GATE Data Science and Artificial Intelligence

Sample Paper – 4

Duration: 128 Minutes **Maximum Marks: 72**

Instructions

- This paper contains **46 questions** covering the core **Data Science and Artificial Intelligence** subject of the GATE paper conducted by the IITs and IISc (General Aptitude and Engineering Mathematics are provided as separate papers).
- **Q1–Q20 carry 1 mark each** and **Q21–Q46 carry 2 marks each**, for a total of **72 marks**.
- Each question carries **negative marking of one-third of its allotted marks**: a wrong 1-mark question costs **0.33** and a wrong 2-mark question costs **0.67**. Unattempted questions carry no penalty.
- Every question here is a single-correct **MCQ**. The real GATE paper also has Multiple Select (MSQ) and Numerical Answer Type (NAT) questions, and **those carry no negative marking**.
- Attempt this paper in one timed sitting of about **128 minutes**, the share this core subject earns at GATE's overall pace of 180 minutes for 65 questions. All logarithms in information-theoretic quantities (entropy, Gini, information gain) are to base 2 and expressed in bits unless stated otherwise.

Probability & Statistics

Q1. Three distinct model-training jobs are to be assigned to 3 of the 6 available GPU servers, with each job going to a different server. The number of distinct assignments possible is: **[1 mark]**

(A) 20

(B) 120

(C) 216

(D) 18

Q2. For two *independent* events A and B with $P(A) = 0.5$ and $P(B) = 0.4$, the probability $P(A \cup B)$ is: **[1 mark]**

(A) 0.90

(B) 0.20

(C) 0.70

(D) 0.50

Q3. In a simple wager, a player wins $\texttt{Rs.}100$ with probability 0.2 and loses $\texttt{Rs.}20$ with probability 0.8. The expected gain per play is: **[1 mark]**

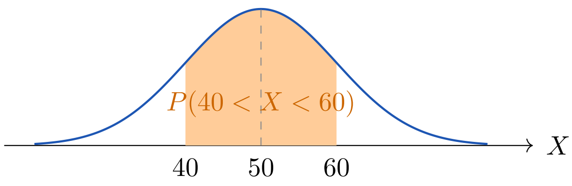
(A) $\texttt{Rs.}4$

(B) $\texttt{Rs.}20$

(C) $\texttt{Rs.}-4$

(D) $\texttt{Rs.}16$

Q4. A feature X is normally distributed with mean $\mu = 50$ and standard deviation $\sigma = 10$. Using the symmetric shaded band shown, the probability $P(40 < X < 60)$ is nearest to: **[1 mark]**



(A) 0.3413

(B) 0.6826

(C) 0.9544

(D) 0.5000

Q5. Requests arrive at a web server following a Poisson process with an average of $\lambda = 2$ requests per second. The probability of receiving *zero* requests in a given one-second window is nearest to (take $e^{-2} = 0.1353$): **[1 mark]**

(A) 0.135

(B) 0.271

(C) 0.865

(D) 0.500

Q6. The four data values 2, 4, 6, 8 form a small population. The population standard deviation of this data set is nearest to: **[1 mark]**

(A) 2.500

(B) 5.000

(C) 2.236

(D) 1.581

Q7. For a bivariate data set the covariance is $\mathrm{Cov}(X, Y) = 12$, and the standard deviations are $\sigma_X = 4$ and $\sigma_Y = 5$. The Pearson correlation coefficient r is: **[1 mark]**

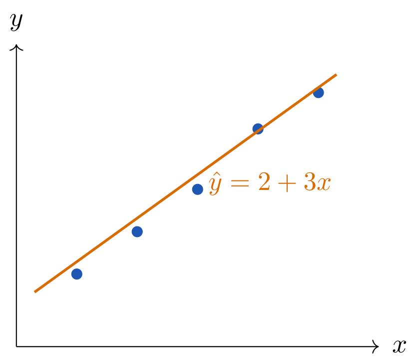
(A) 0.48

(B) 0.60

(C) 0.75

(D) 0.30

Q8. A least-squares regression line $\hat{y} = 2 + 3x$ is fitted to the scatter shown. The value predicted by the line at $x = 4$ is: **[1 mark]**



(A) 11

(B) 14

(C) 12

(D) 5

Q9. A fair coin is tossed 3 times independently. The probability of obtaining *at least one* head is: **[1 mark]**

(A) 0.875

(B) 0.125

(C) 0.500

(D) 0.375

Q10. A random response time X (in milliseconds) is uniformly distributed on the interval $[0, 10]$. The probability $P(X > 7)$ is: **[1 mark]**

(A) 0.70

(B) 0.50

(C) 0.10

(D) 0.30

Linear Algebra, Calculus & Optimization

Q11. Consider the matrix $A = \begin{bmatrix} 5 & 2 \\ 1 & 4 \end{bmatrix}$. The sum of the eigenvalues of A is: **[1 mark]**

(A) 9

(B) 18

(C) 20

(D) 7

Q12. The determinant of the upper-triangular matrix $B = \begin{bmatrix} 2 & 5 & 1 \\ 0 & 3 & 4 \\ 0 & 0 & 4 \end{bmatrix}$ is: **[1 mark]**

(A) 9

(B) 24

(C) 0

(D) 12

Q13. A square matrix is invertible only when its determinant is non-zero. The determinant of $C = \begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}$ is: **[1 mark]**

(A) 0

(B) 12

(C) -12

(D) 24

Q14. For the vector $v = (6, 8)$, the Euclidean (ℓ_2) norm $\|v\|_2$ is: **[1 mark]**

