Part 1: Reasoning Under Uncertainty Basics

Q1.1)

To compute the table of the joint distribution $P(X, Y, Z)$ we can use the probability chain rule
$P(X, Y, Z) = P(Z \mid Y, X) * P(Y \mid X) * P(X)$
However, as Z and X are conditionally independent given Y, we can simplify the rule to
$P(X, Y, Z) = P(Z \mid Y) * P(Y \mid X) * P(X)$

Calculating the joint probabilities
$P(X = 0, Y = 0, Z = 0)$
$= P(Z = 0 \mid Y = 0) * P(Y = 0 \mid X = 0) * P(X = 0) = 0.7 * 0.1 * 0.35 = 0.0245$

$P(X = 0, Y = 0, Z = 1)$
$= P(Z = 1 \mid Y = 0) * P(Y = 0 \mid X = 0) * P(X = 0) = 0.3 * 0.1 * 0.35 = 0.0105$

$P(X = 0, Y = 1, Z = 0)$
$= P(Z = 0 \mid Y = 1) * P(Y = 1 \mid X = 0) * P(X = 0) = 0.2 * 0.9 * 0.35 = 0.063$

$P(X = 0, Y = 1, Z = 1)$
$= P(Z = 1 \mid Y = 1) * P(Y = 1 \mid X = 0) * P(X = 0) = 0.8 * 0.9 * 0.35 = 0.252$

$P(X = 1, Y = 0, Z = 0)$
$= P(Z = 0 \mid Y = 0) * P(Y = 0 \mid X = 1) * P(X = 1) = 0.7 * 0.6 * 0.65 = 0.273$

$P(X = 1, Y = 0, Z = 1)$
$= P(Z = 1 \mid Y = 0) * P(Y = 0 \mid X = 1) * P(X = 1) = 0.3 * 0.6 * 0.65 = 0.117$

$P(X = 1, Y = 1, Z = 0)$
$= P(Z = 0 \mid Y = 1) * P(Y = 1 \mid X = 1) * P(X = 1) = 0.2 * 0.4 * 0.65 = 0.052$

$P(X = 1, Y = 1, Z = 1)$
$= P(Z = 1 \mid Y = 1) * P(Y = 1 \mid X = 1) * P(X = 1) = 0.8 * 0.4 * 0.65 = 0.208$

X	Y	Z	P(Z | Y)	P(Y | X)	P(X)	P(X, Y, Z)
0	0	0	0.7	0.1	0.35	0.0245
0	0	1	0.3	0.1	0.35	0.0105
0	1	0	0.2	0.9	0.35	0.063
0	1	1	0.8	0.9	0.35	0.252
1	0	0	0.7	0.6	0.65	0.273
1	0	1	0.3	0.6	0.65	0.117
1	1	0	0.2	0.4	0.65	0.052
1	1	1	0.8	0.4	0.65	0.208

Q1.2)
To create the full joint table of X and Y we can use the product rule
$P(X, Y) = P(Y \mid X) * P(X)$

Calculating the joint probabilities
$P(X = 0, Y = 0)$
$= P(Y = 0 \mid X = 0) * P(X = 0) = 0.1 * 0.35 = 0.035$

$P(X = 0, Y = 1)$
$= P(Y = 1 \mid X = 0) * P(X = 0) = 0.9 * 0.35 = 0.315$

$P(X = 1, Y = 0)$
$= P(Y = 0 \mid X = 1) * P(X = 1) = 0.6 * 0.65 = 0.39$

$P(X = 1, Y = 1)$
$= P(Y = 1 \mid X = 1) * P(X = 1) = 0.4 * 0.65 = 0.26$

X	**Y**	**P(Y | X)**	**P(X)**	**P(X, Y)**
0	0	0.1	0.35	0.035
0	1	0.9	0.35	0.315
1	0	0.6	0.65	0.39
1	1	0.4	0.65	0.26

Q1.3)

a) To find $P(Z = 0)$ we can use the Sum rule for where $Z = 0$

$$P(Z = 0) = P(X = 0, Y = 0, Z = 0) + P(X = 0, Y = 1, Z = 0)$$
$$+ P(X = 1, Y = 0, Z = 0) + P(X = 1, Y = 1, Z = 0)$$

$$P(Z = 0) = 0.0245 + 0.063 + 0.273 + 0.052$$

$$P(Z = 0) = 0.4125$$

b) To find $P(X = 0, Z = 0)$ we can use the Sum rule for where $X = 0, Z = 0$

$$P(X = 0, Z = 0) = P(X = 0, Y = 0, Z = 0) + P(X = 0, Y = 1, Z = 0)$$

$$P(X = 0, Z = 0) = 0.0245 + 0.063$$

$$P(X = 0, Z = 0) = 0.0875$$

c) To find $P(X = 1, Y = 0 \mid Z = 1)$ we can rearrange the product/chain rule to
$P(X, Y \mid Z) = \frac{P(X, Y, Z)}{P(Z)}$
We can then use the Sum rule to find $P(Z)$

