GATE Data Science and Artificial Intelligence

Sample Paper – 6

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 box contains 4 red and 6 blue balls. Two balls are drawn one after another *without* replacement. The probability that *both* drawn balls are red is: **[1 mark]**

(A) $\dfrac{2}{15}$

(B) $\dfrac{4}{25}$

(C) $\dfrac{1}{3}$

(D) $\frac{8}{45}$

Q2. Two *independent* events A and B have probabilities $P(A) = 0.4$ and $P(B) = 0.5$. The probability of the union $P(A \cup B)$ is: **[1 mark]**

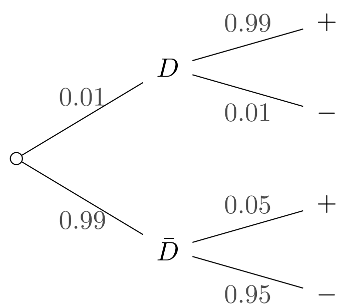
(A) 0.20

(B) 0.70

(C) 0.90

(D) 0.45

Q3. A rare disease affects 1% of a population. A screening test flags a diseased person as positive with probability 0.99 and flags a healthy person as positive with probability 0.05, as shown in the tree. A randomly chosen person tests positive. The probability that the person actually has the disease is nearest to: **[1 mark]**



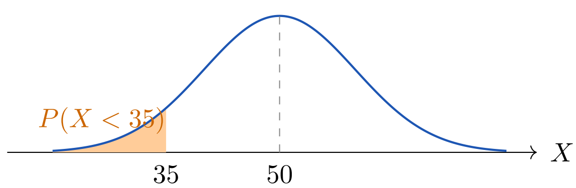
(A) 0.990

(B) 0.050

(C) 0.167

(D) 0.500

Q4. A measurement X is normally distributed with mean $\mu = 50$ and standard deviation $\sigma = 10$. Using the shaded left tail of the standard normal curve shown, the probability $P(X < 35)$ is nearest to (take $P(Z \le 1.5) = 0.9332$): **[1 mark]**



(A) 0.9332

(B) 0.0668

(C) 0.4332

(D) 0.5000

Q5. E-mails arrive at an inbox following a Poisson process with an average of $\lambda = 2$ e-mails per minute. The probability of receiving *exactly* 1 e-mail in a given one-minute window is nearest to (take $e^{-2} = 0.135$): **[1 mark]**

(A) 0.271

(B) 0.135

(C) 0.406

(D) 0.500

Q6. In a game a player wins Rs.10 with probability 0.3 and loses Rs.5 with probability 0.7. The expected net gain of the player per play (a loss shown as a negative amount) is: **[1 mark]**

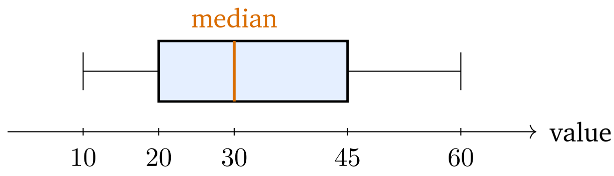
(A) Rs. 0.50

(B) $-$Rs. 0.50

(C) Rs. 3.00

(D) $-$Rs. 3.50

Q7. The box-and-whisker plot below summarizes a data set with five-number summary minimum $= 10$, first quartile $Q_1 = 20$, median $= 30$, third quartile $Q_3 = 45$ and maximum $= 60$. The interquartile range (IQR) of the data is: **[1 mark]**



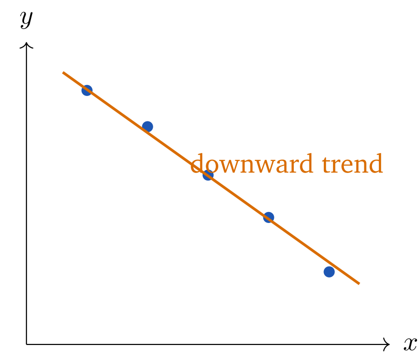
(A) 50

(B) 25

(C) 30

(D) 15

Q8. For a bivariate data set the sums of squares and cross-products about the means are $S_{xx} = 16$, $S_{yy} = 25$ and $S_{xy} = -12$, and the scatter shows a clear downward trend. The Pearson correlation coefficient r is: **[1 mark]**



