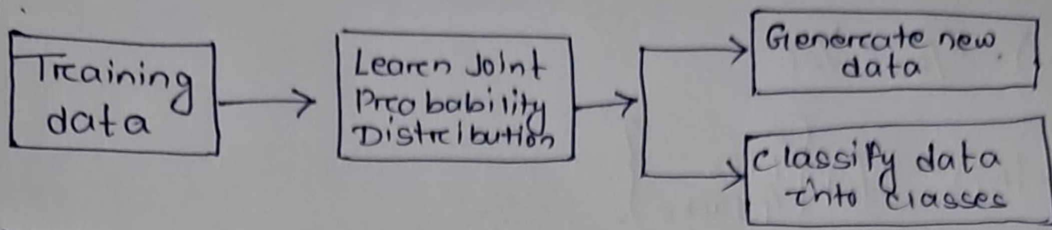
Question-1 Given a set of training data, Probabilistic Generative models learns the Joint probability distribution that represents the training data and uses this underlying distribution to generate new data similar to training data



Applications :-

1) Image Generation: used in GANs and VAEs for creating realistic images

2) Speech Recognition: HMMs are widely used in speech-to-text application

3) Anamoly Detection: Identifies fraud or defects.

Question 2

Deterministic Model	Generative Model
• Given a set of training data it constructs a deterministic function and uses the function to generate new output for new input.	• Given a set of training data, it constructs a Joint probability distribution and uses the distribution to generate new data.
• Produce the same output for the same input	• can produce different outputs for the same input.
• predictable, consistent and operate on a set of rules and algo.	• creative, versatile and can generates new data.
ex Regression, Decesion Tree	ex Naive Bayes Classification.

Question-3

A1 is the event that the 1st card is a red ace (Ace of hearts or Ace of diamonds)

A_2 is the event that the 2nd card is a 10 or a Jack.

A_3 is the event that the 3rd card is greater than 3 but less than 7

$P(A_1) = 2/52 = 1/26$

$P(A_2|A_1) = 8/51 \left\{\begin{matrix} 4 \text{ 10s} \\ 4 \text{ Jacks} \end{matrix}\right\}$.

$P(A_3|A_1 \cap A_2) = 12/50 = 6/25$

$P(A_1 \cap A_2 \cap A_3) = P(A_1) \times P(A_2|A_1) \times P(A_3|A_1 \cap A_2)$

$= 1/26 \times 8/51 \times 6/25$

$= 8/5525 = 0.145\%$

Question 4

Let A_1, A_2, A_3, A_4 be the events of selecting a good quart in each draw without replacement

since there are 15 good quarts out of 20

$P(A_1) = 15/20 = 3/4$

After selecting one good quartz 14 good one out of 19.

$P(A_2|A_1) = 14/19$

For 3rd Good quartz there is 13 good quartz out of 18

$P(A_3|A_1 \cap A_2) = 13/18$

for the 4th good quartz there is 12 good out of 17

$P(A_4|A_1 \cap A_2 \cap A_3 \cap A_3) = 12/17$

$P(A_1 \cap A_2 \cap A_3 \cap A_4) = P(A_1) \times P(A_2|A_1) \times P(A_3|A_1 \cap A_2) \times P(A_4/A_1 \cap A_2 \cap A_3)$

$= 15/20 \times 14/19 \times 13/18 \times 12/17 = 91/323 \approx 0.2818$

Question-5

$f(x) = x/2 \quad 0 \leq x < 2$

$y = \frac{1 - \sqrt{4 - x^2}}{2}$

$\Rightarrow 1 - y = \frac{\sqrt{4 - x^2}}{2}$

$\Rightarrow 2(1-y) = \sqrt{4 - x^2}$

$\Rightarrow 4(1-y)^2 = 4 - x^2$

$\Rightarrow x^2 = 4 - 4(1-y)^2$

$\Rightarrow x = \sqrt{4 - 4(1-y)^2}$

$\Rightarrow \frac{dx}{dy} = \frac{d}{dy}\sqrt{4 - 4(1-y)^2}$

$= \frac{-4(1-y)}{\sqrt{4 - 4(1-y)^2}}$

$g(y) = f(x|y) \left| \frac{dx(y)}{dy} \right|$

$= \frac{\sqrt{4-4(1-y)^2}}{2} \left| \frac{-4^2(1-y)}{\sqrt{4-4(1-y)^2}} \right|$

