

GATE Data Science and Artificial Intelligence

Sample Paper – 5

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

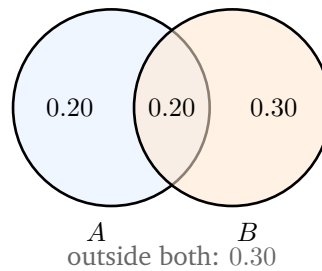
Q1. A data set has 10 candidate features. A model is to use *exactly* 3 of them. The number of distinct feature subsets of size 3 that can be formed is: [1 mark]

- (A) 720
- (B) 30
- (C) 120

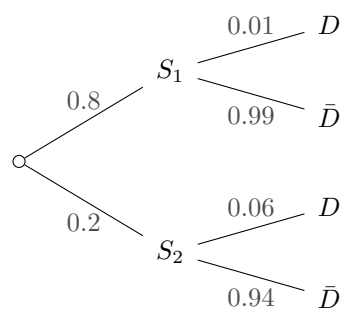


(D) 1000

- Q2.** For two events A and B the probabilities are shown on the Venn diagram below (the region outside both circles has probability 0.30). The conditional probability $P(B | A)$ is: [1 mark]



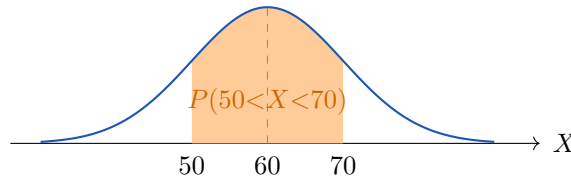
- (A) 0.50
 (B) 0.40
 (C) 0.20
 (D) 0.67
- Q3.** A store stocks the same sensor from two suppliers. Supplier S_1 provides 80% of the units with a defect rate of 1%; supplier S_2 provides the remaining 20% with a defect rate of 6%. The probability tree is shown. A unit drawn at random is found to be defective. The probability that it came from S_2 is: [1 mark]



- (A) 0.40
 (B) 0.60
 (C) 0.012
 (D) 0.008



- Q4.** A feature X is normally distributed with mean $\mu = 60$ and standard deviation $\sigma = 10$. Using the shaded central band of the standard normal curve shown, the probability $P(50 < X < 70)$ is nearest to (take $P(Z \leq 1) = 0.8413$): [1 mark]



- (A) 0.3413
(B) 0.1587
(C) 0.8413
(D) 0.6826
- Q5.** Typographical errors in a printed page follow a Poisson process at an average rate of 3 errors per page. The probability of finding *exactly* 1 error on a randomly chosen page is nearest to (take $e^{-3} = 0.0498$): [1 mark]
- (A) 0.224
(B) 0.149
(C) 0.050
(D) 0.100
- Q6.** The number of defective chips X in a small batch follows the distribution $P(X=0) = 0.1$, $P(X=1) = 0.2$, $P(X=2) = 0.4$, $P(X=3) = 0.3$. The expected number of defective chips $E[X]$ is: [1 mark]

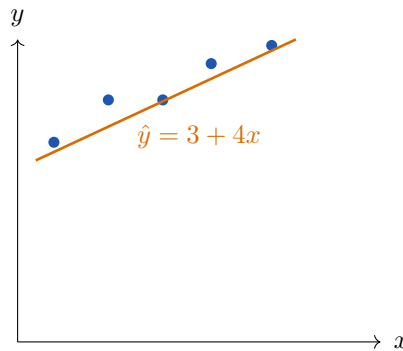
- (A) 1.5
(B) 2.0
(C) 1.0
(D) 1.9



Q7. A sample of size $n = 25$ is drawn from a population whose standard deviation is known to be $\sigma = 20$. The standard error of the sample mean, $\sigma/\sqrt{n}$, is: [1 mark]

- (A) 0.8
- (B) 20
- (C) 4
- (D) 5

Q8. The least-squares regression line $\hat{y} = 3 + 4x$ is fitted to the scatter shown. The value of $\hat{y}$ predicted at $x = 5$ is: [1 mark]



- (A) 7
- (B) 20
- (C) 12
- (D) 23

Q9. A biased coin has probability 0.4 of landing heads. It is tossed 3 times independently. The probability of obtaining *exactly* 2 heads is: [1 mark]

- (A) 0.500
- (B) 0.096
- (C) 0.144
- (D) 0.288

Q10. The time between successive requests at a server follows an exponential distribution with a mean of 2 seconds. The probability that the next request arrives *within* 2 seconds is (take $e^{-1} = 0.3679$): [1 mark]



- (A) 0.368
- (B) 0.135
- (C) 0.632
- (D) 0.865

Linear Algebra, Calculus & Optimization

Q11. Consider the matrix $A = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 3 & 0 \\ 0 & 1 & 4 \end{bmatrix}$. The rank of A is: [1 mark]

- (A) 1
- (B) 2
- (C) 3
- (D) 0

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

- (A) 5
- (B) 6
- (C) 8
- (D) 12

Q13. The larger eigenvalue of the matrix $C = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ is: [1 mark]

- (A) 3
- (B) -1
- (C) 1
- (D) 2

Q14. For a real matrix A , the matrix $A^T A$ has eigenvalues 36 and 16. The *smallest* singular value of A is: [1 mark]

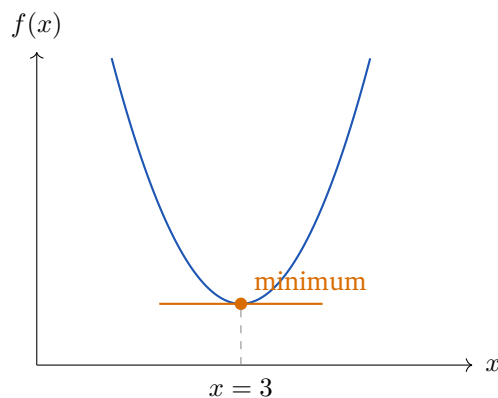


- (A) 4
- (B) 6
- (C) 16
- (D) 36

Q15. For the vector $v = (-7, 4, 2)$, the ℓ_∞ norm $\|v\|_\infty$ is: [1 mark]