(A) 0.60

(B) -0.75

(C) 0.48

(D) -0.60

Q9. A fair coin is tossed 5 times independently. The probability of obtaining *exactly* 2 heads is: **[1 mark]**

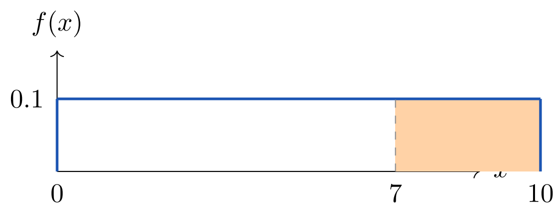
(A) 0.50000

(B) 0.15625

(C) 0.25000

(D) 0.31250

Q10. A continuous random variable X is uniformly distributed on the interval $[0, 10]$, so its density is the constant 0.1 shown below. The probability $P(X > 7)$, equal to the shaded area, is: **[1 mark]**



(A) 0.7

(B) 0.1

(C) 0.3

(D) 0.5

Linear Algebra, Calculus & Optimization

Q11. Consider the upper-triangular matrix $A = \begin{bmatrix} 3 & 1 & 7 \\ 0 & 2 & 4 \\ 0 & 0 & 5 \end{bmatrix}$. The sum of the eigenvalues of A is: **[1 mark]**

(A) 10

(B) 30

(C) 6

(D) 9

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

(A) 10

(B) 5

(C) 7

(D) 11

Q13. The *smaller* eigenvalue of the symmetric matrix $C = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}$ is: **[1 mark]**

(A) 7

(B) 5

(C) 3

(D) 10

Q14. A symmetric 2×2 matrix A has eigenvalues 5 and 2. The determinant of A is: **[1 mark]**

(A) 7

(B) 10

(C) 3

(D) 25

Q15. For the vector $u = (1,\ 2,\ 2)$, the Euclidean (ℓ_2) norm $\|u\|_2$ is: **[1 mark]**

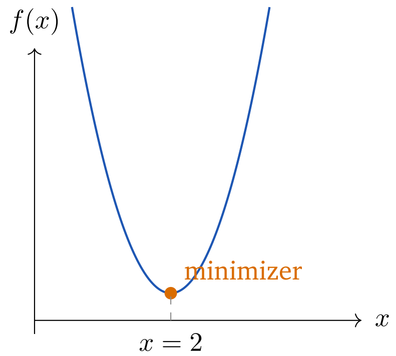
(A) 5

(B) 9

(C) 4

(D) 3

Q16. The convex function $f(x) = 2x^2 - 8x + 3$ is minimized over the real line, as suggested by its graph below. The value of x at which the minimum occurs (the minimizer) is: **[1 mark]**



(A) 2

(B) -2

(C) 3

(D) 4

Q17. For the scalar field $f(x, y) = x^2 + 3y$, the gradient ∇f evaluated at the point $(2,\ 1)$ is: **[1 mark]**

(A) $(2, 3)$

(B) $(4, 1)$

(C) $(7, 0)$

(D) $(4, 3)$

Q18. Gradient descent is used to minimize $f(x) = x^2 - 6x$ with a fixed learning rate $\eta = 0.1$. Starting from $x_0 = 0$, the value of x_1 after one update $x_1 = x_0 - \eta f'(x_0)$ is: **[1 mark]**

(A) -0.6

(B) 6.0

(C) 0.6

(D) 0.0

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

(A) convex everywhere

(B) neither convex nor concave

(C) linear (affine)

(D) concave everywhere

Programming, Data Structures & Algorithms

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

```
x = "hello"
print(x[::-1])
```