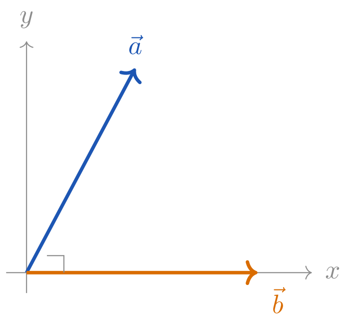
(A) 14

(B) 8

(C) 10

(D) 100

Q15. The two vectors $\vec{a} = (1, 2, 2)$ and $\vec{b} = (2, -1, 0)$ are drawn from a common origin. Their dot product $\vec{a} \cdot \vec{b}$ is: **[1 mark]**



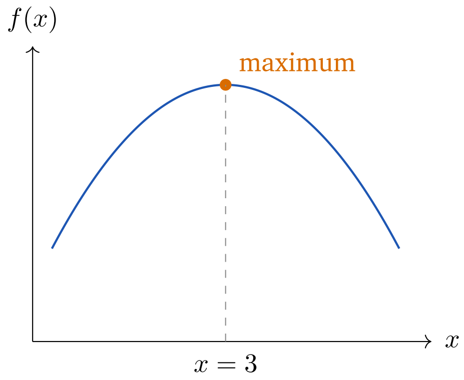
(A) 4

(B) 0

(C) 2

(D) -2

Q16. The function $f(x) = -x^2 + 6x + 1$ is to be *maximized* over the real line, as suggested by the concave graph below. The maximum *value* of f is: **[1 mark]**



(A) 3

(B) 1

(C) 10

(D) 0

Q17. The value of the definite integral $\int_0^2 3x^2\,dx$ is: **[1 mark]**

(A) 4

(B) 8

(C) 12

(D) 6

Q18. Gradient descent is used to minimize $f(x) = x^2$ with a fixed learning rate $\eta = 0.25$. Starting from $x_0 = 8$, the value of x_1 after one update $x_1 = x_0 - \eta\, f'(x_0)$ is: **[1 mark]**

(A) 3.2

(B) 7.5

(C) 6.0

(D) 4.0

Q19. For the function $f(x, y) = x^2 + 3y^2$, the gradient $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$ evaluated at the point $(1, 1)$ is: **[1 mark]**

(A) $(2, 6)$

(B) $(1, 3)$

(C) $(2, 3)$

(D) $(4, 6)$

Programming, Data Structures & Algorithms

Q20. What is the output of the following Python snippet?

```
print(2 ** 3 ** 2)
```

[1 mark]

(A) 64

(B) 12

(C) 256

(D) 512

Q21. Consider the following Python code:

```
x = [i*i for i in range(4)]
print(x[-1])
```

The value printed is: **[2 marks]**

(A) 0

(B) 1

(C) 4

(D) 9

Q22. The following operations are performed on an initially empty *queue* (first-in first-out), in order: `enqueue(5)`, `enqueue(10)`, `enqueue(15)`, `dequeue()`, `enqueue(20)`, `dequeue()`. The value returned by the *last* `dequeue()` is: **[2 marks]**

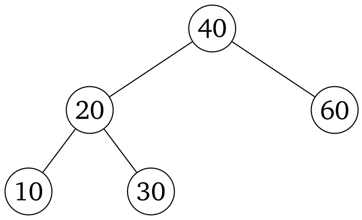
(A) 5

(B) 10

(C) 15

(D) 20

Q23. Starting from an empty binary search tree, the keys $40, 20, 60, 10, 30$ are inserted in the given order, producing the tree shown. The *height* of the resulting tree (measured as the number of edges on the longest root-to-leaf path) is: **[2 marks]**



(A) 2

(B) 3

(C) 4

(D) 5

Q24. A hash table of size 7 uses the hash function $h(k) = k \bmod 7$ with *separate chaining* for collision resolution. The keys $15, 22, 8$ are inserted in this order into an otherwise empty table. The number of keys stored in the chain at index 1 is: **[2 marks]**

(A) 0

(B) 1

(C) 2

(D) 3

Q25. Bubble sort is applied to an array of 5 elements that is initially in strictly *descending* order (the worst case). The total number of swaps performed is: **[2 marks]**

(A) 5

(B) 25

(C) 20

(D) 10

Q26. The worst-case time complexity of the standard Merge Sort algorithm on an array of n elements is: **[2 marks]**

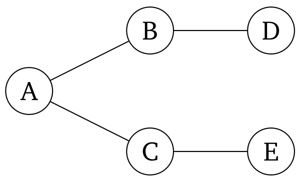
(A) $O(n^2)$

(B) $O(n)$

(C) $O(n \log n)$

(D) $O(\log n)$

Q27. A depth-first search (DFS) is started from vertex A in the undirected graph shown, always visiting adjacent vertices in increasing alphabetical order. The order in which the vertices are first visited is: **[2 marks]**



(A) A, B, C, D, E

(B) A, B, D, C, E

(C) A, C, E, B, D

(D) A, B, D, E, C

Q28. The running time of an algorithm satisfies the recurrence $T(n) = 2T(n/2) + 1$, with $T(1) = 1$. By the Master theorem, $T(n)$ is: **[2 marks]**

(A) $\Theta(n \log n)$

(B) $\Theta(\log n)$

(C) $\Theta(n)$

(D) $\Theta(n^2)$

Database Management & Warehousing

Q29. Relation R has 4 attributes and 6 tuples; relation S has 3 attributes and 5 tuples. The number of *attributes* (the degree) of the Cartesian product $R \times S$ is: **[2 marks]**

(A) 7

(B) 12

(C) 4

(D) 3

Q30. A table `EMP` has 10 rows. In the `salary` column, 2 of these rows contain `NULL` and the remaining 8 contain non-null values. What does the query `SELECT COUNT(*) FROM EMP;` return? **[2 marks]**

(A) 8

(B) 2

(C) 0

(D) 10

Q31. A relation $R(A, B, C)$ has the candidate key $\{A, B\}$ and the functional dependency $A \to C$ (a non-prime attribute depending on *part* of the key). The highest normal form that R satisfies is: **[2 marks]**

