3 + 0.68 × 0.9~~

$y_{in_3} = 0.8013$

$y_3 = 0.690$

$T = 0.5$

$P = 0.69$

$\therefore$ Error $= 0.5 - 0.69 = -0.19$

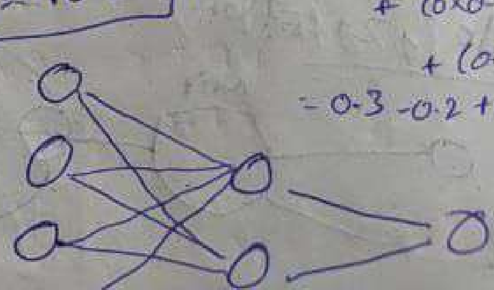
$x_1 = 1$
$x_2 = 1$
$x_3 = 0$
$x_4 = 1$

T =

$w_{11} = 0.3$
$w_{12} = 0.1$
$w_{21} = -0.2$
$w_{22} = 0.4$
$w_{31} = 0.2$
$w_{32} = -0.3$
$w_{41} = 0.1$
$w_{42} = 0.4$

$w_{51} = -0.3$
$w_{52} = +0.2$

$y_{in_1} = 1 \times 0.3 + 1 \times (-0.2) + (0 \times 0.2) + (0.1 \times 1)$
$= 0.3 - 0.2 + 0.1 = 0.2$



$y_1 = \frac{1}{1+e^{0.2}} = 0.5498$

$= 0.1\times1 + 0.4\times1 + (-0.3\times0) + 0.4\times1 = 0.1 + 0.4 + 0.4 = 0.9$

$= \frac{1}{1+e^{-0.9}} = 0.711$

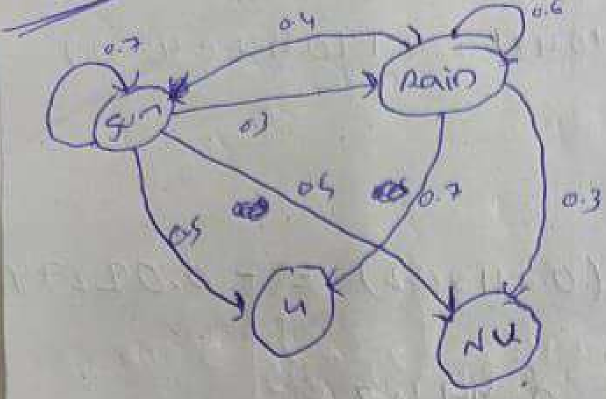
$= 0.5498 \times (-0.3) + (0.711\times0.2) = -0.02274$

$y_3 = \frac{1}{1+e^{+0.02274}} = 0.4943$

$\therefore E = 0.5 - 0.494 = 0.005684755$

$= 0.006$

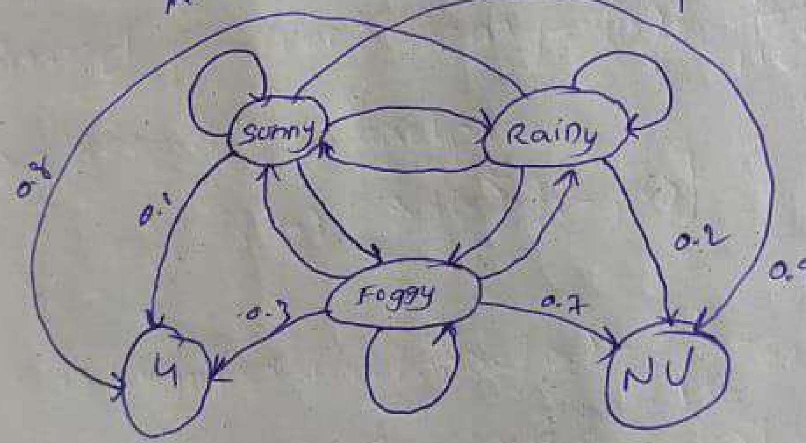
HMM



(Q) Sunny

	S	R	F
S	0.8	0.05	0.5
R	0.2	0.6	0.2
F	0.2	0.3	0.5

weather	P(U)	P(NU)
S	0.1	0.9
R	0.8	0.2
F	0.3	0.7



- Suppose a day you are looked was sunny. The next day the care taker carried an umbrella into the room. He would like to know that the weather was like this second day.

$$P(S|U) * P(S) = 0.8 \times 0.10 \Rightarrow 0.08$$
$$= 0.05 \times 0.8 \Rightarrow 0.04$$
$$= 0.5 \times 0.3 = 0.15$$

	S_1	S_2	S_3	a	b	c
S_1	0.0	0.5	0.5	0.5	0.5	0.0
S_2	1.0	0.0	0.0	0.3	0.3	0.4
S_3	0.0	1.0	0.0	0.25	0.0	0.75

Gaussian Mixture Model (GMM)

1.1
1.9
3.3
4.0
5.8
6.5
7.2
8.0
8.5
9.1

$$P(x_i \mid \mu_j, \sigma_j^2) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(x_i-\mu_j)^2}{2\sigma_j^2}\right)$$

$K = 2$ (No. of cluster)

$\mu_1 = 2.0$ $\mu_2 = 7.0$ (Mean)

$\sigma_1^2 = 1.0$ $\sigma_2^2 = 1.0$

$\pi_1 = 0.5$ $\pi_2 = 0.5$

$$\cancel{= \frac{1}{\sqrt{2\times 0.5\times 1.0}} e^{-\frac{(1.1-2.0)}{2\times 1.0}} = \frac{1}{1}\times e^{0.9/2} = 0.82}$$

$$\cancel{G1 = \frac{1}{\sqrt{2\times 0.5\times 1.0}} \times e^{-\frac{(1.1-2.0)^2}{\ldots}}}$$

$$\cancel{G1 = \frac{1}{\sqrt{2\times 0.5\times 1.0}} \times e^{-\left(\frac{(1.1-2.0)^2}{2}\right)} = \frac{1}{1}\times e^{-0.405}}$$

$$\cancel{G2 = \frac{1}{\sqrt{2\times 0.5\times 1.0}} e^{-\left(\frac{(1.1-7.0)^2}{2\times 1}\right)}}$$

$$\cancel{\mu(x)} = \pi_j \times P(x_i \mid \mu, \sigma_j^2) \Big/ \pi_j P(x_i \mid \mu_j, \sigma_j^2) + \mu_2 \times P(x_i)$$

$\gamma_{1,1}$

$\gamma_{1,2}$