[1 mark]

(A) `hello`

(B) `hell`

(C) `olleh`

(D) `ohell`

Q21. Consider the following Python code:

```
d = {'a': 1, 'b': 2}
d['a'] = d.get('a', 0) + 10
print(d['a'])
```

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

(A) 11

(B) 1

(C) 10

(D) 0

Q22. The following operations are performed on an initially empty *queue* (FIFO), in order: `enqueue(5)`, `enqueue(8)`, `dequeue()`, `enqueue(12)`, `dequeue()`. The value returned by the *last* `dequeue()` is: **[2 marks]**

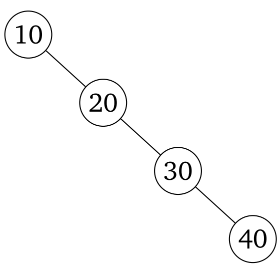
(A) 12

(B) 5

(C) 8

(D) 10

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



(A) 3

(B) 1

(C) 4

(D) 2

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) 5

(B) 3

(C) 4

(D) 24

Q25. An array of 64 distinct elements is repeatedly halved during a binary search. The maximum number of times the search interval can be halved until only one element remains is: **[2 marks]**

(A) 64

(B) 32

(C) 6

(D) 7

Q26. The *best-case* time complexity of the insertion sort algorithm (for example, when the input array is already sorted in ascending order) is: **[2 marks]**

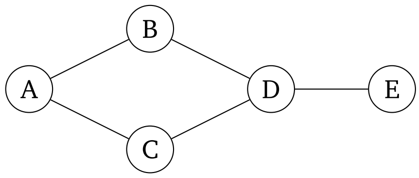
(A) $O(n)$

(B) $O(n^2)$

(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, D, B, E

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

Q28. The running time of an algorithm satisfies the recurrence $T(n) = 4\,T(n/2) + n$, with $T(1) = 1$. By the Master theorem, $T(n)$ is: **[2 marks]**

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

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

(C) $\Theta(n)$

(D) $\Theta(n^3)$

Database Management & Warehousing

Q29. Relation R has 5 attributes and relation S has 3 attributes, and the two relations share exactly 1 common attribute. The number of attributes (the degree) of the natural join $R \bowtie S$ is: **[2 marks]**

(A) 7

(B) 8

(C) 15

(D) 2

Q30. A table `EMP` has a `dept` column whose 6 rows contain the values `HR`, `HR`, `IT`, `IT`, `IT`, `Finance` (in some order, no `NULL`s). What does the query `SELECT COUNT(DISTINCT dept) FROM EMP;` return? **[2 marks]**

(A) 6

(B) 3

(C) 2

(D) 1

Q31. A relation $R(A, B, C)$ has the functional dependencies $AB \to C$ and $C \to B$. Its candidate keys are AB and AC. The highest normal form that R satisfies is: **[2 marks]**

(A) BCNF

(B) 3NF

(C) 1NF

(D) 2NF

Q32. Relation $R(A, B)$ contains the tuples $\{(1, x), (2, x), (3, y)\}$ and relation $S(B, C)$ contains $\{(x, 10), (x, 20), (y, 30), (z, 40)\}$. The number of tuples in the natural join $R \bowtie S$ (on the common attribute B) is: **[2 marks]**

(A) 4

(B) 7

(C) 3

(D) 5

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

(A) drill-down

(B) slice

(C) pivot

(D) roll-up

Q34. The figure shows an internal node of a B$^+$ tree of order $m = 4$ (each internal node holds up to m child pointers, drawn as •, separated by keys). For a *non-root* internal node of this tree, the *minimum* number of child pointers it must have is $\lceil m/2 \rceil$, which equals: **[2 marks]**