(A) BCNF

(B) 3NF

(C) 1NF

(D) 2NF

Q32. Two union-compatible relations R and S contain 5 and 3 tuples respectively, and they have exactly 2 tuples in common. The number of tuples in $R \cup S$ is: **[2 marks]**

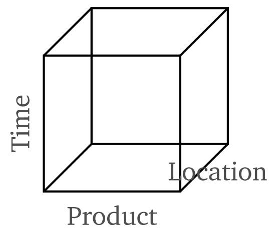
(A) 8

(B) 5

(C) 6

(D) 3

Q33. In the OLAP data cube shown (dimensions Time, Product, Location), the operation that navigates from a finer level of detail to a more *summarized* (coarser) level, such as from monthly totals to yearly totals, is called: **[2 marks]**



(A) drill-down

(B) slice

(C) pivot

(D) roll-up

Q34. A fact table in a data warehouse holds $10{,}000$ records. Each disk block can store 100 records, and records are packed fully. The number of disk blocks required to store the fact table is: **[2 marks]**

(A) $10{,}000$

(B) $1{,}000$

(C) 10

(D) 100

Q35. A relation has attributes $\{A, B, C, D\}$ with functional dependencies $A \to B$ and $B \to C$. The attribute closure A^+ is: **[2 marks]**

(A) $\{A, B, C\}$

(B) $\{A, B\}$

(C) $\{A, B, C, D\}$

(D) $\{A\}$

Machine Learning & Artificial Intelligence

Q36. A binary classifier is evaluated on 100 test examples, giving the confusion matrix shown (TP $= 40$, FP $= 10$, FN $= 20$, TN $= 30$). The classification *accuracy* is: **[2 marks]**

	Pred +	Pred −
Actual +	40	10
Actual −	20	30

(A) 0.70

(B) 0.80

(C) 0.60

(D) 0.50

Q37. A very simple linear model attains poor accuracy on *both* the training set and the test set. In terms of the bias–variance decomposition, this underfitting behaviour is best described as: **[2 marks]**

(A) low bias, high variance

(B) low bias, low variance

(C) high bias, low variance

(D) high bias, high variance

Q38. In a feature space, the Manhattan (ℓ_1) distance between the points $P(1, 2)$ and $Q(4, 6)$ is: **[2 marks]**

(A) 7

(B) 5

(C) 25

(D) 12

Q39. A training set for a binary classification task contains 10 examples, of which 5 are positive and 5 are negative. The entropy of this set (in bits) is: **[2 marks]**

(A) 0.500

(B) 0.000

(C) 1.000

(D) 0.811

Q40. A node in a decision tree contains 4 positive and 4 negative examples, as shown. The Gini impurity of this node is: **[2 marks]**

$4+,\ 4-$ (8 examples)

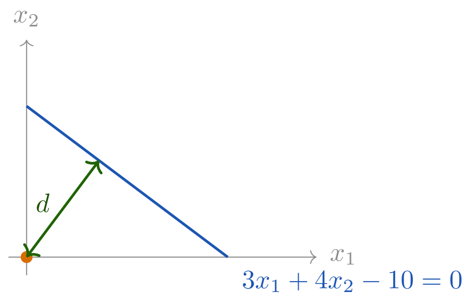
(A) 1.00

(B) 0.25

(C) 0.50

(D) 0.75

Q41. The decision boundary of a linear classifier is the line $3x_1 + 4x_2 - 10 = 0$, drawn below. The perpendicular distance of the origin $(0, 0)$ from this line is: **[2 marks]**



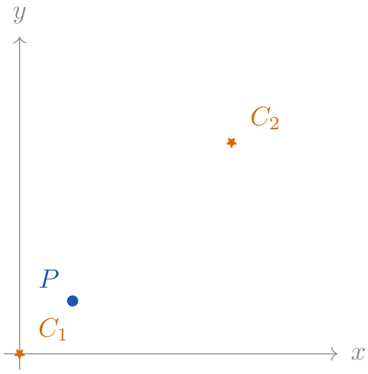
(A) 5

(B) 2

(C) 10

(D) 0.5

Q42. In one iteration of the k-means algorithm there are two cluster centroids, $C_1 = (0, 0)$ and $C_2 = (4, 4)$. The point $P = (1, 1)$, shown below, is assigned to the nearest centroid (Euclidean distance). Point P is assigned to: **[2 marks]**



(A) cluster C_2

(B) a tie (equidistant)

(C) cannot be determined

(D) cluster C_1

Q43. Principal component analysis is applied to a data set whose covariance matrix has eigenvalues $6, 2, 1, 1$. The fraction of the total variance explained by the *first* principal component alone is: **[2 marks]**

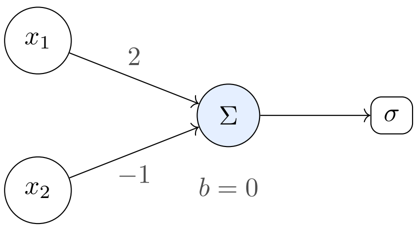
(A) 60%

(B) 80%

(C) 40%

(D) 20%

Q44. A single neuron computes the sigmoid activation $\sigma(z) = 1/(1+e^{-z})$ of its pre-activation $z = w^{\mathsf{T}}x + b$, as shown. With weights $w = (2, -1)$, inputs $x = (1, 2)$ and bias $b = 0$, the output of the neuron is: **[2 marks]**



(A) 1.00

(B) 0.73

(C) 0.00

(D) 0.50

Q45. In the A* search algorithm, a heuristic function $h(n)$ is said to be *admissible* when, for every node n, it never: **[2 marks]**

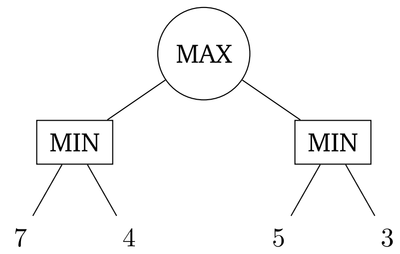
(A) equals the true cost to the goal

(B) overestimates the true cost to the goal

(C) underestimates the true cost to the goal

(D) ignores the true cost to the goal

Q46. In the two-ply game tree shown, the root is a MAX node and the second level are MIN nodes; the leaf values are given. The minimax value backed up to the root is: **[2 marks]**



(A) 3

(B) 4

(C) 5

(D) 7

Detailed Solutions

Q1.

