To be filled in by the examiner											
Question Number	1	2	3	4	5	6	7	8	9	10	**Total**
Marks Obtained											

Marks obtained (in words)	**Signature of the Examiner**	**Signature of the Scrutineer**

Indian Institute of Technology Kharagpur
Department of Computer Science and Engineering

Foundations of Algorithm Design and Machine Learning (CS60020)   **Spring 2025-2026**

22-February-2026 (AN)   **Mid-Semester Examination**   Maximum Marks: 60

Instructions:

- Write your answers in the question paper itself. Be brief and precise. Answer _all_ questions.

- Write the answers only in the respective spaces provided. The last three blank pages may be used for rough work or leftover answers.

- In case you may need more space/pages, please ask for additional sheets in the exam hall and attach the same with this booklet while submitting.

- If you use any algorithm / result / formula covered in the class, just mention it, do not elaborate (unless the same thing has been explicitly asked to answer in the question).

— Question Paper Starts from Next Page —

Q1. **[Recursive Formulation and Algorithm Complexity]** 12 marks

Consider an $n \times n$ 2-dimensional array called A, consisting of only 0s and 1s, where every row (from left to right) and every column (from top to bottom) have a sequence of 0s followed by a sequence of 1s. Some rows or columns may have only 0s and some rows or columns may have only 1s. However, in no case will we have a row or a column where a 0 follows a 1. You are to find the total number of 1s in the array A using an efficient algorithm. For this, answer the following questions.

(a) Present a recursive algorithm (using pseudo-code) to solve the problem efficiently. Clearly define the arguments of the recursive function, and explain them. Highlight the base and recursive portions of your algorithm with proper comments. Clearly indicate the return value. **(5)**

Solution:

(b) Show the working of your algorithm on the following example of a 5×5 array of elements. **(3)**

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Solution:

(c) Present (with a brief justification) the asymptotic time and space complexity of your algorithm. **(2+2)**
Solution:

Q2. **[Dynamic Programming]** $\boxed{\textbf{18 marks}}$

Consider the problem of finding the Longest Common Subsequence (LCS) from three lists, namely L_1, L_2 and L_3, where each list represents a sequence of characters. For this, answer the following questions.

(a) Present an example to show that solving the LCS of 3 lists cannot be guaranteed by finding the LCS of any 2 lists first and then using that result with the third list to get the final solution. **(4)**

Solution:

(b) Present the recursive formulation to solve the LCS of 3 lists. Explain the steps with comments. **(4)**

Solution:

(c) Convert the solution to a top-down dynamic programming algorithm (using *memoization*) and explain the algorithm with comments. Analyze its time and space complexity. **(4+2)**

Solution:

(d) Show the working of your proposed algorithm on an example of your choice with each list having at least 5 characters. **(4)**

Solution:

Q3. [Naïve Bayes Classifier]

A dataset of 6 students containing 4 attributes and 1 outcome, Placed (P), is shown on right. Note that, $\mathbb{P}(P = Yes) = \mathbb{P}(P = No) = \frac{1}{2}$. Here, Course-Type ($C$) and Attend-Class ($A$) are discrete attributes, whereas Study-Hours (S) and Marks (M) are continuous attributes. Assume that all continuous attributes follow normal (gaussian) distribution having probability density function,

Course Type	Attend Class	Study Hours	Marks (%)	Placed (Y/N)
Theory	*Average*	5	74	*No*
Theory	*Low*	10	70	*No*
Theory	*High*	8	92	*Yes*
Lab	*Average*	9	83	*Yes*
Lab	*High*	12	66	*No*
Lab	*Low*	16	87	*Yes*

$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, where μ and σ are the mean and standard deviation. Accordingly, the mean and variance of a set of n i.i.d. samples, $x_1, x_2, \ldots, x_n \sim p(x)$ can be computed (respectively) as follows:

$$[\,\texttt{mean}\,] \quad \mu = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad and \qquad [\,\texttt{variance}\,] \quad \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

Your task is to predict whether two new students, having the following attribute values, will be placed.

X_1 : $\langle$ Course-Type $= Theory$, Attend-Class $= Average$, Study-Hours $= 6$, Marks $= 75$ $\rangle$

X_2 : $\langle$ Course-Type $= Lab$, Attend-Class $= Low$, Study-Hours $= 9$, Marks $= 82$ $\rangle$

Applying (Gaussian) Naïve Bayes method, answer every step of computation as asked below.

(a) Compute the required conditional probabilities of the discrete attributes given the outcome. **(2)**

Solution:

$\mathbb{P}(C = Theory \mid P = Yes) \;=\; \underline{\hspace{3cm}}$ $\qquad$ $\mathbb{P}(A = Low \mid P = Yes) \;=\; \underline{\hspace{3cm}}$

$\mathbb{P}(C = Theory \mid P = No) \;=\; \underline{\hspace{3cm}}$ $\qquad$ $\mathbb{P}(A = Low \mid P = No) \;=\; \underline{\hspace{3cm}}$

$\mathbb{P}(C = Lab \mid P = Yes) \;=\; \underline{\hspace{3cm}}$ $\qquad$ $\mathbb{P}(A = Average \mid P = Yes) \;=\; \underline{\hspace{3cm}}$

$\mathbb{P}(C = Lab \mid P = No) \;=\; \underline{\hspace{3cm}}$ $\qquad$ $\mathbb{P}(A = Average \mid P = No) \;=\; \underline{\hspace{3cm}}$

(b) Compute the required mean and variance for the continuous attributes given the outcome. **(2)**

Solution:

$\mu_{S \mid P=Yes} \;=\; \underline{\hspace{3cm}}$ $\qquad$ $\sigma_{S \mid P=Yes} \;=\; \underline{\hspace{4cm}}$