•	k_1	•	k_2	•	k_3	•

up to 4 pointers (•) and 3 keys

(A) 4

(B) 3

(C) 1

(D) 2

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

(A) $\{A\}$

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

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

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

Machine Learning & Artificial Intelligence

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

(A) 0.73

(B) 0.27

(C) 0.88

(D) 0.50

Q37. A very simple model produces a *high* error on the training set *and* a similarly high error on the test set. In terms of the bias–variance decomposition, this behaviour (underfitting) is best described as: **[2 marks]**

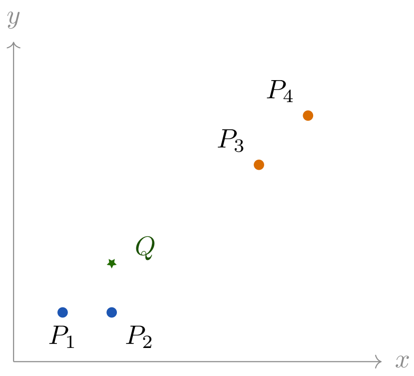
(A) high bias, low variance

(B) low bias, high variance

(C) low bias, low variance

(D) high bias, high variance

Q38. A 1-nearest-neighbour classifier (Euclidean distance) is trained on the labelled points $P_1(1,1)$: Class 0, $P_2(2,1)$: Class 0, $P_3(5,4)$: Class 1, $P_4(6,5)$: Class 1, shown below. The query point $Q(2,2)$ is classified as: **[2 marks]**



(A) Class 1

(B) cannot be determined

(C) Class 0

(D) a tie (undecided)

Q39. A node in a decision tree contains 4 positive and 4 negative examples. The entropy of this node (in bits) is: **[2 marks]**

(A) 0.000

(B) 0.500

(C) 0.811

(D) 1.000

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]**

$\boxed{3+,\ 1- \ \text{(total 4)}}$

(A) 0.500

(B) 0.250

(C) 0.375

(D) 0.750

Q41. The perpendicular distance from the point $(1,\ 1)$ to the line $x+y-4=0$, computed as $\dfrac{|w^{\top}x+b|}{\|w\|}$ with $w=(1,1)$ and $b=-4$, is nearest to: **[2 marks]**

(A) 1.414

(B) 2.000

(C) 4.000

(D) 0.500

Q42. In one iteration of the k-means algorithm a cluster is assigned the four points $(0,0)$, $(2,0)$, $(0,2)$ and $(2,2)$. The updated centroid of this cluster is: **[2 marks]**

(A) $(2,\ 2)$

(B) $(0,\ 0)$

(C) $(4,\ 4)$

(D) $(1,\ 1)$

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]**

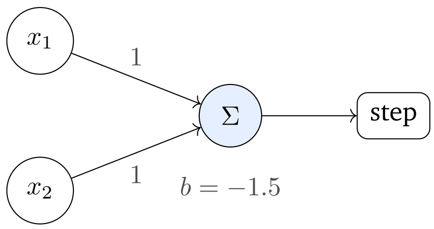
(A) 20%

(B) 80%

(C) 60%

(D) 40%

Q44. A perceptron with a step (threshold) activation computes step$(w_1x_1 + w_2x_2 + b)$, as shown, where step$(z) = 1$ if $z > 0$ and 0 otherwise. With inputs $x_1 = 1$, $x_2 = 1$, weights $w_1 = 1$, $w_2 = 1$ and bias $b = -1.5$, the output is: **[2 marks]**



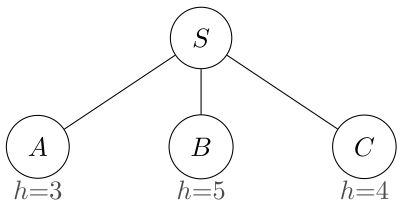
(A) 0

(B) 0.5

(C) -0.5

(D) 1

Q45. A *greedy best-first* search expands the frontier node with the smallest heuristic value h (it ignores the path cost g). The three frontier nodes A, B and C have the h values shown in the tree below. The node the search selects for expansion next is: **[2 marks]**



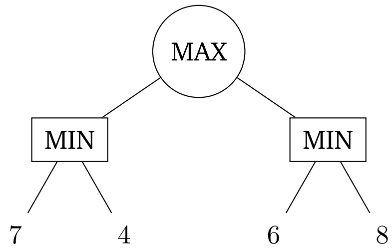
(A) A

(B) B

(C) C

(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) 6

(C) 7

(D) 8

Detailed Solutions

Q1.