Solution

Concept – ordered selection (permutation): The jobs are distinct and each server can host at most one job, so order matters; count arrangements $P(n,r) = \frac{n!}{(n-r)!}$.

Step 1 – choose a server for the first job: There are 6 available servers, so 6 choices.

Step 2 – choose a server for the second job: One server is used, leaving 5 choices.

Step 3 – choose a server for the third job: Two servers are used, leaving 4 choices.

Step 4 – multiply the independent choices: $6 \times 5 \times 4 = 120$

Final Answer: 120 assignments $\Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q1

Q2.

Solution

Concept – union of two events: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Step 1 – use independence for the intersection: For independent events, $P(A \cap B) = P(A)\,P(B)$.

Step 2 – compute the intersection: $P(A \cap B) = 0.5 \times 0.4 = 0.20$

Step 3 – substitute into the union formula: $P(A \cup B) = 0.5 + 0.4 - 0.20$

Step 4 – simplify: $P(A \cup B) = 0.9 - 0.20 = 0.70$

Final Answer: $P(A \cup B) = 0.70 \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q2

Q3.

Solution

Concept – expected value of a discrete outcome: $E[\text{gain}] = \sum(\text{value}) \times (\text{probability})$.

Step 1 – contribution of the winning outcome: $100 \times 0.2 = 20$

Step 2 – contribution of the losing outcome: $(-20) \times 0.8 = -16$

Step 3 – add the contributions: $E[\text{gain}] = 20 + (-16)$

Step 4 – simplify: $E[\text{gain}] = 4$

Final Answer: expected gain $= \texttt{Rs}.\ 4 \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q3

Q4.

Solution

Concept – standardize, then use the normal table: Convert X to $Z = \frac{X - \mu}{\sigma}$; the band 40 to 60 is symmetric about the mean.

Step 1 – standardize the lower cut-off $X = 40$: $Z = \frac{40 - 50}{10} = \frac{-10}{10} = -1$

Step 2 – standardize the upper cut-off $X = 60$: $Z = \frac{60 - 50}{10} = \frac{10}{10} = 1$

Step 3 – write the required probability in Z: $P(40 < X < 60) = P(-1 < Z < 1)$

Step 4 – use the standard normal values: $P(Z \le 1) = 0.8413$ and $P(Z \le -1) = 0.1587$

Step 5 – subtract: $P(-1 < Z < 1) = 0.8413 - 0.1587 = 0.6826$

Final Answer: $P(40 < X < 60) = 0.6826 \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q4

Q5.

Solution

Concept – Poisson probability: $P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$.

Step 1 – list the parameters: $\lambda = 2$ per second, $k = 0$.

Step 2 – substitute into the formula: $P(X = 0) = \frac{e^{-2}\,2^0}{0!}$

Step 3 – evaluate the power and factorial: $2^0 = 1$ and $0! = 1$

$$P(X=0) = \frac{e^{-2} \times 1}{1} = e^{-2}$$

Step 4 – substitute $e^{-2} = 0.1353$: $P(X=0) = 0.135$

Final Answer: $P(X=0) \approx 0.135 \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q5

Q6.

Solution

Concept – population standard deviation: $\sigma = \sqrt{\frac{1}{N}\sum(x_i - \bar{x})^2}$.

Step 1 – compute the mean: $\bar{x} = \frac{2+4+6+8}{4} = \frac{20}{4} = 5$

Step 2 – compute the deviations from the mean: $2-5=-3, \quad 4-5=-1, \quad 6-5=1, \quad 8-5=3$

Step 3 – square the deviations: $(-3)^2 = 9, \quad (-1)^2 = 1, \quad 1^2 = 1, \quad 3^2 = 9$

Step 4 – sum the squared deviations: $9+1+1+9=20$

Step 5 – divide by $N=4$ to get the variance: $\sigma^2 = \frac{20}{4} = 5$

Step 6 – take the square root: $\sigma = \sqrt{5} = 2.236$

Final Answer: $\sigma \approx 2.236 \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q6

Q7.

Solution

Concept – Pearson correlation from covariance: $r = \frac{\text{Cov}(X,Y)}{\sigma_X\,\sigma_Y}$.

Step 1 – compute the denominator: $\sigma_X\,\sigma_Y = 4 \times 5 = 20$

Step 2 – substitute into the formula: $r = \frac{12}{20}$

Step 3 – simplify: $r = 0.60$

Step 4 – sanity check the range: $-1 \le r \le 1$, and 0.60 lies inside this range, so the value is valid.

Final Answer: $r = 0.60 \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q7

Q8.

Solution

Concept – prediction from a regression line: Substitute the given x into $\hat{y} = 2 + 3x$.

Step 1 – write the line with $x = 4$: $\hat{y} = 2 + 3(4)$

Step 2 – evaluate the product: $3 \times 4 = 12$

Step 3 – add the intercept: $\hat{y} = 2 + 12 = 14$

Final Answer: predicted $\hat{y} = 14 \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q8

Q9.

Solution

Concept – complement rule: $P(\text{at least one head}) = 1 - P(\text{no head})$.

Step 1 – probability of a tail on one toss: $P(\text{tail}) = \frac{1}{2}$

Step 2 – probability of all three tosses being tails: $P(\text{no head}) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$

Step 3 – take the complement: $P(\text{at least one head}) = 1 - \frac{1}{8} = \frac{7}{8}$

Step 4 – convert to a decimal: $\frac{7}{8} = 0.875$

Final Answer: $P(\text{at least one head}) = 0.875 \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q9

Q10.

