[Block 1: MLE & MAP]

Q 1.1 Consider we choose number Uniformly from a set {1,2,3,4, …θ} (θ is an integer) with replacement. Suppose the numbers we choose are 2,5,7, then based on MLE, what value of θ do you estimate?

Options:

1) 3
2) 5
3) 7
4) 9

[Block 1: MLE & MAP]

Q 1.1 Consider we choose number Uniformly from a set {1,2,3,4, …θ} (θ is an integer) with replacement. Suppose the numbers we choose are 2,5,7, then based on MLE, what value of θ do you estimate?

Options:

5) 3
6) 5
7) 7
8) 9

Q 1.2 For the above example, if beforehand we know that Pr(θ=5) = 1/3, Pr(θ=8) = 1/6, Pr(θ=9) = 1/2. Then after we see the numbers we choose are 2,5,7, what value of θ do you think it is most likely to be?

Options:

1) 5
2) 7
3) 8
4) 9

Q 1.2 For the above example, if beforehand we know that Pr(θ=5) = 1/3, Pr(θ=8) = 1/6, Pr(θ=9) = 1/2. Then after we see the numbers we choose are 2,5,7, what value of θ do you think it is most likely to be?

Options:

1) 5
2) 7
3) 8
4) **9**

[Block 2: Naïve Bayes]

Q 2.1 Consider a classification problem with n = 32, $y \in \{1, 2, 3, \ldots, n\}$, and two binary features, $x_1, x_2 \in \{0,1\}$. Suppose $P(Y=y) = 1/32$, $P(x_1 = 1 | Y = y) = y / 46$, $P(x_2 = 1 | Y = y) = y / 62$. Which class will naive Bayes classifier produce on a test item with $x_1 = 1$ and $x_2 = 0$?
Options:

1) 16
2) 26
3) 31
4) 32

[Block 2: Naïve Bayes]

Q 2.1 Consider a classification problem with n = 32, $y \in \{1, 2, 3, \ldots, n\}$, and two binary features, $x_1, x_2 \in \{0,1\}$. Suppose $P(Y=y) = 1/32$, $P(x_1 = 1 | Y = y) = y / 46$, $P(x_2 = 1 | Y = y) = y / 62$. Which class will naive Bayes classifier produce on a test item with $x_1 = 1$ and $x_2 = 0$?

Options:

1) 16
2) 26
3) **31**
4) 32

Q 2.2 Consider the problem of detecting if an email message contains a virus. Say we use four random variables to model this problem: Boolean (binary) class variable V indicates if the message contains a virus or not, and three Boolean feature variables: A,B,C. We decide to use a Naive Bayes Classifier to solve this problem so we create a Bayesian network with arcs from V to each of A,B,C. Their associated CPTs (Conditional Probability Table) are created from the following data: $P\{V=1\} = 0.92$, $P\{A=1|V=1\} = 0.65$, $P\{A=1|V=0\} = 0.9$, $P\{B=1|V=1\} = 0.32$, $P\{B=1|V=0\} = 0.78$, $P\{C=1|V=1\} = 0.12$, $P\{C=1|V=0\} = 0.94$. Compute $P\{A=a,B=b,C=c\}$ for $a,b,c = 1, 0, 1$.

Options:

1) 0.0637
2) 0.0149
3) 0.0488
4) 0.0766

Q 2.2 Consider the problem of detecting if an email message contains a virus. Say we use four random variables to model this problem: Boolean (binary) class variable V indicates if the message contains a virus or not, and three Boolean feature variables: A,B,C. We decide to use a Naive Bayes Classifier to solve this problem so we create a Bayesian network with arcs from V to each of A,B,C. Their associated CPTs (Conditional Probability Table) are created from the following data: $P\{V=1\} = 0.92$, $P\{A=1|V=1\} = 0.65$, $P\{A=1|V=0\} = 0.9$, $P\{B=1|V=1\} = 0.32$, $P\{B=1|V=0\} = 0.78$, $P\{C=1|V=1\} = 0.12$, $P\{C=1|V=0\} = 0.94$. Compute $P\{A=a,B=b,C=c\}$ for $a,b,c = 1, 0, 1$.

Options:

1) **0.0637**
2) 0.0149
3) 0.0488
4) 0.0766

[Block 3: Various Naïve Bayes Models]

Q 3.1 Consider the below dataset showing the result whether a person is pass or fail in the exam based on various factors. We want to classify an instance 'X' with Confident=Yes, Studied=Yes and Sick=No. Suppose the factors are independent to each other.

Confident	Studied	Sick	Result
Yes	No	No	Fail
Yes	No	Yes	Pass
No	Yes	Yes	Fail
No	Yes	No	Pass
Yes	Yes	Yes	Pass

Options:

1. Pass
2. Fail

[Block 3: Various Naïve Bayes Models]

Q 3.1 Consider the below dataset showing the result whether a person is pass or fail in the exam based on various factors. We want to classify an instance 'X' with Confident=Yes, Studied=Yes and Sick=No. Suppose the factors are independent to each other.

Confident	Studied	Sick	Result
Yes	No	No	Fail
Yes	No	Yes	Pass
No	Yes	Yes	Fail
No	Yes	No	Pass
Yes	Yes	Yes	Pass

Options:

1. **Pass**
2. Fail

Q 3.2 In a lake, there are 2 kinds of fish, salmon and tilapia. Their lengths are both in Normal distribution: P(length=x | salmon) = $\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-3)^2}{2}\right)$, P(length=x | tilapia) = $\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-6)^2}{2}\right)$. Their weights are also both in Normal distribution: P(w = x | salmon) = $\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-9)^2}{2}\right)$, P(w = x | tilapia) = $\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-5)^2}{2}\right)$.
When you catch a fish, the chance to be a salmon is 0.8. If now you catch a fish with length 5 and weight 7, which fish do you think it would be? (suppose weight and length are independent).

Options:

1. Salmon
2. Tilapia

Q 3.2 In a lake, there are 2 kinds of fish, salmon and tilapia. Their lengths are both in Normal distribution: P(length=x | salmon) = $\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-3)^2}{2}\right)$, P(length=x | tilapia) = $\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-6)^2}{2}\right)$. Their weights are also both in Normal distribution: P(w = x | salmon) = $\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-9)^2}{2}\right)$, P(w = x | tilapia) = $\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-5)^2}{2}\right)$.
When you catch a fish, the chance to be a salmon is 0.8. If now you catch a fish with length 5 and weight 7, which fish do you think it would be? (suppose weight and length are independent).

Options:

1. Salmon
2. **Tilapia**