$= 2(1-y)$

range of y =

when $x = 0$

$y(0) = \frac{1 - \sqrt{4 - 0}}{2} = 0$

when $x = 2$

$y(1) = \frac{1 - \sqrt{4 - 4}}{2} = 1$

$g(y) = \begin{cases} 2(1-y), & 0 \leq y \leq 1 \\ 0, & \text{otherwise} \end{cases}$

Question-6

From Question 5 we already derived,

$$y = 1 - \frac{\sqrt{4-x^2}}{2}$$

$$y(x) = \frac{1-\sqrt{4-x^2}}{2}, \quad 0 \le x < 2$$

Question-7

Empirical distribution $F_n(x) = \frac{\text{No. of observation} \le x}{n}$

data = $[0, 1, 2, 2, 4, 6, 8, 7]$

Here $n = 8$

The Empirical CDF increases

$$F_n(x) = \begin{cases} 0 & x < 0 \\ 1/8 & 0 \le x < 1 \\ 2/8 & 1 \le x < 2 \\ 4/8 & 2 \le x < 4 \\ 5/8 & 4 \le x < 6 \\ 7/8, & 6 \le x < 7 \\ 1, & x \ge 7 \end{cases}$$

Question-8

$$F(x, \theta) = \frac{\theta}{x^{\theta+1}}, \quad x > 1$$

0, elsewhere

For a sample $x_1, x_2, x_3, \dots x_n$ the likelihood function is

$$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta) = \prod_{i=1}^{n} \frac{\theta}{x_i^{\theta+1}}$$

$$L(\theta) = \prod_{i=1}^{n} \theta \prod_{i=1}^{n} \frac{1}{x_i^{\theta+1}}$$

$$L(\theta) = \theta^n \cdot \prod_{i=1}^{n} x_i^{-(\theta+1)}$$

Applying Log

$$\text{Log } L(\theta) = \log(\theta^n) + \log\left(\prod_{i=1}^{n} x_i^{-(\theta+1)}\right)$$

$$\Rightarrow n \log\theta + \sum_{i=1}^{n} \log x_i^{-(\theta+1)}$$

$$\Rightarrow n \log\theta - (\theta+1) \sum_{i=1}^{n} \log x_i$$

$$\frac{d}{d\theta}\left(n(\log\theta)) - (\theta+1)\sum_{i=1}^{n}\log x_i\right) = 0$$

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

$$\frac{n}{\theta} = \sum_{i=1}^{n}\log x_i$$

$$\frac{\theta}{n} = \frac{1}{\sum_{i=1}^{n}\log x_i}$$

$$\theta = \frac{n}{\sum_{i=1}^{n}\log x_i}$$

Question 9

KL divergence between two probability distribution P and Q -

$$D_{KL}(P||Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)}$$

$$D_{KL}(Q||P) = \sum_x Q(x) \log \frac{Q(x)}{P(x)}$$

$P(x) \rightarrow$ Binomial Distribution

$n = 2$ $\quad p = 0.4$

$$P(x) = \binom{2}{x} 0.4^x (0.6)^{2-x}$$

for $x = 0, 1, 2$:-

$$P(0) = \binom{2}{0}(0.4)^0(0.6)^2 = 0.36$$

$$P(1) = \binom{2}{1}(0.4)^1(0.6)^1 = 2\times0.4\times0.6 = 0.48$$

$$P(2) = \binom{2}{2}(0.4)^2(0.6)^0 = 1\times0.16\times1 = 0.16$$

Compute $Q(x) \rightarrow$ uniform distribution)

$$Q(0) = Q(1) = Q(2) = 1/3$$

$$D_{KL}(P||Q) = \sum_x P(x)\log_2\frac{P(x)}{Q(x)}$$

$$= 0.36\times\log_2\frac{0.36}{1/3} + 0.48\times\log_2\frac{0.48}{1/3} + 0.16\log_2\frac{0.16}{1/3}$$

$$= 0.134$$

$$D_{KL}(Q\|P) = \sum_x Q(x)\log_2 \frac{Q(x)}{P(x)}$$

$$= \tfrac{1}{3}\log_2 \frac{1/3}{0.36} + \tfrac{1}{3}\log_2 \frac{1/3}{0.48} + \tfrac{1}{3}\log_2 \frac{1/3}{0.16}$$

$$= 0.141$$

Question 10

Chills(C)	Runnynose(R)	Headache(H)	Fever(F)	Flu
Y	N	mild	Y	No
Y	Y	NO	N	yes
Y	N	Strong	Y	yes
N	Y	mild	Y	yes
N	N	NO	N	NO
N	Y	strong	Y	yes
N	Y	Strong	N	NO
Y	Y	mild	Y	yes