Solution

Concept – uniform distribution probability: For $X \sim \text{Uniform}(a, b)$, $P(X > c) = \frac{b-c}{b-a}$.

Step 1 – identify the interval: $a = 0$, $b = 10$, so the total length is $b - a = 10$.

Step 2 – length of the favourable region $X > 7$: $b - c = 10 - 7 = 3$

Step 3 – form the ratio: $P(X > 7) = \frac{3}{10}$

Step 4 – convert to a decimal: $P(X > 7) = 0.30$

Final Answer: $P(X > 7) = 0.30 \Rightarrow \boxed{\text{D}}$

$\boxed{\text{Answer: (D)}}$ Go Back to Q10

Q11.

Solution

Concept – sum of eigenvalues equals the trace: For any square matrix, $\lambda_1 + \lambda_2 + \cdots = \text{tr}(A)$, the sum of the diagonal entries.

Step 1 – identify the diagonal entries of A: The diagonal entries are 5 and 4.

Step 2 – add them: $\text{tr}(A) = 5 + 4 = 9$

Step 3 – conclude: The sum of the eigenvalues equals the trace, so it is 9.

Final Answer: sum of eigenvalues $= 9 \Rightarrow \boxed{\text{A}}$

$\boxed{\text{Answer: (A)}}$ Go Back to Q11

Q12.

Solution

Concept – determinant of a triangular matrix: The determinant of an upper- (or lower-) triangular matrix is the product of its diagonal entries.

Step 1 – identify the diagonal entries of B: The diagonal entries are 2, 3 and 4.

Step 2 – multiply the first two: $2 \times 3 = 6$

Step 3 – multiply by the third: $6 \times 4 = 24$

Final Answer: $\det B = 24 \Rightarrow \boxed{\text{B}}$

$\boxed{\text{Answer: (B)}}$ Go Back to Q12

Q13.

Solution

Concept – 2×2 determinant: $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$

Step 1 – form the main-diagonal product: $ad = 2 \times 6 = 12$

Step 2 – form the anti-diagonal product: $bc = 4 \times 3 = 12$

Step 3 – subtract: $\det C = 12 - 12 = 0$

Step 4 – interpret: Since $\det C = 0$, the matrix is singular (not invertible); note that row 2 is exactly 1.5 times row 1.

Final Answer: $\det C = 0 \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q13

Q14.

Solution

Concept – Euclidean norm: $\|v\|_2 = \sqrt{v_1^2 + v_2^2}$.

Step 1 – square the components: $6^2 = 36$ and $8^2 = 64$

Step 2 – add the squares: $36 + 64 = 100$

Step 3 – take the square root: $\|v\|_2 = \sqrt{100} = 10$

Final Answer: $\|v\|_2 = 10 \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q14

Q15.

Solution

Concept – dot product: $\vec{a} \cdot \vec{b} = a_1b_1 + a_2b_2 + a_3b_3$.

Step 1 – multiply the corresponding components: $a_1b_1 = 1 \times 2 = 2$

$a_2b_2 = 2 \times (-1) = -2$

$a_3b_3 = 2 \times 0 = 0$

Step 2 – add the products: $\vec{a} \cdot \vec{b} = 2 + (-2) + 0 = 0$

Step 3 – interpret: Since the dot product is 0, the vectors are orthogonal (the right angle shown in the figure is consistent).

Final Answer: $\vec{a} \cdot \vec{b} = 0 \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q15

Q16.

Solution

Concept – maximize by setting the derivative to zero: For a concave parabola, the stationary point is the global maximum.

Step 1 – differentiate: $f'(x) = -2x + 6$

Step 2 – set the derivative to zero: $-2x + 6 = 0$

$x = 3$

Step 3 – confirm it is a maximum: $f''(x) = -2 < 0$, so the point is a maximum.

Step 4 – evaluate f at $x = 3$: $f(3) = -(3)^2 + 6(3) + 1$

$= -9 + 18 + 1$

$= 10$

Final Answer: maximum value $= 10 \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q16

Q17.

Solution

Concept – definite integral of a power: $\int x^n\, dx = \dfrac{x^{n+1}}{n+1}$, then evaluate between limits.

Step 1 – find the antiderivative: $\int 3x^2\, dx = 3 \cdot \dfrac{x^3}{3} = x^3$

Step 2 – evaluate at the upper limit $x = 2$: $2^3 = 8$

Step 3 – evaluate at the lower limit $x = 0$: $0^3 = 0$

Step 4 – subtract: $\int_0^2 3x^2\, dx = 8 - 0 = 8$

Final Answer: the integral $= 8 \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q17

Q18.

Solution

Concept – one gradient-descent update: $x_1 = x_0 - \eta\, f'(x_0)$, with $f(x) = x^2$ so $f'(x) = 2x$.

Step 1 – evaluate the gradient at $x_0 = 8$: $f'(8) = 2 \times 8 = 16$

Step 2 – multiply by the learning rate: $\eta\, f'(x_0) = 0.25 \times 16 = 4$

Step 3 – subtract from the current point: $x_1 = 8 - 4 = 4$

Final Answer: $x_1 = 4 \Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q18

Q19.

Solution

Concept – gradient of a multivariable function: $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$.

Step 1 – partial derivative with respect to x: $\frac{\partial f}{\partial x} = 2x$

Step 2 – partial derivative with respect to y: $\frac{\partial f}{\partial y} = 6y$

Step 3 – evaluate each at $(1, 1)$: $2x = 2(1) = 2$ and $6y = 6(1) = 6$

Step 4 – assemble the gradient vector: $\nabla f(1, 1) = (2,\ 6)$

Final Answer: $\nabla f(1, 1) = (2, 6) \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q19

Q20.