- (A) 13
- (B) 4
- (C) 9
- (D) 7

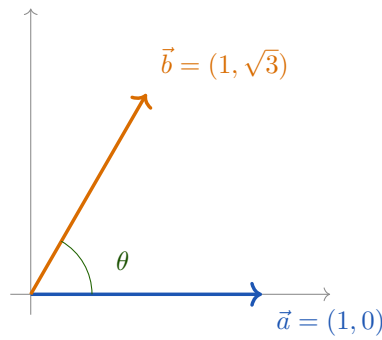
Q16. The function $f(x) = x^2 - 6x + 13$ is minimized over the real line, as suggested by its convex graph below. The minimum value of f is: [1 mark]



- (A) 3
- (B) 13
- (C) 4
- (D) 0

Q17. The angle θ between the vectors $\vec{a} = (1, 0)$ and $\vec{b} = (1, \sqrt{3})$, shown below, is: [1 mark]





- (A) 30°
- (B) 45°
- (C) 60°
- (D) 90°

Q18. Gradient descent is used to minimize $f(x) = x^2$ with a fixed learning rate $\eta = 0.2$. Starting from $x_0 = 5$, the value of x_1 after one update $x_1 = x_0 - \eta f'(x_0)$ is: [1 mark]

- (A) 3
- (B) 2
- (C) 4
- (D) 1

Q19. Consider the function $f(x, y) = xy$ on $\mathbb{R}^2$. Based on its Hessian, the function is: [1 mark]

- (A) convex everywhere
- (B) concave everywhere
- (C) linear (affine)
- (D) neither convex nor concave

Programming, Data Structures & Algorithms

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

```
s = "data"
print(s[::-1])
```

[1 mark]



- (A) atad
- (B) data
- (C) dat
- (D) ata

Q21. Consider the following Python code:

```
x = [1, 2, 3]
y = x[:]
y.append(4)
print(len(x))
```

The value printed is:

[2 marks]

- (A) 4
- (B) 3
- (C) 2
- (D) 1

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

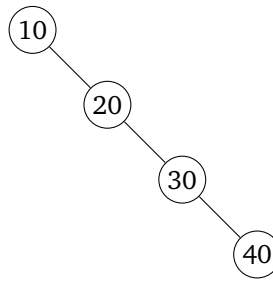
[2 marks]

- (A) 20
- (B) 10
- (C) 15
- (D) 5

Q23. Starting from an empty binary search tree, the keys 10, 20, 30, 40 are inserted in this (ascending) order, producing the right-skewed tree shown. The *height* of the tree, measured as the number of edges on the longest root-to-leaf path, is:

[2 marks]





- (A) 3
- (B) 4
- (C) 1
- (D) 0

Q24. A hash table of size 7 uses the hash function $h(k) = k \bmod 7$ with *linear probing* for collision resolution. The keys 10, 17, 24 are inserted in this order into an otherwise empty table. The final index occupied by key 24 is: [2 marks]

- (A) 3
- (B) 4
- (C) 5
- (D) 24

Q25. Binary search is performed on a sorted array of 63 distinct elements. The maximum number of key comparisons required in the worst case is: [2 marks]

- (A) 63
- (B) 32
- (C) 7
- (D) 6

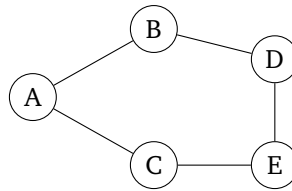
Q26. The worst-case time complexity of the standard Heap Sort algorithm on an array of n elements is: [2 marks]

- (A) $O(n^2)$



- (B) $O(n \log n)$
- (C) $O(n)$
- (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, D, E, C
- (B) A, B, C, D, E
- (C) A, C, E, D, B
- (D) A, B, C, E, D

Q28. The running time of an algorithm satisfies the recurrence $T(n) = 3T(n/2) + n$, with $T(1) = 1$. By the Master theorem, $T(n)$ is: [2 marks]

- (A) $\Theta(n)$
- (B) $\Theta(n^2)$
- (C) $\Theta(n \log n)$
- (D) $\Theta(n^{\log_2 3})$

Database Management & Warehousing

Q29. Relation R has 4 attributes and 8 tuples; relation S has 3 attributes and 6 tuples. The number of *attributes* (the degree) of the Cartesian product $R \times S$ is: [2 marks]

- (A) 12
- (B) 7
- (C) 1



(D) 48

Q30. A table EMP has 6 rows whose salary values are {3000, 6000, 5000, 8000, 4500, 7000}. What does the query `SELECT COUNT(*) FROM EMP WHERE salary > 5000;` return? [2 marks]

(A) 6

(B) 1

(C) 2

(D) 3

Q31. A relation $R(A, B, C)$ has *only* the functional dependencies $A \rightarrow B$ and $A \rightarrow C$, with A as the sole candidate key. The highest normal form that R satisfies is: [2 marks]

(A) 1NF

(B) 2NF

(C) BCNF

(D) 3NF only (not BCNF)

Q32. Relation $R(A, B)$ contains the tuples $\{(10, a), (20, a), (30, b)\}$ and relation $S(B, C)$ contains $\{(a, 1), (a, 2), (b, 3)\}$. The number of tuples in the natural join $R \bowtie S$ (on the common attribute B) is: [2 marks]

(A) 3

(B) 5

(C) 6

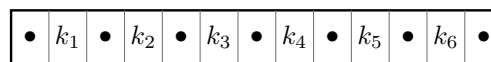
(D) 9

Q33. In OLAP, the operation that selects a single value for one dimension of a data cube (for example, fixing Time = 2024) to produce a lower-dimensional sub-cube is called: [2 marks]



- (A) roll-up
- (B) drill-down
- (C) slice
- (D) pivot

Q34. The figure shows the structure of an internal node of a B^+ tree of order $m = 7$ (each internal node holds up to m child pointers, drawn as $\bullet$, separated by keys). The maximum number of keys an internal node of this tree can hold is: [2 marks]



up to 7 pointers ($\bullet$) and 6 keys

- (A) 7
- (B) 6
- (C) 3
- (D) 8

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

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

Machine Learning & Artificial Intelligence

Q36. A logistic regression model has weight vector $w = (1, -1)$ and bias $b = 0$, and uses the sigmoid $\sigma(z) = 1/(1 + e^{-z})$. For the input $x = (3, 3)$, the predicted probability $\sigma(w^T x + b)$ is: [2 marks]