Let B = {Chills = Y, RunnyNose = N, Headache = mild, and fever = Y}

$$P(\text{flu}=\text{yes}\mid B) = \frac{P(B\mid \text{flu}=\text{yes})\,P(\text{flu}=\text{yes})}{P(B)}$$

$= P(\text{chills}=Y \mid \text{flu}=\text{yes})\,P(\text{Runnynose}=N \mid \text{flu}=\text{yes})$
$P(\text{Headache}=\text{mild} \mid \text{flu}=\text{yes})\,P(\text{fever}=y \mid \text{flu}=\text{yes})$
$P(\text{flu}=\text{yes})$

$= 0.5 \times 0.25 \times 0.5 \times 1 \times 0.5$

$= 0.03125$

$$P(\text{flu}=\text{NO}\mid B) = \frac{P(B\mid \text{flu}=\text{NO})\,P(\text{flu}=\text{NO})}{P(B)}$$

$= P(\text{chills}=y \mid \text{flu}=\text{NO})\,P(\text{Runnynose}=N \mid \text{flu}=\text{NO})$
$P(\text{Headache}=\text{mild} \mid \text{flu}=\text{NO})\,P(\text{fever}=Y \mid \text{flu}=\text{NO})$
$P(\text{flu}=\text{NO})$

$= 0.5 \times 0.5 \times 0.25 \times 0.25 \times 0.5 = 0.0078$

Since $0.03125 > 0.0078$

The patient is likely to have flu = yes.

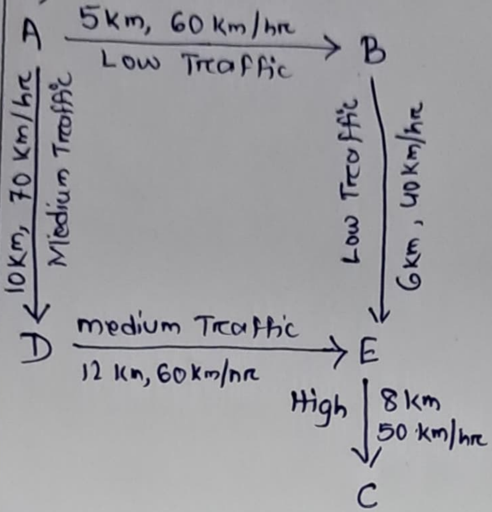
Question = 11

Graph in AI -

Graph are produced / Powerful structures used in AI to represent relationships betwn entities. A graph cosist of -

(i) Nodes
(ii) Edges

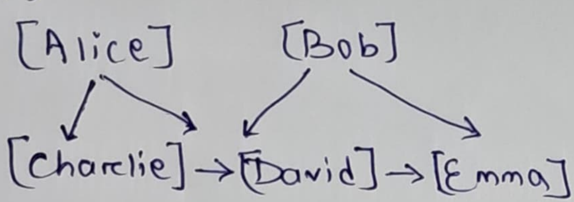
a)



Node A, B, C, D, E represent intersections
Edges represent roads betwn intersections

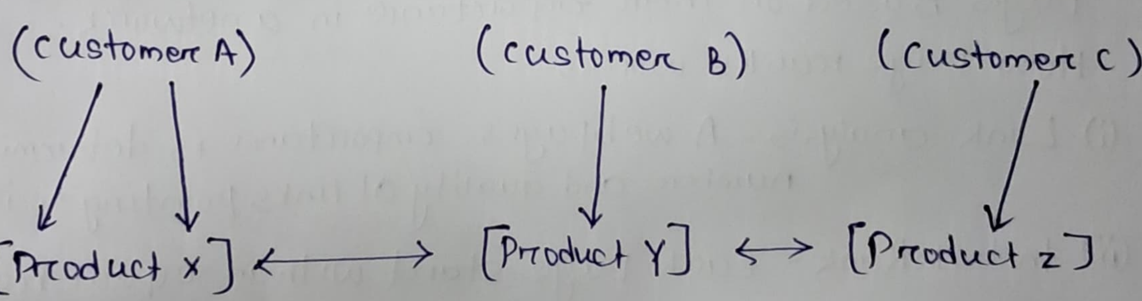
Best Route → A → B → C = 13Km

b. (b)