Solution

Concept – the $**$ operator is right-associative in Python: $2 ** 3 ** 2$ is evaluated as $2 ** (3 ** 2)$, not $(2 ** 3) ** 2$.

Step 1 – evaluate the top exponent first: $3 ** 2 = 3^2 = 9$

Step 2 – apply the base to that result: $2 ** 9 = 2^9$

Step 3 – evaluate the power: $2^9 = 512$

Step 4 – note the wrong grouping: Left grouping would give $(2^3)^2 = 8^2 = 64$, which is *not* how Python evaluates it.

Final Answer: the output is $512 \Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q20

Q21.

Solution

Concept – list comprehension and negative indexing: Build the list, then read the last element with index -1.

Step 1 – expand `range(4)`: It yields $0,\ 1,\ 2,\ 3$.

Step 2 – apply $i * i$ to each value: $0 \times 0 = 0, \quad 1 \times 1 = 1, \quad 2 \times 2 = 4, \quad 3 \times 3 = 9$

Step 3 – form the list: $x = [0,\ 1,\ 4,\ 9]$

Step 4 – read index -1 (the last element): $x[-1] = 9$

Final Answer: the value printed is $9 \Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q21

Q22.

Solution

Concept – a queue is first-in first-out (FIFO): The element removed by `dequeue` is always the one that has waited longest.

Step 1 – state after the three enqueues: Front $\rightarrow [5, 10, 15] \rightarrow$ rear.

Step 2 – first `dequeue`: Removes the front element 5; queue is now $[10, 15]$.

Step 3 – `enqueue(20)`: Adds 20 at the rear; queue is now $[10, 15, 20]$.

Step 4 – second (last) `dequeue`: Removes the current front element, which is 10.

Final Answer: the last `dequeue` returns $10 \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q22

Q23.

Solution

Concept – height of a BST is the longest root-to-leaf edge count: Insert the keys and trace the deepest path.

Step 1 – insert 40: It becomes the root.

Step 2 – insert 20 and 60: $20 < 40$ goes left of the root; $60 > 40$ goes right of the root.

Step 3 – insert 10 and 30: $10 < 40$ then $10 < 20$, so 10 becomes the left child of 20; $30 < 40$ then $30 > 20$, so 30 becomes the right child of 20.

Step 4 – find the longest path: Root $40 \rightarrow 20 \rightarrow 10$ (or $40 \rightarrow 20 \rightarrow 30$) uses 2 edges; the path to 60 uses only 1 edge.

Step 5 – read the height: The longest root-to-leaf path has 2 edges, so the height is 2.

Final Answer: height $= 2 \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q23

Q24.

Solution

Concept – separate chaining groups all colliding keys in a linked list at one index: Compute $h(k) = k \bmod 7$ for each key.

Step 1 – hash key 15**:** $15 \bmod 7 = 1$ (since $15 = 2 \times 7 + 1$).

Step 2 – hash key 22**:** $22 \bmod 7 = 1$ (since $22 = 3 \times 7 + 1$).

Step 3 – hash key 8**:** $8 \bmod 7 = 1$ (since $8 = 1 \times 7 + 1$).

Step 4 – collect the chain at index 1**:** All three keys $15, 22, 8$ hash to index 1, so its chain holds 3 keys.

Final Answer: chain length at index 1 is $3 \Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q24

Q25.

Solution

Concept – worst-case swaps in bubble sort: A fully reversed array requires a swap for every inverted pair; the number of swaps equals $\frac{n(n-1)}{2}$.

Step 1 – identify n**:** $n = 5$ elements.

Step 2 – substitute into the formula: $\frac{n(n-1)}{2} = \frac{5 \times 4}{2}$

Step 3 – evaluate the numerator: $5 \times 4 = 20$

Step 4 – divide: $\frac{20}{2} = 10$

Final Answer: number of swaps $= 10 \Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q25

Q26.

Solution

Concept – merge sort divides and merges: It splits the array into halves ($\log_2 n$ levels) and merges each level in linear time.

Step 1 – count the levels of recursion: Halving n repeatedly gives about $\log_2 n$ levels.

Step 2 – cost of merging at each level: Merging all sub-arrays at a level touches every element once, costing $O(n)$.

Step 3 – multiply levels by per-level cost: $O(n) \times O(\log n) = O(n \log n)$.

Step 4 – note it is the same in all cases: Merge sort's split-and-merge cost does not depend on the input order, so worst case is also $O(n \log n)$.

Final Answer: worst-case time $= O(n \log n) \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q26

Q27.

Solution

Concept – DFS goes as deep as possible before backtracking: At each vertex, visit the smallest unvisited neighbour first.

Step 1 – start at A; list its neighbours: A is adjacent to B and C; the smaller is B, so go to B.

Step 2 – from B, go deeper: B is adjacent to A (visited) and D; visit D.

Step 3 – from D, backtrack: D has no unvisited neighbour, so backtrack to B, then to A.

Step 4 – from A, take the next branch: The next unvisited neighbour of A is C; visit C.

Step 5 – from C, go deeper: C is adjacent to E; visit E.

Step 6 – assemble the visit order: A, B, D, C, E.

Final Answer: order $= A, B, D, C, E \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q27

Q28.

Solution

Concept – Master theorem for $T(n) = aT(n/b) + f(n)$: Compare $f(n)$ with $n^{\log_b a}$.

Step 1 – read off the parameters: $a = 2$, $b = 2$, $f(n) = 1 = \Theta(1)$.

Step 2 – compute the critical exponent: $\log_b a = \log_2 2 = 1$, so $n^{\log_b a} = n^1 = n$.

Step 3 – compare $f(n)$ with n: $f(n) = \Theta(1)$ grows slower than n^1, so Case 1 applies.

Step 4 – conclude from Case 1: $T(n) = \Theta(n^{\log_b a}) = \Theta(n)$.

Final Answer: $T(n) = \Theta(n) \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q28

Q29.