- (A) 0.50
- (B) 0.73



(C) 0.27

(D) 0.88

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

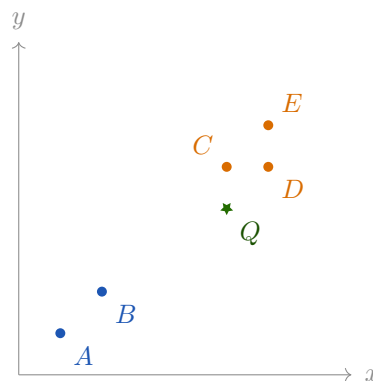
(A) high bias, low variance

(B) low bias, high variance

(C) low bias, low variance

(D) high bias, high variance

Q38. A k -nearest-neighbour classifier with $k = 3$ (Euclidean distance) is trained on the labelled points $A(1, 1)$: Class 0, $B(2, 2)$: Class 0, $C(5, 5)$: Class 1, $D(6, 5)$: Class 1, $E(6, 6)$: Class 1, shown below. The query point $Q(5, 4)$ is classified as: [2 marks]



(A) Class 0

(B) Class 1

(C) a tie (undecided)

(D) cannot be determined

Q39. A training set for a binary classification task contains 8 examples, of which 4 are positive and 4 are negative. The entropy of this set (in bits) is: [2 marks]

(A) 0



- (B) 0.5
- (C) 0.811
- (D) 1.0

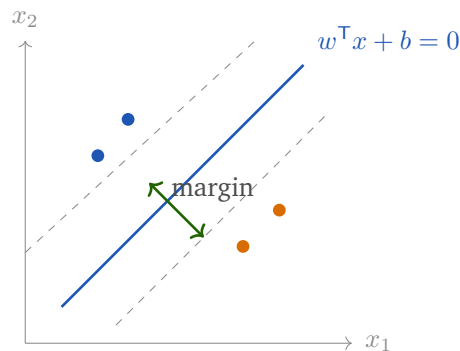
Q40. A node in a decision tree contains 3 positive and 1 negative example, as shown. The Gini impurity of this node, $1 - \sum_i p_i^2$, is: [2 marks]

3+, 1-

4 examples at the node

- (A) 0.5
- (B) 0.375
- (C) 0.25
- (D) 0.75

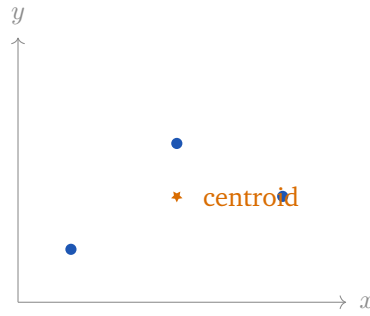
Q41. For a linear support vector machine the separating hyperplane $w^T x + b = 0$ is shown with its two margin boundaries. If $\|w\| = 0.4$, the width of the margin, equal to $2/\|w\|$, is: [2 marks]



- (A) 0.4
- (B) 5
- (C) 2
- (D) 2.5

Q42. In one iteration of the k -means algorithm, a cluster is assigned the three points $(1, 1)$, $(3, 3)$ and $(5, 2)$, shown below. The updated centroid of this cluster is: [2 marks]



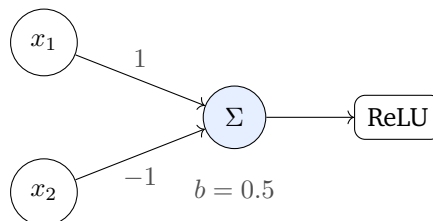


- (A) (3, 2)
- (B) (3, 3)
- (C) (2, 3)
- (D) (9, 6)

Q43. Principal component analysis is applied to a data set whose covariance matrix has eigenvalues 5, 3, 1, 1. The fraction of the total variance explained by the first two principal components is: [2 marks]

- (A) 50%
- (B) 30%
- (C) 80%
- (D) 70%

Q44. A single neuron with a ReLU activation computes $\text{ReLU}(w_1x_1 + w_2x_2 + b)$, as shown, where $\text{ReLU}(z) = \max(0, z)$. With inputs $x_1 = 2$, $x_2 = 1$, weights $w_1 = 1$, $w_2 = -1$ and bias $b = 0.5$, the output of the neuron is: [2 marks]

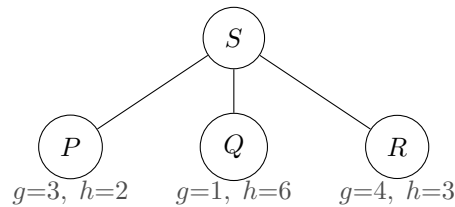


- (A) 1.5
- (B) 0



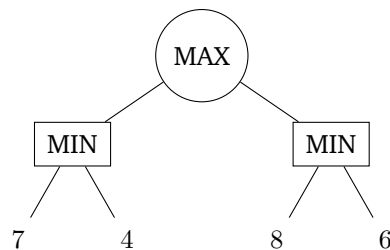
(C) -1.5 (D) 0.5

- Q45.** A* search maintains three frontier nodes P , Q and R with g (path cost so far) and h (heuristic) values shown in the tree below, where $f = g + h$. The node A* selects for expansion next is: [2 marks]

(A) P (B) Q (C) R

(D) all three at once

- 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) 4

(B) 7

(C) 8

(D) 6



Detailed Solutions

Q1.

Solution

Concept – unordered selection (combination): Choosing a subset where order does not matter uses the binomial coefficient $\binom{n}{r}$.

Step 1 – identify n and r : $n = 10$ features, $r = 3$ chosen.

Step 2 – write the combination: $\binom{10}{3} = \frac{10 \times 9 \times 8}{3 \times 2 \times 1}$

Step 3 – evaluate the numerator: $10 \times 9 \times 8 = 720$

Step 4 – evaluate the denominator: $3 \times 2 \times 1 = 6$

Step 5 – divide: $\frac{720}{6} = 120$

Why (A) is wrong: 720 is the number of *ordered* arrangements $\binom{10}{3} \times 3!$ split differently, i.e. $P(10, 3)$; subsets are unordered.

Final Answer: 120 subsets $\Rightarrow$

Answer: (C) [Go Back to Q1](#)

Q2.

Solution

Concept – conditional probability: $P(B | A) = \frac{P(A \cap B)}{P(A)}$.

Step 1 – read the overlap from the Venn diagram: $P(A \cap B) = 0.20$ (the shared region).