Solution

Concept – probability without replacement: Multiply the probability of the first red draw by the conditional probability of the second red draw.

Step 1 – probability the first ball is red: There are 4 red out of 10 balls, so $P(\text{1st red}) = \frac{4}{10}$.

Step 2 – probability the second ball is red given the first was red: Now 3 red remain out of 9 balls, so $P(\text{2nd red} \mid \text{1st red}) = \frac{3}{9}$.

Step 3 – multiply the two: $P(\text{both red}) = \frac{4}{10} \times \frac{3}{9}$

$= \frac{12}{90}$

Step 4 – simplify: $= \frac{2}{15}$

Final Answer: $P(\text{both red}) = \frac{2}{15} \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q1

Q2.

Solution

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

Step 1 – compute the intersection using independence: $P(A \cap B) = 0.4 \times 0.5 = 0.20$

Step 2 – apply the inclusion–exclusion formula: $P(A \cup B) = 0.4 + 0.5 - 0.20$

Step 3 – add and subtract: $= 0.9 - 0.20$

$= 0.70$

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

Answer: (B) Go Back to Q2

Q3.

Solution

Concept – Bayes' theorem (reverse the tree): $P(D \mid +) = \dfrac{P(D)\,P(+\mid D)}{P(D)\,P(+\mid D) + P(\bar{D})\,P(+\mid \bar{D})}$.

Step 1 – probability along the diseased branch: $P(D)\,P(+\mid D) = 0.01 \times 0.99 = 0.0099$

Step 2 – probability along the healthy branch: $P(\bar{D})\,P(+\mid \bar{D}) = 0.99 \times 0.05 = 0.0495$

Step 3 – total probability of a positive test: $P(+) = 0.0099 + 0.0495 = 0.0594$

Step 4 – apply Bayes' theorem: $P(D\mid +) = \dfrac{0.0099}{0.0594}$

Step 5 – divide: $P(D\mid +) = 0.1667 \approx 0.167$

Final Answer: $P(D\mid +) \approx 0.167 \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q3

Q4.

Solution

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

Step 1 – standardize the cut-off $X = 35$: $Z = \dfrac{35-50}{10}$

$Z = \dfrac{-15}{10} = -1.5$

Step 2 – write the required probability in Z: $P(X < 35) = P(Z < -1.5)$

Step 3 – use symmetry of the normal curve: $P(Z < -1.5) = 1 - P(Z \leq 1.5)$

Step 4 – substitute $P(Z \leq 1.5) = 0.9332$: $= 1 - 0.9332 = 0.0668$

Final Answer: $P(X < 35) = 0.0668 \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q4

Q5.

Solution

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

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

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

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

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

Step 4 – substitute $e^{-2} = 0.135$**:** $P(X = 1) = 0.135 \times 2$

Step 5 – multiply: $P(X = 1) = 0.271$

Final Answer: $P(X = 1) \approx 0.271 \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q5

Q6.

Solution

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

Step 1 – contribution of the winning outcome: $(+10) \times 0.3 = +3.0$

Step 2 – contribution of the losing outcome: $(-5) \times 0.7 = -3.5$

Step 3 – add the two contributions: $E[\text{gain}] = 3.0 + (-3.5)$

$= -0.5$

Step 4 – interpret: The expected net result is a loss of Rs.0.50 per play, i.e. $-$Rs.0.50.

Final Answer: $E[\text{gain}] = -\text{Rs.}\,0.50 \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q6

Q7.

Solution

Concept – interquartile range: IQR $= Q_3 - Q_1$, the width of the box in a box plot.