$$P(X = 1, Y = 0 \mid Z = 1) = \frac{P(X=1, Y=0, Z=1)}{P(X=0, Y=0, Z=1)+P(X=0, Y=1, Z=1)+P(X=1, Y=0, Z=1)+P(X=1, Y=1, Z=1)}$$

$$P(X = 1, Y = 0 \mid Z = 1) = \frac{0.117}{0.0105+0.252+0.117+0.208}$$

$$P(X = 1, Y = 0 \mid Z = 1) = 0.199$$

d) To find $P(X = 0 \mid Y = 0, Z = 0)$ we can rearrange the product/chain rule to
$P(X \mid Y, Z) = \frac{P(X, Y, Z)}{P(Y, Z)}$
We can then use the Sum rule to find $P(Y, Z)$

$$P(X = 0 \mid Y = 0, Z = 0) = \frac{P(X=0, Y=0, Z=0)}{P(X=0, Y=0, Z=0)+P(X=1, Y=0, Z=0)}$$

$$P(X = 0 \mid Y = 0, Z = 0) = \frac{0.0245}{0.0245+0.273}$$

$$P(X = 0 \mid Y = 0, Z = 0) = 0.0823$$

Q2)

1) To find $P(B = t, C = t)$ we can use the product rule
$P(B = t, C = t) = P(B = t \mid C = t) * P(C = t)$
$P(B = t, C = t) = 0.2 * 0.4$
$P(B = t, C = t) = 0.08$

2) A can either be t or f, and all probabilities must sum to 1. We also know that $P(A = t \mid B = t) = 0.3$. So we can find $P(A = f \mid B = t)$ by subtracting $P(A = t \mid B = t)$ from 1

$P(A = f \mid B = t) = 1 - P(A = t \mid B = t)$
$P(A = f \mid B = t) = 1 - 0.3$
$P(A = f \mid B = t) = 0.7$

3) As A and B are conditionally independent given C, we can find $P(A = t, B = t \mid C = t)$

$P(A = t, B = t \mid C = t) = P(A = t \mid C = t) * P(B = t \mid C = t)$
$P(A = t, B = t \mid C = t) = 0.5 * 0.2$
$P(A = t, B = t \mid C = t) = 0.1$

4) As A and B are conditionally independent given C, we can simplify
$P(A = t \mid B = t, C = t)$ to $P(A = t \mid C = t)$
We already know that $P(A = t \mid C = t) = 0.5$, so
$P(A = t \mid B = t, C = t) = 0.5$

5) We can use the product/chain rule to find $P(A = t, B = t, C = t)$ by
$P(A = t, B = t, C = t) = P(C = t) * P(A = t, B = t \mid C = t)$
$P(A = t, B = t, C = t) = 0.4 * 0.1$
$P(A = t, B = t, C = t) = 0.04$

Part 2: Naive Bayes Method

Q2.1

No-recurrence-events

$P(Age = 10 - 19 \mid Class = no - recurrence - events) = 0.005263157894736842$
$P(Age = 20 - 29 \mid Class = no - recurrence - events) = 0.010526315789473684$
$P(Age = 30 - 39 \mid Class = no - recurrence - events) = 0.11578947368421053$
$P(Age = 40 - 49 \mid Class = no - recurrence - events) = 0.3263157894736842$
$P(Age = 50 - 59 \mid Class = no - recurrence - events) = 0.34210526315789475$
$P(Age = 60 - 69 \mid Class = no - recurrence - events) = 0.2$
$P(Age = 70 - 79 \mid Class = no - recurrence - events) = 0.031578947368421054$
$P(Age = 80 - 89 \mid Class = no - recurrence - events) = 0.012658227848101266$
$P(Age = 90 - 99 \mid Class = no - recurrence - events) = 0.005263157894736842$

$P(Menopause = lt40 \mid Class = no - recurrence - events) = 0.031578947368421054$
$P(Menopause = ge40 \mid Class = no - recurrence - events) = 0.4631578947368421$
$P(Menopause = premeno \mid Class = no - recurrence - events) = 0.5157894736842106$