Step 2 – read $P(A)$ as the whole of circle A: $P(A) = 0.20 + 0.20 = 0.40$

Step 3 – substitute: $P(B | A) = \frac{0.20}{0.40}$

Step 4 – divide: $P(B | A) = 0.50$

Step 5 – sanity check the whole sample space: $0.20 + 0.20 + 0.30 + 0.30 = 1.00$, so the diagram is consistent.

Final Answer: $P(B | A) = 0.50 \Rightarrow$

Answer: (A) [Go Back to Q2](#)



Q3.

Solution

Concept – Bayes' theorem (reverse the tree): $P(S_2 | D) = \frac{P(S_2) P(D | S_2)}{P(S_1) P(D | S_1) + P(S_2) P(D | S_2)}$.

Step 1 – probability from the S_1 branch: $P(S_1) P(D | S_1) = 0.80 \times 0.01 = 0.008$

Step 2 – probability from the S_2 branch: $P(S_2) P(D | S_2) = 0.20 \times 0.06 = 0.012$

Step 3 – total probability of a defective unit: $P(D) = 0.008 + 0.012 = 0.020$

Step 4 – apply Bayes' theorem: $P(S_2 | D) = \frac{0.012}{0.020}$

Step 5 – divide: $P(S_2 | D) = 0.60$

Final Answer: $P(S_2 | D) = 0.60 \Rightarrow \boxed{\text{B}}$

Answer: (B) [Go Back to Q3](#)

Q4.

Solution

Concept – standardize, then use the normal table: Convert X to the standard normal $Z = \frac{X - \mu}{\sigma}$.

Step 1 – standardize the lower cut-off $X = 50$: $Z_1 = \frac{50 - 60}{10} = \frac{-10}{10} = -1$

Step 2 – standardize the upper cut-off $X = 70$: $Z_2 = \frac{70 - 60}{10} = \frac{10}{10} = 1$

Step 3 – write the required probability in Z : $P(50 < X < 70) = P(-1 < Z < 1)$

Step 4 – express through the CDF: $P(-1 < Z < 1) = P(Z \leq 1) - P(Z \leq -1)$

Step 5 – use symmetry $P(Z \leq -1) = 1 - P(Z \leq 1)$: $P(-1 < Z < 1) = 0.8413 - (1 - 0.8413)$

$$= 0.8413 - 0.1587$$

$$= 0.6826$$

Final Answer: $P(50 < X < 70) = 0.6826 \Rightarrow \boxed{\text{D}}$

Answer: (D) [Go Back to Q4](#)



Q5.

Solution

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

Step 1 – list the parameters: $\lambda = 3$ per page, $k = 1$.

Step 2 – substitute into the formula: $P(X = 1) = \frac{e^{-3} 3^1}{1!}$

Step 3 – simplify the power and factorial: $3^1 = 3$ and $1! = 1$

$$P(X = 1) = e^{-3} \times 3$$

Step 4 – substitute $e^{-3} = 0.0498$: $P(X = 1) = 0.0498 \times 3$

Step 5 – multiply: $P(X = 1) = 0.1494 \approx 0.149$

Final Answer: $P(X = 1) \approx 0.149 \Rightarrow \boxed{\text{B}}$

Answer: (B) [Go Back to Q5](#)

Q6.

Solution

Concept – expected value of a discrete variable: $E[X] = \sum_x x P(X = x)$.

Step 1 – form each term $x P(X = x)$: $0 \times 0.1 = 0$

$$1 \times 0.2 = 0.2$$

$$2 \times 0.4 = 0.8$$

$$3 \times 0.3 = 0.9$$

Step 2 – add the terms: $E[X] = 0 + 0.2 + 0.8 + 0.9$

Step 3 – finish the addition: $E[X] = 1.9$

Step 4 – sanity check the probabilities: $0.1 + 0.2 + 0.4 + 0.3 = 1.0$, so the distribution is valid.

Final Answer: $E[X] = 1.9 \Rightarrow \boxed{\text{D}}$

Answer: (D) [Go Back to Q6](#)

Q7.

Solution

Concept – standard error of the mean: For a sample mean with known population σ , the standard error is $SE = \frac{\sigma}{\sqrt{n}}$.

Step 1 – take the square root of the sample size: $\sqrt{n} = \sqrt{25} = 5$



Step 2 – substitute into the formula: $SE = \frac{20}{5}$

Step 3 – divide: $SE = 4$

Why (B) is wrong: 20 is the population standard deviation σ itself, not the standard error of the *mean*, which shrinks by the factor $\sqrt{n}$.

Final Answer: $SE = 4 \Rightarrow$ C

Answer: (C) [Go Back to Q7](#)

Q8.

Solution

Concept – prediction from a fitted regression line: Substitute the given x into $\hat{y} = 3 + 4x$.

Step 1 – write the line and the query point: $\hat{y} = 3 + 4x$, with $x = 5$.

Step 2 – compute the slope term: $4 \times 5 = 20$

Step 3 – add the intercept: $\hat{y} = 3 + 20$

Step 4 – finish: $\hat{y} = 23$

Final Answer: $\hat{y} = 23 \Rightarrow$ D

Answer: (D) [Go Back to Q8](#)

Q9.

Solution

Concept – binomial probability: $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

Step 1 – list the parameters: $n = 3, k = 2, p = 0.4, 1 - p = 0.6$.

Step 2 – evaluate the binomial coefficient: $\binom{3}{2} = 3$

Step 3 – evaluate the powers: $p^2 = 0.4^2 = 0.16$ and $(1 - p)^1 = 0.6$

Step 4 – multiply the pieces: $P(X = 2) = 3 \times 0.16 \times 0.6$

Step 5 – finish the arithmetic: $3 \times 0.16 = 0.48$

$0.48 \times 0.6 = 0.288$

Final Answer: $P(X = 2) = 0.288 \Rightarrow$ D

Answer: (D) [Go Back to Q9](#)



Q10.

Solution

Concept – exponential CDF: For an exponential variable with rate λ , $P(X \leq t) = 1 - e^{-\lambda t}$.

Step 1 – find the rate from the mean: Mean = $\frac{1}{\lambda} = 2$ s, so $\lambda = \frac{1}{2} = 0.5 \text{ s}^{-1}$.