Step 1 – read the quartiles from the plot: $Q_1 = 20$ (left edge of the box) and $Q_3 = 45$ (right edge of the box).

Step 2 – subtract: IQR $= 45 - 20$

$= 25$

Step 3 – note what IQR does not use: The minimum (10), median (30) and maximum (60) are not needed for the IQR.

Final Answer: IQR $= 25 \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q7

Q8.

Solution

Concept – Pearson correlation coefficient: $r = \dfrac{S_{xy}}{\sqrt{S_{xx}\,S_{yy}}}$.

Step 1 – compute the product under the root: $S_{xx}\,S_{yy} = 16 \times 25 = 400$

Step 2 – take the square root: $\sqrt{400} = 20$

Step 3 – substitute into the formula: $r = \dfrac{-12}{20}$

Step 4 – simplify: $r = -0.60$

Step 5 – sanity check the sign: $S_{xy} < 0$ matches the downward scatter, so r must be negative.

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

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: $n = 5$, $k = 2$, $p = 0.5$, $1 - p = 0.5$.

Step 2 – evaluate the binomial coefficient: $\binom{5}{2} = \dfrac{5 \times 4}{2 \times 1} = 10$

Step 3 – evaluate the power term: $p^2(1-p)^3 = 0.5^5 = \dfrac{1}{32} = 0.03125$

Step 4 – multiply: $P(X = 2) = 10 \times 0.03125$

$= 0.3125$

Final Answer: $P(X = 2) = 0.3125 \Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q9

Q10.

Solution

Concept – probability for a uniform distribution: For $X \sim \text{Uniform}(a, b)$, $P(X > c) = \dfrac{b-c}{b-a}$ (area of the shaded rectangle).

Step 1 – identify the interval and cut-off: $a = 0$, $b = 10$, $c = 7$.

Step 2 – length of the shaded part: $b - c = 10 - 7 = 3$

Step 3 – total length of the support: $b - a = 10 - 0 = 10$

Step 4 – form the ratio (area = base × height $= 3 \times 0.1$**):** $P(X > 7) = \frac{3}{10} = 0.3$

Final Answer: $P(X > 7) = 0.3 \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q10

Q11.

Solution

Concept – eigenvalues of a triangular matrix: For a triangular matrix the eigenvalues are exactly the diagonal entries, and their sum equals the trace.

Step 1 – read the diagonal entries: The diagonal of A is $3, 2, 5$, so the eigenvalues are 3, 2 and 5.

Step 2 – add the eigenvalues: $3 + 2 + 5$

Step 3 – compute the sum: $= 10$

Step 4 – cross-check with the trace: $\mathrm{tr}(A) = 3 + 2 + 5 = 10$, which matches.

Final Answer: sum of eigenvalues $= 10 \Rightarrow \boxed{\text{A}}$

Answer: (A) 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: $b_{11} = 2$, $b_{12} = 0$, $b_{13} = 1$.

Step 2 – minor for b_{11}: $\begin{vmatrix} 1 & 2 \\ 0 & 3 \end{vmatrix} = (1)(3) - (2)(0) = 3$

Step 3 – the middle term vanishes: Since $b_{12} = 0$, its contribution is 0.

Step 4 – minor for b_{13}: $\begin{vmatrix} 3 & 1 \\ 1 & 0 \end{vmatrix} = (3)(0) - (1)(1) = -1$

Step 5 – combine with the sign pattern $(+, -, +)$: $\det B = 2(3) - 0 + 1(-1)$

$= 6 - 1$

$= 5$

Final Answer: $\det B = 5 \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q12

Q13.

Solution

Concept – eigenvalues of a 2×2 symmetric matrix: For $\begin{bmatrix} a & b \\ b & a \end{bmatrix}$ the eigenvalues are $a + b$ and $a - b$.

Step 1 – identify a and b: Here $a = 5$ and $b = 2$.