Solution

Concept – degree of a Cartesian product: The Cartesian product concatenates the attributes of both relations, so its degree is the *sum* of the two degrees (tuple counts do not affect the degree).

Step 1 – degree of R: R has 4 attributes.

Step 2 – degree of S: S has 3 attributes.

Step 3 – add the degrees: $4 + 3 = 7$

Step 4 – note on the tuple count: The number of tuples would be $6 \times 5 = 30$, but the question asks for the degree, which is 7.

Final Answer: degree of $R \times S$ is $7 \Rightarrow \boxed{\text{A}}$

Answer: (A) **Go Back to Q29**

Q30.

Solution

Concept – `COUNT(*)` versus `COUNT(column)`: `COUNT(*)` counts *all* rows, including those with `NULL` in any column; only `COUNT(column)` skips `NULL`s in that column.

Step 1 – total number of rows in `EMP`: There are 10 rows.

Step 2 – does `COUNT(*)` skip the `NULL` salaries? No; `COUNT(*)` counts rows, not column values, so the 2 `NULL`-salary rows are still counted.

Step 3 – state the result: `COUNT(*)` $= 10$

Step 4 – contrast: `COUNT(salary)` would return 8 (skipping the 2 `NULL`s), but that is not what was asked.

Final Answer: the query returns $10 \Rightarrow \boxed{\text{D}}$

Answer: (D) **Go Back to Q30**

Q31.

Solution

Concept – 2NF forbids partial dependencies: 2NF requires that no non-prime attribute depend on only *part* of a composite candidate key.

Step 1 – identify the candidate key and non-prime attribute: The candidate key is $\{A, B\}$; the non-prime attribute is C.

Step 2 – examine the dependency $A \to C$: C depends on A alone, which is only *part* of the key $\{A, B\}$; this is a partial dependency.

Step 3 – conclude which normal form fails: A partial dependency violates 2NF, so R is *not* in 2NF.

Step 4 – state the highest normal form satisfied: Every relation with atomic attributes is at least in 1NF, and R cannot rise above that here, so the highest normal form is 1NF.

Final Answer: highest normal form $= 1\text{NF} \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q31

Q32.

Solution

Concept – cardinality of a set union: $|R \cup S| = |R| + |S| - |R \cap S|$ (subtract the common tuples counted twice).

Step 1 – list the given counts: $|R| = 5$, $|S| = 3$, $|R \cap S| = 2$.

Step 2 – add the two relation sizes: $5 + 3 = 8$

Step 3 – subtract the common tuples: $8 - 2 = 6$

Final Answer: $|R \cup S| = 6 \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q32

Q33.

Solution

Concept – OLAP roll-up: Roll-up aggregates a data cube from a finer level to a coarser level along a dimension hierarchy (for example month $\rightarrow$ year).

Step 1 – direction of the operation: Going from monthly totals to yearly totals means moving to *less* detail, i.e. upward in the hierarchy.

Step 2 – name that direction: Moving to less detail is roll-up; moving to more detail would be drill-down.

Step 3 – eliminate the other options: Slice fixes one dimension value, and pivot rotates the cube's axes; neither changes the level of aggregation.

Final Answer: the operation is roll-up $\Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q33

Q34.

Solution

Concept – blocking factor: Number of blocks $= \left\lceil \dfrac{\text{total records}}{\text{records per block}} \right\rceil$.

Step 1 – list the given values: Total records $= 10{,}000$; records per block $= 100$.

Step 2 – divide: $\dfrac{10{,}000}{100} = 100$

Step 3 – apply the ceiling: The division is exact, so the ceiling leaves 100.

Final Answer: blocks required $= 100 \Rightarrow \boxed{\text{D}}$

Answer: (D) **Go Back to Q34**

Q35.

Solution

Concept – attribute closure: A^+ is the set of all attributes functionally determined by A, found by repeatedly applying the FDs.

Step 1 – start with the attribute itself: $A^+ = \{A\}$

Step 2 – apply $A \to B$: Since $A \in A^+$, add B: $A^+ = \{A, B\}$

Step 3 – apply $B \to C$: Since $B \in A^+$, add C: $A^+ = \{A, B, C\}$

Step 4 – check for D: No functional dependency has D on its right-hand side reachable from these attributes, so D is not added.

Step 5 – state the closure: $A^+ = \{A, B, C\}$

Final Answer: $A^+ = \{A, B, C\} \Rightarrow \boxed{\text{A}}$

Answer: (A) **Go Back to Q35**

Q36.

Solution

Concept – accuracy from a confusion matrix: Accuracy $= \dfrac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}$.

Step 1 – sum the correct predictions (diagonal): $\text{TP} + \text{TN} = 40 + 30 = 70$

Step 2 – sum all predictions: $40 + 10 + 20 + 30 = 100$

Step 3 – form the ratio: Accuracy $= \dfrac{70}{100}$

Step 4 – simplify: Accuracy $= 0.70$

Final Answer: accuracy $= 0.70 \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q36

Q37.

Solution

Concept – underfitting in the bias–variance picture: A model too simple to capture the pattern makes consistent but large errors, so it has high bias; because it barely changes with the data, its variance is low.

Step 1 – interpret poor training accuracy: Failing even on the training data signals the model cannot represent the relationship, i.e. high bias.

Step 2 – interpret similar poor test accuracy: Training and test errors are both poor and close together, meaning the model is stable across samples, i.e. low variance.

Step 3 – combine: High bias together with low variance is the signature of underfitting.

Final Answer: high bias, low variance $\Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q37

Q38.

Solution

Concept – Manhattan (ℓ_1) distance: $d_1(P,Q) = |x_2 - x_1| + |y_2 - y_1|$.