Step 2 – form the exponent: $\lambda t = 0.5 \times 2 = 1$

Step 3 – write the “within t ” probability: $P(X \leq 2) = 1 - e^{-1}$

Step 4 – substitute $e^{-1} = 0.3679$: $P(X \leq 2) = 1 - 0.3679$

Step 5 – subtract: $P(X \leq 2) = 0.6321 \approx 0.632$

Final Answer: $P(X \leq 2) \approx 0.632 \Rightarrow \boxed{\text{C}}$

Answer: (C) [Go Back to Q10](#)

Q11.

Solution

Concept – rank via the determinant: A 3×3 matrix has full rank 3 exactly when its determinant is non-zero.

Step 1 – expand the determinant along the first row: $\det A = 2 \begin{vmatrix} 3 & 0 \\ 1 & 4 \end{vmatrix} - 0 \begin{vmatrix} 1 & 0 \\ 0 & 4 \end{vmatrix} + 1 \begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix}$

Step 2 – evaluate the first minor: $\begin{vmatrix} 3 & 0 \\ 1 & 4 \end{vmatrix} = (3)(4) - (0)(1) = 12$

Step 3 – evaluate the third minor: $\begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} = (1)(1) - (3)(0) = 1$

Step 4 – combine: $\det A = 2(12) - 0 + 1(1) = 24 + 1 = 25$

Step 5 – conclude: $\det A = 25 \neq 0$, so all three rows are independent and $\text{rank}(A) = 3$.

Final Answer: $\text{rank}(A) = 3 \Rightarrow \boxed{\text{C}}$

Answer: (C) [Go Back to Q11](#)



Q12.

Solution

Concept – cofactor expansion along the first row: $\det B = b_{11}M_{11} - b_{12}M_{12} + b_{13}M_{13}$.

Step 1 – first-row entries: $b_{11} = 1, b_{12} = 2, b_{13} = 0$.

Step 2 – minor for b_{11} : $\begin{vmatrix} 3 & 0 \\ 0 & 2 \end{vmatrix} = (3)(2) - (0)(0) = 6$

Step 3 – minor for b_{12} : $\begin{vmatrix} 0 & 0 \\ 1 & 2 \end{vmatrix} = (0)(2) - (0)(1) = 0$

Step 4 – the third term vanishes: Since $b_{13} = 0$, its contribution is 0.

Step 5 – combine with the sign pattern: $\det B = 1(6) - 2(0) + 0$
 $= 6 - 0$
 $= 6$

Final Answer: $\det B = 6 \Rightarrow \boxed{\text{B}}$

Answer: (B) [Go Back to Q12](#)

Q13.

Solution

Concept – eigenvalues from the characteristic equation: Solve $\det(C - \lambda I) = 0$; use trace and determinant as a shortcut.

Step 1 – trace and determinant of C : $\text{tr}(C) = 1 + 1 = 2$ and $\det(C) = (1)(1) - (2)(2) = 1 - 4 = -3$.

Step 2 – write the characteristic equation: $\lambda^2 - \text{tr}(C)\lambda + \det(C) = 0$
 $\lambda^2 - 2\lambda - 3 = 0$

Step 3 – factor: $(\lambda - 3)(\lambda + 1) = 0$

Step 4 – read the roots: $\lambda = 3$ or $\lambda = -1$.

Step 5 – pick the larger: The larger eigenvalue is 3.

Final Answer: $\lambda_{\max} = 3 \Rightarrow \boxed{\text{A}}$

Answer: (A) [Go Back to Q13](#)



Q14.

Solution

Concept – singular values are square roots of the eigenvalues of $A^T A$: $\sigma_i = \sqrt{\lambda_i(A^T A)}$.

Step 1 – take the square roots of the given eigenvalues: $\sqrt{36} = 6$ and $\sqrt{16} = 4$.

Step 2 – list the singular values: $\sigma_1 = 6$ and $\sigma_2 = 4$.

Step 3 – identify the smallest: $\sigma_{\min} = 4$.

Why (B) is wrong: 6 is the *largest* singular value (the spectral norm), not the smallest.

Final Answer: $\sigma_{\min} = 4 \Rightarrow \boxed{A}$

Answer: (A) [Go Back to Q14](#)

Q15.

Solution

Concept – the ℓ_∞ norm is the largest absolute component: $\|v\|_\infty = \max_i |v_i|$.

Step 1 – take absolute values of the components: $|-7| = 7$, $|4| = 4$, $|2| = 2$.

Step 2 – take the maximum: $\max(7, 4, 2) = 7$

Step 3 – contrast with the other norms: $\|v\|_1 = 7 + 4 + 2 = 13$ and $\|v\|_2 = \sqrt{49 + 16 + 4} = \sqrt{69}$; the ℓ_∞ value is 7.

Final Answer: $\|v\|_\infty = 7 \Rightarrow \boxed{D}$

Answer: (D) [Go Back to Q15](#)

Q16.

Solution

Concept – minimize by setting the derivative to zero: For a convex parabola, the stationary point is the global minimum.

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 minimum: $f''(x) = 2 > 0$, so the point is a minimum.

Step 4 – evaluate f at $x = 3$: $f(3) = 3^2 - 6(3) + 13$

$$= 9 - 18 + 13$$



$= 4$

Final Answer: minimum value $= 4 \Rightarrow \boxed{\text{C}}$

Answer: (C) [Go Back to Q16](#)

Q17.

Solution

Concept – angle between vectors: $\cos \theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|}$.

Step 1 – compute the dot product: $\vec{a} \cdot \vec{b} = (1)(1) + (0)(\sqrt{3}) = 1$

Step 2 – compute the magnitudes: $\|\vec{a}\| = \sqrt{1^2 + 0^2} = 1$

$\|\vec{b}\| = \sqrt{1^2 + (\sqrt{3})^2} = \sqrt{1 + 3} = 2$

Step 3 – form the cosine: $\cos \theta = \frac{1}{1 \times 2} = 0.5$

Step 4 – invert to find the angle: $\theta = \cos^{-1}(0.5) = 60^\circ$

Final Answer: $\theta = 60^\circ \Rightarrow \boxed{\text{C}}$

Answer: (C) [Go Back to Q17](#)

Q18.

Solution

Concept – one gradient-descent step: $x_1 = x_0 - \eta f'(x_0)$.