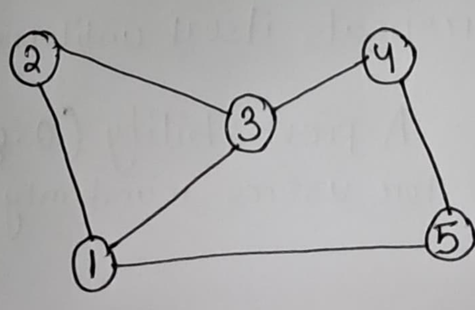


Three features influencing
- Common Interest
- Interaction frequency
- Mutual friends

(c)



Question 12



The differcent types of graph representation are:-

(i) AdJacency Matrix

(ii) AdJacency List

(i) AdJacency Matrix

	1	2	3	4	5
1	0	1	1	0	1
2	1	0	1	0	0
3	1	1	0	1	1
4	0	0	1	0	1
5	1	0	1	1	0

(ii) AdJacency List:

$1 \rightarrow 2,3,5$

$2 \rightarrow 1,3$

$3 \rightarrow 1,2,4,5$

$4 \rightarrow 3,5$

$5 \rightarrow 1,3,5$

Question 13

Page rank is an algorithm developed by larry Page rank web Pages based on their importance in a network.

How page ranks works-

(i) Link analysis - A webpage's importance is determined by the number and quality of links pointing to it.

(ii) Initial Rank - Each page start with an equal rank

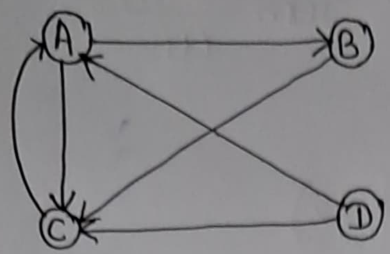
(iii) equal rank distributions

(iv) The process repeats itself until ranks stabilize

(v) Damping Factor - A probability (0.85) is used to account for users randomly Jumping to another page

Question-14

(a)



(b)

	1st Iteration	2nd Iteration	3rd iteration
A	0.25	0.375	0.5
B	0.25	0.125	0.1875
C	0.25	0.5	0.3125
D	0.25	0	0

For 2nd Iteration:

$$PR(A) = \frac{0.25}{1} + \frac{0.25}{2} = 0.375$$

$$PR(B) = \frac{0.25}{2} = 0.125$$

$$PR(C) = \frac{0.25}{2} + \frac{0.25}{1} + \frac{0.25}{2} = 0.5$$

$$PR(D) = 0$$

For 3rd Iteration:

$$PR(A) = 0.5 + 0 = 0.5$$

$$PR(B) = \frac{0.375}{2} = 0.1875$$

$$PR(C) = \frac{0.375}{2} + \frac{0.125}{1} + 0 = 0.3125$$

$$PR(D) = 0$$

(c)

	1st iteration	2nd iteration	3rd iteration
A	0.25	0.35625	0.44656
B	0.25	0.14375	0.18891
C	0.25	0.4625	0.32703
D	0.25	0.0375	0.0375

$d = 0.85$
$N = 4$

$$PR^{(t+1)}(i) = (1-d)\frac{1}{N} + d\sum_{j\in K(i)}\frac{PR^{(t)}(j)}{C(j)}$$

For 2nd Iteration:

$PR(A) = (1-0.85)\frac{1}{4} + 0.85\left(0.25 + \frac{0.25}{2}\right)$

$= 0.35625$

$PR(B) = (1-0.85)\frac{1}{4} + 0.85\left(\frac{PR(A)}{C(A)}\right)$

$= 0.14375$

$PR(C) = (1-0.85)\frac{1}{4} + 0.85\left(\frac{PR(A)}{C(A)} + \frac{PR(B)}{C(B)} + \frac{PR(D)}{C(D)}\right)$

$= 0.4625$

$PR(D) = (1-0.85)\frac{1}{4} + 0.85\times 0$

$= 0.0375$

For 3rd Iteration:

