U.S.N.										

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

Autonomous Institute Affiliated to VTU

June 2025 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.

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.

		UNIT - I	*CO*	*PO*	Marks
1	a)	What are some real-world applications of probability in machine learning?	*CO2*	*PO1*	**05**
	b)	Let A_1, A_2, A_3 *and* A_4 be mutually independent events such that $P(A_1) = 1/2$, $P(A_2) = 1/4, P(A_3) = 1/8$ *and* $P(A_4) = 1/16$. What is the value of $P(A_1 \cup A_2) + P(A_3 \cup A_4) - P(A_1 \cap A_2 \cap A_3 \cap A_4)$	*CO1*	*PO2*	**08**
	c)	The joint probability distribution of X and Y is given below:	*CO1*	*PO2*	**07**

X \ Y	1	3	9
2	1/8	1/24	1/12
4	1/4	1/4	0
6	1/8	1/24	1/12

i. Find the marginal probability distribution of Y.
ii. Find the conditional distribution of Y given that X = 2
iii. Are X and Y independent?

		OR			
2	a)	Explain the intuition behind joint, marginal, and conditional probability.	*CO2*	*PO1*	**06**
	b)	The odds against Algorithm A successfully classifying data are 8:6, and the odds in favor of Algorithm B successfully classifying the same data are 14:16. i. What is the chance that neither algorithm successfully classifies the data, if they both try, independently of each other? ii. What is the probability that the data will be successfully classified by either algorithm?	*CO3*	*PO1*	**08**

		c)	Explain the following with an example: i. Random experiment ii. Event iii. Trials	*CO2*	*PO1*	**06**
			UNIT - II			
	3	a)	How do Maximum A Posteriori (MAP) and Maximum Likelihood Estimation (MLE) techniques differ?	*CO2*	*PO1*	**06**
		b)	Given three urns, where the first urn contains 7 white and 10 black balls, the second urn contains only black balls, and the third urn contains 17 white balls, a person randomly selects an urn and draws a ball, which turns out to be white. What is the total probability that the white ball drawn came from the first urn and the third urn, respectively?	*CO2*	*PO2*	**08**
		c)	How does the Gaussian Naive Bayes Classifier work, and what assumptions does it make about the features? What are its advantages and limitations?	*CO1*	*PO2*	**06**
			OR			
	4	a)	Use the naive Bayes method to determine if a loan with attributes X= (Home Owner =No, Marital Status=Married, Income=High) should be classified as a Defaulted Borrower or not. Determine which is larger, $P(Yes\|X)$ or $P(No\|X)$.	*CO3*	*PO3*	**10**

Tid	Home Owner	Marital Status	Annual Income	Defaulted Borrower
1	Yes	Single	High	**No**
2	No	Married	High	**No**
3	No	Single	Low	**No**
4	Yes	Married	High	**No**
5	No	Divorced	Low	**Yes**
6	No	Married	Low	**No**
7	Yes	Divorced	High	**No**
8	No	Single	Low	**Yes**
9	No	Married	Low	**No**
10	No	Single	Low	**Yes**

		b)	A medical test detects a rare disease affecting 1 in 1000 people. The test has a 99% accuracy rate. Using Bayes' theorem, explain prior, posterior, and likelihood probabilities. Calculate the probability that a person who tests positive actually has the disease.	*CO3*	*PO2*	**10**
			UNIT - III			
	5	a)	Define a random variable and explain its types with appropriate examples.	*CO1*	*PO1*	**06**
		b)	Calculate the expectation of the number of failures preceding the first success in an infinite series of independent trials, each with constant probability ‘p’ of success.	*CO1*	*PO2*	**08**