Step 1 – differentiate the objective: $f(x) = x^2 \Rightarrow f'(x) = 2x$

Step 2 – evaluate the gradient at $x_0 = 5$: $f'(5) = 2 \times 5 = 10$

Step 3 – apply the update with $\eta = 0.2$: $x_1 = 5 - 0.2 \times 10$

Step 4 – finish the arithmetic: $x_1 = 5 - 2 = 3$

Final Answer: $x_1 = 3 \Rightarrow \boxed{\text{A}}$

Answer: (A) [Go Back to Q18](#)



Q19.

Solution

Concept – convexity via the Hessian: A twice-differentiable function is convex iff its Hessian is positive semi-definite everywhere, and concave iff negative semi-definite.

Step 1 – compute the second partials: $f_{xx} = 0, f_{yy} = 0, f_{xy} = f_{yx} = 1$.

Step 2 – write the Hessian: $H = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Step 3 – find the eigenvalues of H : $\det(H - \lambda I) = \lambda^2 - 1 = 0 \Rightarrow \lambda = +1, -1$.

Step 4 – classify from the signs: One eigenvalue is positive and one is negative, so H is indefinite (a saddle shape).

Step 5 – conclude: An indefinite Hessian means f is neither convex nor concave.

Final Answer: neither convex nor concave $\Rightarrow$ D

Answer: (D) [Go Back to Q19](#)

Q20.

Solution

Concept – extended slice with a negative step: $s[::-1]$ walks the string from the last character to the first, giving the reverse.

Step 1 – index the string: $s = \left[\underbrace{d}_0, \underbrace{a}_1, \underbrace{t}_2, \underbrace{a}_3 \right]$

Step 2 – read the slice $[::-1]$ from the end: Take index 3 $\rightarrow$ a, index 2 $\rightarrow$ t, index 1 $\rightarrow$ a, index 0 $\rightarrow$ d.

Step 3 – concatenate: The reversed string is atad.

Final Answer: atad $\Rightarrow$ A

Answer: (A) [Go Back to Q20](#)

Q21.

Solution

Concept – slicing makes a shallow copy: $y = x[:]$ builds a *new* list with the same elements, so y is not an alias of x .

Step 1 – after $x = [1, 2, 3]$: x refers to a list of length 3.

Step 2 – after $y = x[:]$: y points to a separate copy; changes to y do not affect x .



Step 3 – after `y.append(4)`: Only `y` grows to `[1, 2, 3, 4]`; `x` is unchanged.

Step 4 – evaluate `len(x)`: `x` still has 3 elements, so `len(x)` is 3.

Why (A) is wrong: 4 would result from `y = x` (an alias); the slice `[:]` breaks the sharing.

Final Answer: 3 $\Rightarrow$

Answer: (B) [Go Back to Q21](#)

Q22.

Solution

Concept – a queue is first-in, first-out (FIFO): Elements leave in the same order they entered; track the front.

Step 1 – `enqueue(5)`, `enqueue(10)`: Queue (front $\rightarrow$ rear): `[5, 10]`.

Step 2 – `dequeue()`: Removes the front 5; queue becomes `[10]`.

Step 3 – `enqueue(15)`, `enqueue(20)`: Queue becomes `[10, 15, 20]`.

Step 4 – `dequeue()`: Removes the front 10; queue becomes `[15, 20]`.

Step 5 – `dequeue()` (the last one): Removes the front 15; this is the value returned.

Final Answer: last `dequeue()` returns 15 $\Rightarrow$

Answer: (C) [Go Back to Q22](#)

Q23.

Solution

Concept – BST insertion in ascending order gives a skewed chain: Each new key is larger than all existing keys, so it always attaches as a right child.

Step 1 – insert 10: 10 becomes the root (level 0).

Step 2 – insert 20: $20 > 10$, so it is the right child of 10 (level 1).

Step 3 – insert 30 and 40: $30 > 20$ (right child of 20, level 2); $40 > 30$ (right child of 30, level 3).

Step 4 – trace the longest path: $10 \rightarrow 20 \rightarrow 30 \rightarrow 40$ uses 3 edges.

Step 5 – read the height: Height as the number of edges on the longest root-to-leaf path = 3.

Final Answer: height = 3 $\Rightarrow$



Answer: (A) [Go Back to Q23](#)

Q24.

Solution

Concept – linear probing scans the next free slot: On a collision at index i , try $i + 1, i + 2, \dots$ (modulo table size).

Step 1 – insert 10: $h(10) = 10 \bmod 7 = 3$; slot 3 is free, so 10 goes to index 3.

Step 2 – insert 17: $h(17) = 17 \bmod 7 = 3$; index 3 is occupied, so probe index 4; it is free, so 17 goes to index 4.

Step 3 – insert 24: $h(24) = 24 \bmod 7 = 3$; index 3 occupied $\rightarrow$ probe index 4 (occupied) $\rightarrow$ probe index 5 (free).

Step 4 – place 24: 24 is stored at index 5.

Final Answer: index 5 $\Rightarrow$ C

Answer: (C) [Go Back to Q24](#)

Q25.

Solution

Concept – binary search halves the range each step: The worst-case number of comparisons is $\lfloor \log_2 n \rfloor + 1$.

Step 1 – take the base-2 logarithm of n : $\log_2 63 \approx 5.98$

Step 2 – take the floor: $\lfloor 5.98 \rfloor = 5$

Step 3 – add one: $5 + 1 = 6$

Step 4 – cross-check by doubling: $2^5 = 32 \leq 63 < 64 = 2^6$, so 6 halvings suffice in the worst case.

Final Answer: 6 comparisons $\Rightarrow$ D

Answer: (D) [Go Back to Q25](#)

Q26.

Solution

Concept – Heap Sort builds a heap and repeatedly extracts the maximum: Building the heap is $O(n)$; each of the n extractions costs a sift-down of $O(\log n)$.

Step 1 – cost of the n extractions: $n \times O(\log n) = O(n \log n)$



Step 2 – add the heap-construction cost: $O(n) + O(n \log n) = O(n \log n)$

Step 3 – note the guarantee: This bound holds in the worst case (and every case), unlike Quicksort whose worst case is $O(n^2)$.

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

Answer: (B) [Go Back to Q26](#)

Q27.

Solution

Concept – DFS goes as deep as possible before backtracking: Visit a vertex, then recurse into its smallest-labelled unvisited neighbour.

Step 1 – list the adjacencies: $A : \{B, C\}, B : \{A, D\}, C : \{A, E\}, D : \{B, E\}, E : \{C, D\}$.