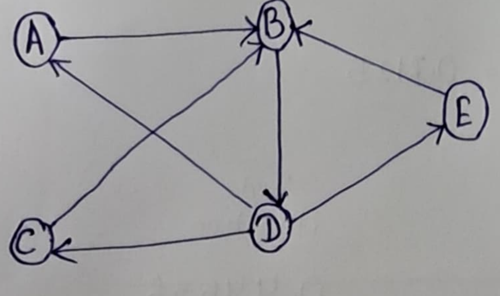
$PR(A) = 0.0375 + 0.85\left(\frac{0.4625}{1} + \frac{0.0375}{2}\right) = 0.44656$

$PR(B) = 0.0375 + 0.85(0.17813) = 0.18891$

$PR(C) = 0.0375 + 0.85\left(\frac{0.35625}{2} + \frac{0.14375}{1} + \frac{0.0375}{2}\right) = 0.32703$

$PR(D) = 0.0375.$

Question 15



(a)

	A	B	C	D	E
A	0	1	0	0	0
B	0	0	0	1	0
C	0	1	0	0	0
D	1/3	0	1/3	0	1/3
E	0	1	0	0	0

(b)

	1st	2nd
A	0.0667	0.0667
B	0.6	0.2
C	0.0667	0.0667
D	0.2	0.6
E	0.0667	0.0667

$PR^{(0)} = [0.2, 0.2, 0.2, 0.2, 0.2]$

1st Iteration :-

$PR^{(1)}(A) = 0.0667$

$PR^{(1)}(B) = 0.6$

$PR^{(1)}(C) = 0.0667$

$PR^{(1)}(D) = 0.2$

$PR'(E) = 0.0667$

(C) $PR^{2}(A) = 0.2/3 = 0.0667$

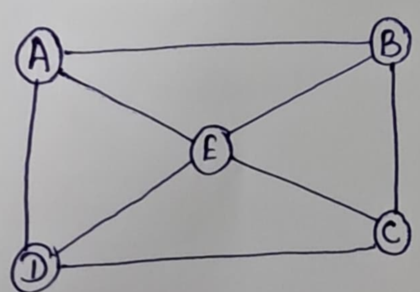
$PR^{(2)}(B) = 0.2$

$PR^{(2)}(C) = 0.2/3$

$PR^{(2)}(D) = 0.6$

$PR^{(2)}(E) = 0.2/3 = 0.0667$

Question 16



(a) Aggregated Message (Node) = Avg. Node values

For node A -

$$A_{new} = \frac{B+D+E}{3} = \left[\frac{1+8+5}{3}, \frac{0+3+1}{3}, \frac{2+0+2}{3}\right] = [4.67, 1.33, 1.33]$$

$B_{new} = \frac{A+C+E}{3} = \left[\frac{7}{3}, \frac{4}{3}, \frac{6}{3}\right] = [2.33, 1.33, 2.00]$

$C_{new} = \frac{B+E+D}{3} = [4.67, 1.33, 1.33]$

$D_{new} = [2.33, 1.33, 2.00]$

$E_{new} = \left[\frac{11}{4}, \frac{6}{4}, \frac{6}{4}\right] = [2.75, 1.50, 1.50]$

(b)

$A_{new} = \frac{1\times8\times5,\ 0\times3\times1,\ 2\times0\times2}{(1*8*5), (0+3+1), (2+0+2)}$

$= [2.86, 0, 0]$

$B_{new} \rightarrow [0.71, 0.50, 1.00]$

$C_{new} \Rightarrow [2.86, 0, 0]$

$D_{new} \rightarrow [0.71, 0.50, 1.00]$

(C)

$A \rightarrow \sqrt{(1+8+5), (0+3+1), (2+0+2)}$

$= [3.74, 2.00, 2.00] = A_{new}$

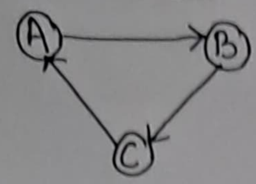
$B \rightarrow [2.65, 2.00, 2.45] = B_{new}$