Step 1 – absolute difference in the x-coordinates: $|4-1| = 3$

Step 2 – absolute difference in the y-coordinates: $|6-2| = 4$

Step 3 – add the two: $d_1 = 3 + 4 = 7$

Step 4 – contrast with Euclidean distance: The straight-line (ℓ_2) distance would be $\sqrt{3^2+4^2} = 5$; the Manhattan value asked for is 7.

Final Answer: $d_1(P,Q) = 7 \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q38

Q39.

Solution

Concept – binary entropy: $H = -p\log_2 p - q\log_2 q$, where p and q are the class proportions.

Step 1 – compute the proportions: $p = \frac{5}{10} = 0.5$ and $q = \frac{5}{10} = 0.5$.

Step 2 – evaluate each logarithm: $\log_2 0.5 = -1$, so $-p\log_2 p = -0.5(-1) = 0.5$ for each class.

Step 3 – add the two contributions: $H = 0.5 + 0.5 = 1.0$

Step 4 – interpret: A perfectly balanced two-class set has the maximum possible entropy of 1 bit.

Final Answer: $H = 1.000$ bit $\Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q39

Q40.

Solution

Concept – Gini impurity: $G = 1 - \sum_i p_i^2$, where p_i are the class proportions.

Step 1 – compute the class proportions: $p_+ = \frac{4}{8} = 0.5$ and $p_- = \frac{4}{8} = 0.5$.

Step 2 – square each proportion: $0.5^2 = 0.25$ for each class.

Step 3 – sum the squares: $0.25 + 0.25 = 0.5$

Step 4 – subtract from 1: $G = 1 - 0.5 = 0.5$

Final Answer: Gini impurity $= 0.50 \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q40

Q41.

Solution

Concept – distance from a point to a line: For the line $ax + by + c = 0$, the distance of the point (x_0, y_0) is $d = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}$.

Step 1 – identify the coefficients: $a = 3$, $b = 4$, $c = -10$, and the point is the origin $(0, 0)$.

Step 2 – evaluate the numerator: $|3(0) + 4(0) - 10| = |-10| = 10$

Step 3 – evaluate the denominator: $\sqrt{3^2 + 4^2} = \sqrt{9 + 16} = \sqrt{25} = 5$

Step 4 – divide: $d = \frac{10}{5} = 2$

Final Answer: distance $= 2 \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q41

Q42.

Solution

Concept – k-means assigns each point to the nearest centroid: Compute the Euclidean distance from P to each centroid and pick the smaller.

Step 1 – distance from $P = (1, 1)$ to $C_1 = (0, 0)$: $d_1 = \sqrt{(1-0)^2 + (1-0)^2} = \sqrt{1+1} = \sqrt{2} \approx 1.41$

Step 2 – distance from $P = (1, 1)$ to $C_2 = (4, 4)$: $d_2 = \sqrt{(1-4)^2 + (1-4)^2} = \sqrt{9+9} = \sqrt{18} \approx 4.24$

Step 3 – compare the distances: $1.41 < 4.24$, so C_1 is nearer.

Step 4 – assign the point: P joins the cluster of C_1.

Final Answer: P is assigned to cluster $C_1 \Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q42

Q43.

Solution

Concept – variance explained by a principal component: Each eigenvalue is the variance along its component; the fraction explained is that eigenvalue divided by the total.

Step 1 – sum all eigenvalues (total variance): $6 + 2 + 1 + 1 = 10$

Step 2 – take the largest eigenvalue (first PC): The first principal component corresponds to $\lambda_1 = 6$.

Step 3 – form the ratio: $\frac{6}{10} = 0.60$

Step 4 – express as a percentage: $0.60 = 60\%$

Final Answer: first PC explains $60\% \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q43

Q44.

Solution

Concept – neuron output with a sigmoid activation: First form the pre-activation $z = w^{\mathsf{T}}x + b$, then apply $\sigma(z) = \dfrac{1}{1+e^{-z}}$.

Step 1 – weighted sum of the inputs: $w^{\mathsf{T}}x = (2)(1) + (-1)(2) = 2 - 2 = 0$

Step 2 – add the bias: $z = 0 + 0 = 0$

Step 3 – apply the sigmoid at $z = 0$: $\sigma(0) = \dfrac{1}{1+e^{0}} = \dfrac{1}{1+1}$

Step 4 – simplify: $\sigma(0) = \dfrac{1}{2} = 0.50$

Final Answer: neuron output $= 0.50 \Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q44

Q45.

Solution

Concept – admissibility of an A* heuristic: A heuristic $h(n)$ is admissible if it never *overestimates* the true (optimal) cost from n to the goal; it must satisfy $h(n) \le h^*(n)$.

Step 1 – recall the guarantee A* needs: A* returns an optimal path when h never exceeds the real remaining cost.

Step 2 – translate "never exceeds": This is exactly the condition that h never overestimates the true cost to the goal.

Step 3 – eliminate the other options: Underestimating is allowed (even $h = 0$ is admissible); equalling the cost is also allowed; "ignores" is not a defined condition.

Final Answer: it never overestimates the true cost $\Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q45

Q46.

Solution

Concept – minimax back-up: MIN nodes take the minimum of their children; the MAX root takes the maximum of the MIN results.

Step 1 – evaluate the left MIN node: $\min(7, 4) = 4$

Step 2 – evaluate the right MIN node: $\min(5, 3) = 3$

Step 3 – evaluate the MAX root: $\max(4, 3) = 4$

Final Answer: minimax value $= 4 \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q46

Answer Key

Q	Ans	Q	Ans	Q	Ans	Q	Ans	Q	Ans
1	B	2	C	3	A	4	B	5	A
6	C	7	B	8	B	9	A	10	D
11	A	12	B	13	A	14	C	15	B
16	C	17	B	18	D	19	A	20	D
21	D	22	B	23	A	24	D	25	D
26	C	27	B	28	C	29	A	30	D
31	C	32	C	33	D	34	D	35	A
36	A	37	C	38	A	39	C	40	C
41	B	42	D	43	A	44	D	45	B
46	B								