Step 2 – start at A, go to B: A's smallest neighbour is B; visit B.

Step 3 – from B, go to D: B's unvisited neighbour is D; visit D.

Step 4 – from D, go to E: D's neighbours are B (visited) and E; visit E.

Step 5 – from E, go to C: E's smallest unvisited neighbour is C; visit C. C has no unvisited neighbours, so DFS stops.

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

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

Answer: (A) [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 – identify the parameters: $a = 3, b = 2, f(n) = n$.

Step 2 – compute the critical exponent: $\log_2 3 \approx 1.585$, so $n^{\log_2 3} = n^{1.585}$.

Step 3 – compare $f(n)$ with $n^{\log_2 3}$: $f(n) = n = n^1$ grows slower than $n^{1.585}$, i.e. $f(n) = O(n^{\log_2 3 - \epsilon})$ – Master-theorem Case 1.

Step 4 – apply Case 1: $T(n) = \Theta(n^{\log_2 3}) = \Theta(n^{1.585})$

Final Answer: $T(n) = \Theta(n^{\log_2 3}) \Rightarrow \boxed{\text{D}}$

Answer: (D) [Go Back to Q28](#)



Q29.

Solution

Concept – degree of a Cartesian product: $R \times S$ has a number of attributes equal to the *sum* of the two degrees (its number of tuples is the product, but that is not asked).

Step 1 – read the degrees: $\text{degree}(R) = 4$, $\text{degree}(S) = 3$.

Step 2 – add them: $\text{degree}(R \times S) = 4 + 3 = 7$

Why (D) is wrong: $48 = 8 \times 6$ is the number of *tuples*, not the number of attributes.

Final Answer: 7 attributes $\Rightarrow$ **B**

Answer: (B) [Go Back to Q29](#)

Q30.

Solution

Concept – COUNT(*) with a WHERE filter: Only the rows that satisfy the predicate are counted.

Step 1 – list the salaries: {3000, 6000, 5000, 8000, 4500, 7000}.

Step 2 – apply the predicate salary > 5000 (strictly greater): 3000 (no), 6000 (yes), 5000 (no – not strictly greater), 8000 (yes), 4500 (no), 7000 (yes).

Step 3 – collect the qualifying rows: {6000, 8000, 7000}.

Step 4 – count them: 3 rows qualify.

Final Answer: the query returns 3 $\Rightarrow$ **D**

Answer: (D) [Go Back to Q30](#)

Q31.

Solution

Concept – BCNF requires every determinant to be a superkey: When the sole candidate key determines everything and nothing else determines any attribute, the relation is in BCNF.

Step 1 – identify the candidate key: A is the only candidate key, so A is a superkey.

Step 2 – examine each FD's left side: $A \rightarrow B$ and $A \rightarrow C$ both have A (a superkey) on the left.

Step 3 – apply the BCNF test: For every non-trivial FD $X \rightarrow Y$, X must be a



superkey; here every determinant is A , which is a superkey.

Step 4 – conclude: All FDs satisfy the BCNF condition, so the highest normal form is BCNF.

Final Answer: highest form = BCNF $\Rightarrow$

Answer: (C) [Go Back to Q31](#)

Q32.

Solution

Concept – natural join matches tuples on the common attribute: Combine a tuple of R with a tuple of S whenever their B values are equal.

Step 1 – match R 's tuple $(10, a)$: S has $(a, 1)$ and $(a, 2)$ with $B = a$, giving $(10, a, 1)$ and $(10, a, 2)$ – 2 tuples.

Step 2 – match R 's tuple $(20, a)$: S again has $(a, 1)$ and $(a, 2)$, giving $(20, a, 1)$ and $(20, a, 2)$ – 2 tuples.

Step 3 – match R 's tuple $(30, b)$: S has $(b, 3)$ with $B = b$, giving $(30, b, 3)$ – 1 tuple.

Step 4 – total the matches: $2 + 2 + 1 = 5$

Final Answer: 5 tuples $\Rightarrow$

Answer: (B) [Go Back to Q32](#)

Q33.

Solution

Concept – OLAP selection operations: Slice fixes one dimension at a single value; dice selects ranges on several dimensions.

Step 1 – describe the operation: Fixing Time = 2024 keeps only the cells for that one value of the Time dimension.

Step 2 – match the definition: Cutting the cube at a single value of one dimension is a **slice**.

Why the other options are wrong: Roll-up aggregates to a coarser level; drill-down goes to finer detail; pivot rotates the cube's axes for presentation.

Final Answer: slice $\Rightarrow$

Answer: (C) [Go Back to Q33](#)



Q34.

Solution

Concept – order and keys in a B^+ tree: An internal node of order m holds at most m child pointers, and the number of keys is always one fewer than the number of pointers.

Step 1 – maximum pointers: Order $m = 7 \Rightarrow$ up to 7 child pointers.

Step 2 – keys are one fewer than pointers: Maximum keys $= m - 1 = 7 - 1$

Step 3 – evaluate: $= 6$

Final Answer: 6 keys $\Rightarrow$ **B**

Answer: (B) [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 $A \rightarrow B$: A determines B , so $A^+ = \{A, B\}$.

Step 3 – apply $B \rightarrow C$: B is now in the closure, so it brings in C : $A^+ = \{A, B, C\}$.

Step 4 – check for more: No FD has C , D , or E alone on its left that could add D or E ; the closure stops at $\{A, B, C\}$.

Why (C) is wrong: D and E are never determined by A (no FD produces them), so the closure is not all five attributes.

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

Answer: (B) [Go Back to Q35](#)

Q36.

Solution

Concept – logistic regression prediction: Compute the linear score $z = w^T x + b$, then apply the sigmoid.

Step 1 – compute the linear score: $z = (1)(3) + (-1)(3) + 0$

$= 3 - 3 + 0$

$= 0$



Step 2 – write the sigmoid at $z = 0$: $\sigma(0) = \frac{1}{1 + e^0}$

Step 3 – use $e^0 = 1$: $\sigma(0) = \frac{1}{1 + 1} = \frac{1}{2}$

Step 4 – read the value: $\sigma(0) = 0.50$

Final Answer: $\sigma(0) = 0.50 \Rightarrow \boxed{A}$

Answer: (A) [Go Back to Q36](#)

Q37.

Solution

Concept – underfitting in bias–variance terms: A model too simple to capture the pattern does poorly everywhere: high bias, low variance.