$C \rightarrow [3.74, 2.00, 2.00] = C_{new}$

$D \rightarrow [2.65, 2.00, 2.45] = D_{new}$

$E \rightarrow [3.32, 2.45, 2.45] = E_{new}$

Question-17



$A \to B$ $\quad$ $A \cdot [1,0,1]$

$C \to A$ $\quad$ $B = [0,2,2]$

$B \to C$ $\quad$ $C = [3,1,2]$

New Values of each node

$A_{new} = C_{old} \times C_{old} = [3,1,2] \times [3,1,2]$
$= [9,1,4]$

$B_{new} = [1,0,1] \times [1,0,1]$
$[1,0,1]$

$C_{new} = [0,2,2] \times [0,2,2]$
$= [0,4,4]$

Question-18

Bayesian Network - is a graphical model that represents probalistic relationship among a set of variables, Each node in the network represents a random variable and the edges (arcs) represents probalistics dependencies. These networks are particularly useful for reasoning under uncertainity, as they allow us to model complex domains with conditional dependencies.

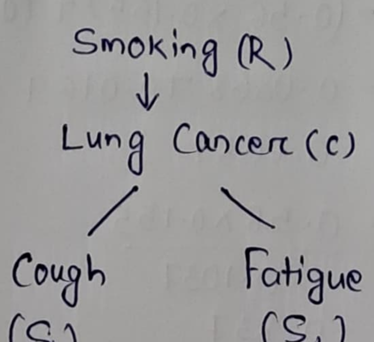
Components:-

(i) Nodes - represent random variables

(ii) Edges - Condition dependencies

Question 19

(a)



(b) Prior Probabilities:

$P(R=T)=0.3$, $P(R=F)=0.7$

Conditional Probabilities:

$P(C=T|R=T)=0.4$ $\quad$ $P(C=F|R=T)=0.6$

$P(C=T|R=F)=0.05$ $\quad$ $P(C=F|R=F)=0.95$

$P(S_1=T|C=T)=0.8$ $\quad$ $P(S_1=F|C=T)=0.2$

$P(S_1=T|C=F)=0.1$ $\quad$ $P(S_1=F|C=F)=0.9$

$P(S_2=T|C=T)=0.7$ $\quad$ $P(S_2=F|C=T)=0.3$

$P(S_2=T|C=F)=0.2$ $\quad$ $P(S_2=F|C=F)=0.8$

(c) $P(S_1=T, S_2=T|C=T) = P(S_1=T|C=T) \cdot P(S_2=T|C=T)$

$= 0.8 \times 0.7 = 0.56$

$P(S_1=T, S_2=T | C=F) = P(S_1=T|C=F)\, P(S_2=T|C=F)$

$= 0.1 \times 0.2 = 0.2$

$P(S_1=T, S_2=T) = P(S_1=T, S_2=T|C=T) + P(S_1=T, S_2=T|C=F)$

$P(C=1)$

$P(C=T) = P(R=T)P(C=T|R=T) + P(R=F)P(C=T|R=F)$

$= 0.155$

$P(C=F) = 1 - P(C=T) = 1 - 0.155 = 0.845$

$=$

Now,

$P(S_1=T, S_2=T) = (0.56 \times 0.155) + (0.02 \times 0.845)$

$= 0.0868 + 0.0169$

$= 0.1037$

$P(C=T|S_1=T, S_2=T) = \frac{0.56 \times 0.155}{0.1037}$

$= 0.837$

Question 20

(a)

$$\begin{tikzcd} \text{suspicious words } (w) \arrow[dr, dash] & & \text{unknown Sender}(U) \arrow[dl, dash] \\ & \text{Spam}(S) & \end{tikzcd}$$

(b) $P(w=1|s=1) = \dfrac{P(S=1|w=1)\ P(w=1)}{P(S=1)}$

$$\begin{aligned} P(S=1) &= P(S=1|w=1, U=1) \cdot P(w=1)\ P(U=1) \\ &\quad + (P\ S=1|w=1/U=0)\ P(w=1) \cdot P(U=0) \\ &\quad + P(S=1|W=0, u=1)\ \ P(w=0)\ P(U=1) \\ &\quad + P(S=1|W=0, U=0)\ P(w=0)\ P(U=0) \\ &= (0.95 \times 0.4 \times 0.3) + (0.8 \times 0.4 \times 0.7) + (0.7 \times 0.6 \times 0.3) + \\ &\qquad (0.1 \times 0.6 \times 0.7) \\ &= 0.506 \end{aligned}$$

$$P(w=1|s=1) = \frac{(0.95 \times 0.4 \times 0.3) + (0.8 \times 0.4 \times 0.7)}{0.506}$$

$$= 0.668$$