$\mu_{S \mid P=No} \;=\; \underline{\hspace{3cm}}$ $\qquad$ $\sigma_{S \mid P=No} \;=\; \underline{\hspace{4cm}}$

$\mu_{M \mid P=Yes} \;=\; \underline{\hspace{3cm}}$ $\qquad$ $\sigma_{M \mid P=Yes} \;=\; \underline{\hspace{4cm}}$

$\mu_{M \mid P=No} \;=\; \underline{\hspace{3cm}}$ $\qquad$ $\sigma_{M \mid P=No} \;=\; \underline{\hspace{4cm}}$

(c) Derive the required conditional probabilities of the continuous attributes given the outcome, using the probability density function for the normal distribution. **(4)**

Solution:

$$\mathbb{P}(S = 6 \mid P = Yes) \quad = \quad \underline{\hspace{10cm}}$$

$$\mathbb{P}(S = 6 \mid P = No) \quad = \quad \underline{\hspace{10cm}}$$

$$\mathbb{P}(S = 9 \mid P = Yes) \quad = \quad \underline{\hspace{10cm}}$$

$$\mathbb{P}(S = 9 \mid P = No) \quad = \quad \underline{\hspace{10cm}}$$

$$\mathbb{P}(M = 75 \mid P = Yes) \quad = \quad \underline{\hspace{10cm}}$$

$$\mathbb{P}(M = 75 \mid P = No) \quad = \quad \underline{\hspace{10cm}}$$

$$\mathbb{P}(M = 82 \mid P = Yes) \quad = \quad \underline{\hspace{10cm}}$$

$$\mathbb{P}(M = 82 \mid P = No) \quad = \quad \underline{\hspace{10cm}}$$

(d) Derive the ratio of required conditional probabilities for the outcome $(P = Yes/No)$ for the given two new student data, namely X_1 and X_2, and indicate the outcome that a Naïve Bayes classifier would predict for these two new students? Show your calculations. **(2)**

Solution:

$$\frac{\mathbb{P}\Big(P = Yes \mid X_1 : \langle C = Theory,\ A = Average,\ S = 6,\ M = 75 \rangle\Big)}{\mathbb{P}\Big(P = No \mid X_1 : \langle C = Theory,\ A = Average,\ S = 6,\ M = 75 \rangle\Big)} =$$

$$\frac{\mathbb{P}\Big(P = Yes \mid X_2 : \langle C = Lab,\ A = Low,\ S = 9,\ M = 82 \rangle\Big)}{\mathbb{P}\Big(P = No \mid X_2 : \langle C = Lab,\ A = Low,\ S = 9,\ M = 82 \rangle\Big)} =$$

Q4. **[Bayesian Networks]**

Suppose that a patient can have a symptom (S) that can be caused by two different diseases (A and B). It is known that the variation of gene G plays a big role in the manifestation of disease A. The Bayes' Net and corresponding conditional probability tables for this situation are shown on right. The conditional and marginal probability values are given as,

$$\mathbb{P}(G) = 0.1, \quad \mathbb{P}(B) = 0.4, \quad \mathbb{P}(A \mid G) = 1.0, \quad \mathbb{P}(A \mid \neg G) = 0.1,$$

$$\mathbb{P}(S \mid A, B) = 1.0, \quad \mathbb{P}(S \mid A, \neg B) = 0.9, \quad \mathbb{P}(S \mid \neg A, B) = 0.8, \quad \mathbb{P}(S \mid \neg A, \neg B) = 1.0.$$
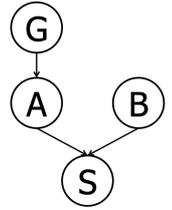
Answer the following. Show your calculations.

(a) Compute the probability of the joint disribution, that is, $\mathbb{P}(G, A, B, S)$. **(2)**

Solution:

(b) Find the probability that a patient has disease A. **(2)**

Solution:

(c) Find the probability that a patient has disease *A* given that they have disease *B*. **(1)**

Solution:

(d) Find the probability that a patient has disease *A* given that they have symptom *S* and disease *B*. **(2)**

Solution:

(e) Find the probability that a patient has the disease carrying gene variation G given that they have disease A. **(2)**

Solution:

(f) Find the probability that a patient has the disease carrying gene variation G given that they have disease B. **(1)**

Solution:

Q5. **[Classifier Evaluation]** 10 marks

You are given with the following data set of 15 samples, $D = \{d_0, d_2, \ldots, d_{14}\}$ with known actual class labels, $L(d_i) \in \{A, B, C\}$ for each of the sample $d_i \in D$ ($0 \le i \le 14$). In order to evaluate the quality of a classifier K, each sample $d_i \in D$ is additionally classified using K, yielding the predicted class label $K(d_i)$. The results are given in the table below.

$i \rightarrow$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Actual $\rightarrow L(d_i)$	A	B	A	C	C	B	A	A	A	B	B	C	C	C	B
Predicted $\rightarrow K(d_i)$	A	A	C	C	B	B	A	A	A	C	A	A	C	C	B

(a) Express the confusion matrix. (2)

Solution:

	Predicted A	Predicted B	Predicted C
Actual A	4	0	1
Actual B	2	2	1
Actual C	1	1	3

(b) Compute the accuracy. Show your calculations. (2)

Solution:

Accuracy $= \dfrac{\text{correct predictions}}{\text{total}} = \dfrac{4+2+3}{15} = \dfrac{9}{15} = 0.6$

(c) Compute Precision, Recall and F_1-score for each class $c \in \{A, B, C\}$. Show your calculations. **(2×3)**

Solution:

Precision:

Recall:

F_1-score:

— Question Paper Ends Here —