		c)	A player tosses two fair coins. He wins Rs.5 if two head occur, Rs. 2 if one head occurs and Re. 1 if no head occurs. Calculate the variance of the player's gain.	*CO1*	*PO3*	**06**
			OR			
	6	a)	Consider a survey where X represents customer satisfaction (1 for satisfied, 0 for not satisfied) and Y represents customer recommendation (1 for yes, 0 for no). The joint probability distribution is: $P(0,0) = 0.5$, $P(0,1) = 0.2$, $P(1,0) = 0.2$, $P(1,1) = 0.1$. i. Are X and Y independent? Explain. ii. Are (X+Y) and (X−Y) independent? Explain.	*CO2*	*PO2*	**10**
		b)	If X is a Poisson variable such that $P(X = 2) = 9P(X = 4)$ i. Compute the parameter λ. ii. Compute the mean and variance of X. iii. Calculate $P(X = 0)$.	*CO1*	*PO2*	**10**
			UNIT - IV			
	7	a)	The Life times of a certain kind of electronic devices have a mean of 300 hours and a S.D of 25 hours. Assuming that the distribution of these life times to be approximated closely to normal curve. i. Find the probability that any one of these electronic devices will have a life time of more than 350 hours ii. What percentage will have lifetimes of 300 hours or less? iii. What percentage will have lifetimes from 220 to 260 hours?	*CO3*	*PO2*	**10**
		b)	The waiting time (in minutes) for the next train follows a uniform distribution over the interval (a, b) with the probability density function $$f(x) = \frac{1}{b-a} \quad a < x < b$$ i. Determine the mean and variance of the waiting time. ii. For a specific case where a = 0 and b = 30 minutes, calculate the probability that a person will wait at least 20 minutes.	*CO2*	*PO2*	**10**
			OR			
	8	a)	A machine in a factory process orders sequentially, one after another. The time (in minutes) required to complete an order follows a continuous distribution with the probability density function: $f(x) = Kx^2$, $0 < x < 5$. i. Determine the value of K. ii. Compute the expectation (mean) processing time. iii. Find the probability that processing takes more than 4 minutes given that it has already exceeded 3 minutes.	*CO1*	*PO2*	**10**
		b)	In a credit card fraud detection system, 5% of transactions are classified as fraudulent. A sample of 3000 transactions is selected for analysis. i. State the Central Limit Theorem (CLT) and explain how it can be applied in this scenario. ii. Explain the concept of the continuity correction factor and why it is necessary when applying CLT to approximate probabilities from a discrete distribution.	*CO3*	*PO2*	**10**

			iii. Using CLT and the continuity correction factor, calculate the probability that the sample contains between 140 and 150 fraudulent transactions (inclusive of both limits).			
			UNIT - V			
	9	a)	Why do researchers often prefer sampling over studying the entire population?	*CO2*	*PO1*	**06**
		b)	Calculate the standard deviation from the following data:	*CO3*	*PO2*	**10**

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of Students	5	7	14	12	9	6	2

		c)	The weekly salaries (in ₹) of five employees are: 12,000, 15,000, 18,000, 22,000, 25,000. Compute the first quartile (Q_1) and the third quartile (Q_3).	*CO3*	*PO3*	**04**
			OR			
	10	a)	An economic analyst is studying the income distribution of a small town. After collecting income data from various households, the analyst needs to summarize this data using a measure of central tendency. Describe the characteristics of a good measure of central tendency, and identify which measure (mean, median, or mode) would be most appropriate for this data and why.	*CO1*	*PO1*	**06**
		b)	The heights (in cm) of five students are given as: 150, 155, 160, 165, 170. i. Compute the mean of the given data. ii. If each student's height is measured relative to 150 cm (i.e., subtract 150 from each value), find the mean of the transformed data. iii. Verify whether the mean remains unchanged due to this shift in origin.	*CO3*	*PO2*	**08**
		c)	A researcher analyses test scores of 200 students. She calculates the mean, median, and standard deviation and creates histograms. She then uses the data to predict overall student performance and tests score differences between two departments. 1. Identify the descriptive statistics used. 2. Identify the inferential statistics used. 3. Explain the difference between descriptive and inferential statistics using this example.	*CO3*	*PO2*	**06**

******