$P(Tumor - size = 0 - 4 \mid Class = no - recurrence - events) = 0.042105263157894736$
$P(Tumor - size = 5 - 9 \mid Class = no - recurrence - events) = 0.02631578947368421$
$P(Tumor - size = 10 - 14 \mid Class = no - recurrence - events) = 0.1368421052631579$
$P(Tumor - size = 15 - 19 \mid Class = no - recurrence - events) = 0.12105263157894737$
$P(Tumor - size = 20 - 24 \mid Class = no - recurrence - events) = 0.18421052631578946$
$P(Tumor - size = 25 - 29 \mid Class = no - recurrence - events) = 0.16842105263157894$
$P(Tumor - size = 30 - 34 \mid Class = no - recurrence - events) = 0.17894736842105263$
$P(Tumor - size = 35 - 39 \mid Class = no - recurrence - events) = 0.06315789473684211$
$P(Tumor - size = 40 - 44 \mid Class = no - recurrence - events) = 0.08947368421052632$
$P(Tumor - size = 45 - 49 \mid Class = no - recurrence - events) = 0.015789473684210527$
$P(Tumor - size = 50 - 54 \mid Class = no - recurrence - events) = 0.02631578947368421$
$P(Tumor - size = 55 - 59 \mid Class = no - recurrence - events) = 0.005263157894736842$

$P(Inv - nodes = 3 - 5 \,|\, Class = no - recurrence - events) = 0.08947368421052632$
$P(Inv - nodes = 6 - 8 \,|\, Class = no - recurrence - events) = 0.042105263157894736$
$P(Inv - nodes = 9 - 11 \,|\, Class = no - recurrence - events) = 0.015789473684210527$
$P(Inv - nodes = 12 - 14 \,|\, Class = no - recurrence - events) = 0.010526315789473684$
$P(Inv - nodes = 15 - 17 \,|\, Class = no - recurrence - events) = 0.021052631578947368$
$P(Inv - nodes = 18 - 20 \,|\, Class = no - recurrence - events) = 0.005263157894736842$
$P(Inv - nodes = 21 - 23 \,|\, Class = no - recurrence - events) = 0.005263157894736842$
$P(Inv - nodes = 24 - 26 \,|\, Class = no - recurrence - events) = 0.005263157894736842$
$P(Inv - nodes = 27 - 29 \,|\, Class = no - recurrence - events) = 0.005263157894736842$
$P(Inv - nodes = 30 - 32 \,|\, Class = no - recurrence - events) = 0.005263157894736842$
$P(Inv - nodes = 33 - 35 \,|\, Class = no - recurrence - events) = 0.005263157894736842$
$P(Inv - nodes = 36 - 39 \,|\, Class = no - recurrence - events) = 0.005263157894736842$

$P(Node - caps = yes \,|\, Class = no - recurrence - events) = 0.12631578947368421$
$P(Node - caps = no \,|\, Class = no - recurrence - events) = 0.8789473684210526$

$P(Deg - malig = 1 \,|\, Class = no - recurrence - events) = 0.29473684210526313$
$P(Deg - malig = 2 \,|\, Class = no - recurrence - events) = 0.5157894736842106$
$P(Deg - malig = 3 \,|\, Class = no - recurrence - events) = 0.2$

$P(Breast = left \,|\, Class = no - recurrence - events) = 0.5105263157894737$
$P(Breast = right \,|\, Class = no - recurrence - events) = 0.49473684210526314$

$P(Breast - quad = left - up \,|\, Class = no - recurrence - events) = 0.3526315789473684$
$P(Breast - quad = left - low \,|\, Class = no - recurrence - events) = 0.3736842105263158$
$P(Breast - quad = right - up \,|\, Class = no - recurrence - events) = 0.11052631578947368$
$P(Breast - quad = right - low \,|\, Class = no - recurrence - events) = 0.09473684210526316$
$P(Breast - quad = central \,|\, Class = no - recurrence - events) = 0.08947368421052632$

$P(Irradiat = yes \,|\, Class = no - recurrence - events) = 0.15789473684210525$
$P(Irradiat = no \,|\, Class = no - recurrence - events) = 0.8473684210526315$

