U.S.N.

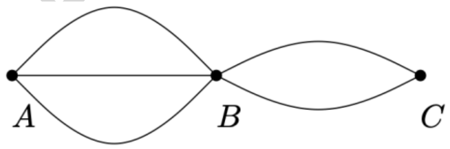
B.M.S. College of Engineering, Bengaluru-560019

Autonomous Institute Affiliated to VTU

April 2024 Semester End Main Examinations

Programme: B.E. **Semester: III**
Branch: Artificial Intelligence and Machine Learning **Duration: 3 hrs.**
Course Code: 23AM3PCPSM **Max Marks: 100**
Course: Probability and Statistics for Machine Learning

Instructions: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

		UNIT - I	***CO***	***PO***	**Marks**
1	a)	Traffic engineers have coordinated the timing of two traffic lights to encourage a run of green lights. In particular, the timing was designed so that with probability 0.8 a driver will find the second light to have the same color as the first. Assuming the first light is equally likely to be red or green: I. What is the probability P[G_2] that the second light is green? II. What is P[W], the probability that you wait for at least one of the first two lights?	*CO1*	*PO2*	**8**
	b)	Consider the following map which shows the roads between villages A, B & C.  Suppose that on a given winter day, a road is blocked with probability *p*. Assume that the conditions of the roads (i.e. whether they are open or blocked) are independent of each other. Compute the probability that you can go from A to C on a winter day.	*CO1*	*PO2*	**6**
	c)	Define the following terms. I. Joint probability II. Marginal Probability III. Conditional Probability	*CO1*	*PO1*	**6**

Important Note: Completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages. Revealing of identification, appeal to evaluator will be treated as malpractice.

<table>
<tr><td></td><td></td><td></td><td>UNIT - II</td><td></td><td></td><td></td></tr>
<tr><td></td><td>2</td><td>a)</td><td>The owner of a house has asked the neighbour to water a sick plant because the owner is going on a vacation. Without water it will die with a probability of 0.8, with water it will die with a probability of 0.15. It is 90% certain that the neighbour will remember to water the plant.
I. What is the probability that the plant will be alive when the owner return from vacation?
II. If the plant is dead, what is the probability that the neighbor forgot to water it?</td><td><i>CO2</i></td><td><i>PO3</i></td><td>6</td></tr>
<tr><td></td><td></td><td>b)</td><td>If A and B are two independent events with $P(A) = 0.4$, $P(B) = 0.2$, compute $P(A\cup B)$, $P(A|B)$, $P(B|A)$ and $P[(A\cap B)|B]$.</td><td><i>CO1</i></td><td><i>PO2</i></td><td>8</td></tr>
<tr><td></td><td></td><td>c)</td><td>State Bayes’ Theorem and elaborate on its applications.</td><td><i>CO1</i></td><td><i>PO2</i></td><td>6</td></tr>
<tr><td></td><td></td><td></td><td>OR</td><td></td><td></td><td></td></tr>
<tr><td></td><td>3</td><td>a)</td><td>A first step towards identifying spam is to create a list of words that are more likely to appear in spam than in normal messages. For instance, words like buy or the brand name of an enhancement drug are more likely to occur in spam messages than in normal messages. Suppose a specified list of words is available and that your database of 5000 messages contains 1700 that are spam. Among the spam messages, 1343 contain words in the list. Of the 3300 normal messages, only 297 contain words in the list. Obtain the probability that a message is spam given that the message contains words in the list.</td><td><i>CO1</i></td><td><i>PO3</i></td><td>8</td></tr>
<tr><td></td><td></td><td>b)</td><td>If A and B are two disjoint events with $P(A) = 0.4$, $P(B) = 0.2$, compute $P(A\cup B)$, $P(A|B)$, $P(B|A)$ and $P[(A\cap B)|B]$.</td><td><i>CO1</i></td><td><i>PO2</i></td><td>8</td></tr>
<tr><td></td><td></td><td>c)</td><td>I. How does MAP estimation contribute to Bayes Classifier?
II. It is known that a letter has come from either LONDON (or) CLIFTON. On the envelope just two consecutive letters 'ON' are visible. What is the probability that the letter has come from LONDON?</td><td><i>CO1</i></td><td><i>PO1</i></td><td>4</td></tr>
<tr><td></td><td></td><td></td><td>UNIT - III</td><td></td><td></td><td></td></tr>
<tr><td></td><td>4</td><td>a)</td><td>Every day, the number of network blackouts has a distribution
<table>
<tr><td>x</td><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>$P(X = x)$</td><td>0.5</td><td>k</td><td>0.1</td><td>k</td></tr>
</table>
A small internet trading company estimates that each network blackout results in a $500 loss. Find the value of k. Compute expectation and variance of this company’s daily loss due to blackouts.</td><td><i>CO2</i></td><td><i>PO3</i></td><td>8</td></tr>
<tr><td></td><td></td><td>b)</td><td>When a person goes for fishing, m hooks are attached to the fishing line. Each time the person cast the line, each hook will be swallowed by a fish with probability h, independent of whether any other hook is swallowed. What is the PMF of K, the number of fish that are hooked on a single cast of the line?</td><td><i>CO2</i></td><td><i>PO2</i></td><td>4</td></tr>
</table>

		c)	In a Poisson distribution frequency corresponding to the 3rd success is 2/3 times frequency corresponding to the 4th success. I. Derive the mean and variance of the distribution. II. Find the mean and standard deviation using the given data. III. Compute the probability of utmost three successes.	*CO2*	*PO1*	**8**
			OR			
	5	a)	Tossing a fair die is an experiment that can result in any integer number from 1 to 6 with probabilities such that the probability of getting an even number is twice that of the probability of getting an odd number. Let *X* be the number of dots on the top face of a die. Compute E(X) and Var(X), where Var is the variance of the associated random variable	*CO2*	*PO3*	**8**
		b)	The probability that an archer hits the target at any given attempt is 0.75. I. What is the probability that archer finally hits the target in his 4th attempt? II. Determine the type of distributions and justify.	*CO2*	*PO3*	**7**
		c)	Suppose *X* and *Y* are two independent random variables with E[*X*] = E[Y] = *k* and the variances Var[X] = Var[Y] = *s*, find the values of E[*X* - *Y*] and Var[*X* - *Y*].	*CO*	*PO1*	**5**
			UNIT - IV			
	6	a)	Let *X* be a continuous random variable with pdf $f_X(x) = ax^2 + bx$ in the interval $0 \leq x \leq 1$ and 0 elsewhere. Note that *a* and *b* are constants. Suppose , find the values of *a* and *b* and variance of *X*.	*CO3*	*PO1*	**10**
		b)	If *X* is a continuous random variable with PDF $f_X(x) = 3x^2$ in the interval (0, 1), find *a* such that $P(X \geq a) = P(X < a)$	*CO3*	*PO1*	**6**
		c)	Obtain the variance for a uniformly distributed random variable *X* in the interval (*a*, *b*).	*CO3*	*PO1*	**4**
			UNIT - V			
	7	a)	Define the following terms: I. Population II. Sample III. Interquartile Range	*CO3*	*PO1*	**6**
		b)	Compute the mean, variance and standard deviation for the following data: 1, 2, 3, 7, 17, 20, 6, 7, 8, 23, 16	*CO3*	*PO2*	**10**
		c)	The class average of 50 students in statistics exam is 20 marks out of 50. The professor decides to add 10 marks to each student to compensate for an incorrect question in the exam. I. Compute the new class average after adding 10 marks to each student? II. What will be the new standard deviation of the marks?	*CO3*	*PO2*	**4**

******