Step 1 – interpret the poor training accuracy: Failing even on the training data means the model cannot represent the signal – *high bias*.

Step 2 – interpret the similar poor test accuracy: Its predictions barely change with the training sample, so it is stable – *low variance*.

Step 3 – combine: Poor on both sets with stable behaviour is the classic signature of underfitting: high bias, low variance.

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

Answer: (A) [Go Back to Q37](#)

Q38.

Solution

Concept – kNN takes a majority vote over the k nearest points: Compute distances from Q to every training point, keep the 3 closest, and vote.

Step 1 – distance $Q(5, 4)$ to $A(1, 1)$: $\sqrt{(5-1)^2 + (4-1)^2} = \sqrt{16+9} = \sqrt{25} = 5.00$

Step 2 – distance to $B(2, 2)$: $\sqrt{(5-2)^2 + (4-2)^2} = \sqrt{9+4} = \sqrt{13} = 3.61$

Step 3 – distance to $C(5, 5)$: $\sqrt{(5-5)^2 + (4-5)^2} = \sqrt{0+1} = 1.00$

Step 4 – distances to D and E : $D(6, 5) : \sqrt{1+1} = 1.41$; $E(6, 6) : \sqrt{1+4} = \sqrt{5} = 2.24$.

Step 5 – pick the 3 nearest and vote: Nearest three are C (1.00, Class 1), D (1.41, Class 1), E (2.24, Class 1). Votes: Class 1 three times.

Step 6 – assign the majority label: Class 1 wins 3 to 0.



Final Answer: Q is Class 1 $\Rightarrow$ B

Answer: (B) [Go Back to Q38](#)

Q39.

Solution

Concept – binary entropy: $H = -p \log_2 p - (1 - p) \log_2 (1 - p)$, with p the fraction of positives.

Step 1 – compute the class proportions: $p = \frac{4}{8} = 0.5$ positive, $1 - p = \frac{4}{8} = 0.5$ negative.

Step 2 – take the base-2 logs: $\log_2 0.5 = -1$

Step 3 – form each term: $-0.5 \times (-1) = 0.5$

$$-0.5 \times (-1) = 0.5$$

Step 4 – add the terms: $H = 0.5 + 0.5 = 1.0$ bit

Step 5 – interpret: A perfectly balanced two-class set has maximum entropy, exactly 1 bit.

Final Answer: $H = 1.0$ bit $\Rightarrow$ D

Answer: (D) [Go Back to Q39](#)

Q40.

Solution

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

Step 1 – compute the class proportions: $p_+ = \frac{3}{4} = 0.75$, $p_- = \frac{1}{4} = 0.25$.

Step 2 – square each proportion: $p_+^2 = 0.75^2 = 0.5625$

$$p_-^2 = 0.25^2 = 0.0625$$

Step 3 – sum the squares: $0.5625 + 0.0625 = 0.625$

Step 4 – subtract from 1: $G = 1 - 0.625 = 0.375$

Final Answer: $G = 0.375 \Rightarrow$ B

Answer: (B) [Go Back to Q40](#)



Q41.

Solution

Concept – SVM margin width: For the hyperplane $w^T x + b = 0$, the full margin (distance between the two boundaries) is $\frac{2}{\|w\|}$.

Step 1 – write the margin formula: $\text{margin} = \frac{2}{\|w\|}$

Step 2 – substitute $\|w\| = 0.4$: $\text{margin} = \frac{2}{0.4}$

Step 3 – divide: $\text{margin} = 5$

Final Answer: margin width = 5 $\Rightarrow$ **B**

Answer: (B) [Go Back to Q41](#)

Q42.

Solution

Concept – k-means centroid update: The new centroid is the coordinate-wise mean of the points assigned to the cluster.

Step 1 – average the x -coordinates: $\bar{x} = \frac{1 + 3 + 5}{3} = \frac{9}{3} = 3$

Step 2 – average the y -coordinates: $\bar{y} = \frac{1 + 3 + 2}{3} = \frac{6}{3} = 2$

Step 3 – assemble the centroid: $(\bar{x}, \bar{y}) = (3, 2)$

Why (D) is wrong: (9, 6) is the coordinate-wise *sum*, not the mean; divide by the 3 points.

Final Answer: centroid = (3, 2) $\Rightarrow$ **A**

Answer: (A) [Go Back to Q42](#)

Q43.

Solution

Concept – variance explained in PCA: Each principal component explains variance proportional to its eigenvalue; the fraction is the eigenvalue sum of the kept components over the total.

Step 1 – total variance (sum of all eigenvalues): $5 + 3 + 1 + 1 = 10$

Step 2 – variance in the first two components: $5 + 3 = 8$

Step 3 – form the fraction: $\frac{8}{10} = 0.80$



Step 4 – express as a percentage: $0.80 = 80\%$

Final Answer: 80% of the variance $\Rightarrow$

Answer: (C) [Go Back to Q43](#)

Q44.

Solution

Concept – neuron forward pass with ReLU: Compute the weighted sum plus bias, then apply $\text{ReLU}(z) = \max(0, z)$.

Step 1 – weighted sum of inputs: $w_1x_1 + w_2x_2 = (1)(2) + (-1)(1)$

$$= 2 - 1$$

$$= 1$$

Step 2 – add the bias: $z = 1 + 0.5 = 1.5$

Step 3 – apply ReLU: $\text{ReLU}(1.5) = \max(0, 1.5) = 1.5$

Final Answer: output = 1.5 $\Rightarrow$

Answer: (A) [Go Back to Q44](#)

Q45.

Solution

Concept – A* expands the node of least $f = g + h$: Compute f for each frontier node and pick the minimum.

Step 1 – f for node P : $f(P) = g + h = 3 + 2 = 5$

Step 2 – f for node Q : $f(Q) = 1 + 6 = 7$

Step 3 – f for node R : $f(R) = 4 + 3 = 7$

Step 4 – choose the smallest f : $\min(5, 7, 7) = 5$, which belongs to node P .

Final Answer: A* expands $P \Rightarrow$

Answer: (A) [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: $\min(8, 6) = 6$

Step 3 – evaluate the MAX root: $\max(4, 6) = 6$

Final Answer: minimax value = 6 $\Rightarrow$ D

Answer: (D) [Go Back to Q46](#)



Answer Key

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