Recurrence-events

$P(Age\ =\ 10-19\,|\,Class\ =\ recurrence-events)\ =\ 0.012658227848101266$
$P(Age\ =\ 20-29\,|\,Class\ =\ recurrence-events)\ =\ 0.012658227848101266$
$P(Age\ =\ 30-39\,|\,Class\ =\ recurrence-events)\ =\ 0.20253164556962025$
$P(Age\ =\ 40-49\,|\,Class\ =\ recurrence-events)\ =\ 0.34177215189873417$
$P(Age\ =\ 50-59\,|\,Class\ =\ recurrence-events)\ =\ 0.27848101265822783$
$P(Age\ =\ 60-69\,|\,Class\ =\ recurrence-events)\ =\ 0.21518987341772153$
$P(Age\ =\ 70-79\,|\,Class\ =\ recurrence-events)\ =\ 0.012658227848101266$
$P(Age\ =\ 80-89\,|\,Class\ =\ recurrence-events)\ =\ 0.012658227848101266$
$P(Age\ =\ 90-99\,|\,Class\ =\ recurrence-events)\ =\ 0.012658227848101266$

$P(Menopause\ =\ lt40\,|\,Class\ =\ recurrence-events)\ =\ 0.012658227848101266$
$P(Menopause\ =\ ge40\,|\,Class\ =\ recurrence-events)\ =\ 0.3924050632911392$
$P(Menopause\ =\ premeno\,|\,Class\ =\ recurrence-events)\ =\ 0.620253164556962$

$P(Tumor-size\ =\ 0-4\,|\,Class\ =\ recurrence-events)\ =\ 0.02531645569620253$
$P(Tumor-size\ =\ 5-9\,|\,Class\ =\ recurrence-events)\ =\ 0.012658227848101266$
$P(Tumor-size\ =\ 10-14\,|\,Class\ =\ recurrence-events)\ =\ 0.02531645569620253$
$P(Tumor-size\ =\ 15-19\,|\,Class\ =\ recurrence-events)\ =\ 0.08860759493670886$
$P(Tumor-size\ =\ 20-24\,|\,Class\ =\ recurrence-events)\ =\ 0.17721518987341772$
$P(Tumor-size\ =\ 25-29\,|\,Class\ =\ recurrence-events)\ =\ 0.24050632911392406$
$P(Tumor-size\ =\ 30-34\,|\,Class\ =\ recurrence-events)\ =\ 0.2911392405063291$
$P(Tumor-size\ =\ 35-39\,|\,Class\ =\ recurrence-events)\ =\ 0.10126582278481013$
$P(Tumor-size\ =\ 40-44\,|\,Class\ =\ recurrence-events)\ =\ 0.08860759493670886$
$P(Tumor-size\ =\ 45-49\,|\,Class\ =\ recurrence-events)\ =\ 0.02531645569620253$
$P(Tumor-size\ =\ 50-54\,|\,Class\ =\ recurrence-events)\ =\ 0.05063291139240506$
$P(Tumor-size\ =\ 55-59\,|\,Class\ =\ recurrence-events)\ =\ 0.012658227848101266$

$P(Inv-nodes\ =\ 0-2\,|\,Class\ =\ recurrence-events)\ =\ 0.5443037974683544$
$P(Inv-nodes\ =\ 3-5\,|\,Class\ =\ recurrence-events)\ =\ 0.20253164556962025$
$P(Inv-nodes\ =\ 6-8\,|\,Class\ =\ recurrence-events)\ =\ 0.13924050632911392$
$P(Inv-nodes\ =\ 9-11\,|\,Class\ =\ recurrence-events)\ =\ 0.0759493670886076$
$P(Inv-nodes\ =\ 12-14\,|\,Class\ =\ recurrence-events)\ =\ 0.0379746835443038$
$P(Inv-nodes\ =\ 15-17\,|\,Class\ =\ recurrence-events)\ =\ 0.05063291139240506$
$P(Inv-nodes\ =\ 18-20\,|\,Class\ =\ recurrence-events)\ =\ 0.012658227848101266$
$P(Inv-nodes\ =\ 21-23\,|\,Class\ =\ recurrence-events)\ =\ 0.012658227848101266$
$P(Inv-nodes\ =\ 24-26\,|\,Class\ =\ recurrence-events)\ =\ 0.02531645569620253$
$P(Inv-nodes\ =\ 27-29\,|\,Class\ =\ recurrence-events)\ =\ 0.012658227848101266$
$P(Inv-nodes\ =\ 30-32\,|\,Class\ =\ recurrence-events)\ =\ 0.012658227848101266$
$P(Inv-nodes\ =\ 33-35\,|\,Class\ =\ recurrence-events)\ =\ 0.012658227848101266$
$P(Inv-nodes\ =\ 36-39\,|\,Class\ =\ recurrence-events)\ =\ 0.012658227848101266$

$P(Node - caps = yes \,|\, Class = recurrence - events) = 0.4050632911392405$
$P(Node - caps = no \,|\, Class = recurrence - events) = 0.6075949367088608$

$P(Deg - malig = 1 \,|\, Class = recurrence - events) = 0.11392405063291139$
$P(Deg - malig = 2 \,|\, Class = recurrence - events) = 0.3670886075949367$
$P(Deg - malig = 3 \,|\, Class = recurrence - events) = 0.5443037974683544$

$P(Breast = left \,|\, Class = recurrence - events) = 0.5569620253164557$
$P(Breast = right \,|\, Class = recurrence - events) = 0.45569620253164556$

$P(Breast - quad = left - up \,|\, Class = recurrence - events) = 0.31645569620253167$
$P(Breast - quad = left - low \,|\, Class = recurrence - events) = 0.4050632911392405$
$P(Breast - quad = right - up \,|\, Class = recurrence - events) = 0.17721518987341772$
$P(Breast - quad = right - low \,|\, Class = recurrence - events) = 0.08860759493670886$
$P(Breast - quad = central \,|\, Class = recurrence - events) = 0.06329113924050633$

$P(Irradiat = yes \,|\, Class = recurrence - events) = 0.3924050632911392$
$P(Irradiat = no \,|\, Class = recurrence - events) = 0.620253164556962$

Q2.2
$P(Class = no - recurrence - events) = 0.7116104868913857$
$P(Class = recurrence - events) = 0.2958801498127341$

Q2.3
Input vector: no-recurrence-events,50-59,premeno,50-54,0-2,yes,2,right,left_up,yes
Instance: 0 Score(Class = no-recurrence-events) = 5.0252361135206404E-6
Instance: 0 Score(Class = recurrence-events) = 1.185144746072391E-5
Predicted class: recurrence-events

Input vector: no-recurrence-events,50-59,ge40,35-39,0-2,no,2,left,left_up,no
Instance: 1 Score(Class = no-recurrence-events) = 4.1732849746988816E-4
Instance: 1 Score(Class = recurrence-events) = 4.345530735598768E-5
Predicted class: no-recurrence-events

Input vector: no-recurrence-events,50-59,premeno,10-14,3-5,no,1,right,left_up,no
Instance: 2 Score(Class = no-recurrence-events) = 5.8878218806890874E-5
Instance: 2 Score(Class = recurrence-events) = 1.622419224516929E-6
Predicted class: no-recurrence-events

Input vector: no-recurrence-events,40-49,premeno,10-14,0-2,no,2,left,left_low,yes
Instance: 3 Score(Class = no-recurrence-events) = 1.8965784203051642E-4
Instance: 3 Score(Class = recurrence-events) = 1.7066084343442433E-5
Predicted class: no-recurrence-events

Input vector: no-recurrence-events,50-59,ge40,15-19,0-2,yes,2,left,central,yes
Instance: 4   Score(Class = no-recurrence-events) = 5.434866736709343E-6
Instance: 4   Score(Class = recurrence-events) = 3.207415542941948E-6
Predicted class: no-recurrence-events

Input vector: no-recurrence-events,50-59,premeno,25-29,0-2,no,1,left,left_low,no
Instance: 5   Score(Class = no-recurrence-events) = 7.504740415831185E-4
Instance: 5   Score(Class = recurrence-events) = 6.480302920326064E-5
Predicted class: no-recurrence-events

Input vector: no-recurrence-events,60-69,ge40,25-29,0-2,no,3,right,left_low,no
Instance: 6   Score(Class = no-recurrence-events) = 2.5906825523534885E-4
Instance: 6   Score(Class = recurrence-events) = 1.2384056710871053E-4
Predicted class: no-recurrence-events

Input vector: recurrence-events,60-69,ge40,20-24,0-2,no,1,right,left_up,no
Instance: 7   Score(Class = no-recurrence-events) = 3.9405165026762883E-4
Instance: 7   Score(Class = recurrence-events) = 1.4921120960543075E-5
Predicted class: no-recurrence-events

Input vector: recurrence-events,40-49,ge40,30-34,3-5,no,3,left,left_low,no
Instance: 8   Score(Class = no-recurrence-events) = 4.8934838332255675E-5
Instance: 8   Score(Class = recurrence-events) = 1.0828136273080714E-4
Predicted class: recurrence-events

Input vector: recurrence-events,50-59,ge40,30-34,3-5,no,3,left,left_low,no
Instance: 9   Score(Class = no-recurrence-events) = 5.130265309026804E-5
Instance: 9   Score(Class = recurrence-events) = 8.822925852139842E-5
Predicted class: recurrence-events