Step 2 – compute the two eigenvalues: $a + b = 5 + 2 = 7$

$a - b = 5 - 2 = 3$

Step 3 – pick the smaller one: $\min(7, 3) = 3$

Step 4 – cross-check with the trace: $7 + 3 = 10 = \mathrm{tr}(C)$, consistent.

Final Answer: smaller eigenvalue $= 3 \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q13

Q14.

Solution

Concept – determinant equals the product of eigenvalues: $\det(A) = \lambda_1 \lambda_2$.

Step 1 – list the eigenvalues: $\lambda_1 = 5$ and $\lambda_2 = 2$.

Step 2 – multiply them: $\det(A) = 5 \times 2$

Step 3 – compute: $= 10$

Step 4 – note the contrast: The *trace* would be $5 + 2 = 7$; the question asks for the determinant, which is the product 10.

Final Answer: $\det(A) = 10 \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q14

Q15.

Solution

Concept – the ℓ_2 (Euclidean) norm: $\|u\|_2 = \sqrt{\sum_i u_i^2}$.

Step 1 – square each component: $1^2 = 1, \quad 2^2 = 4, \quad 2^2 = 4$.

Step 2 – add the squares: $1 + 4 + 4 = 9$

Step 3 – take the square root: $\|u\|_2 = \sqrt{9} = 3$

Step 4 – contrast with the ℓ_1 norm: $\|u\|_1 = 1+2+2 = 5$; the Euclidean norm asked for is 3.

Final Answer: $\|u\|_2 = 3 \Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q15

Q16.

Solution

Concept – minimizer of a convex parabola: For $f(x) = ax^2 + bx + c$ with $a > 0$, the minimum is at $x^* = -\dfrac{b}{2a}$.

Step 1 – identify the coefficients: $a = 2$, $b = -8$, $c = 3$.

Step 2 – substitute into the formula: $x^* = -\dfrac{-8}{2 \times 2}$

Step 3 – simplify the denominator: $= \dfrac{8}{4}$

Step 4 – divide: $x^* = 2$

Step 5 – confirm via the derivative: $f'(x) = 4x - 8 = 0 \Rightarrow x = 2$, and $f''(x) = 4 > 0$ confirms a minimum.

Final Answer: minimizer $x^* = 2 \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q16

Q17.

Solution

Concept – the gradient vector: $\nabla f = \left(\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}\right)$.

Step 1 – partial derivative with respect to x: $\dfrac{\partial}{\partial x}(x^2 + 3y) = 2x$

Step 2 – partial derivative with respect to y: $\dfrac{\partial}{\partial y}(x^2 + 3y) = 3$

Step 3 – evaluate at the point $(2, 1)$: $2x = 2 \times 2 = 4$ and the second component is the constant 3.

Step 4 – assemble the gradient: $\nabla f(2, 1) = (4,\ 3)$

Final Answer: $\nabla f(2, 1) = (4, 3) \Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q17

Q18.

Solution

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

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

Step 2 – evaluate the derivative at $x_0 = 0$: $f'(0) = 2(0) - 6 = -6$

Step 3 – substitute into the update rule: $x_1 = 0 - 0.1 \times (-6)$

Step 4 – simplify: $x_1 = 0 + 0.6$

$x_1 = 0.6$

Final Answer: $x_1 = 0.6 \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q18

Q19.

Solution

Concept – convexity from the Hessian: A function is convex if its Hessian is positive semidefinite everywhere, and concave if it is negative semidefinite everywhere.

Step 1 – second partial derivatives: $f_{xx} = -2, \quad f_{yy} = -2, \quad f_{xy} = 0.$

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

Step 3 – find its eigenvalues: Both diagonal entries are eigenvalues: -2 and -2.

Step 4 – classify: Both eigenvalues are negative, so H is negative definite everywhere; hence f is concave everywhere.

Final Answer: f is concave everywhere $\Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q19

Q20.

Solution

Concept – Python extended slice for reversal: The slice `[::-1]` steps through the sequence with step -1, producing it in reverse.

Step 1 – identify the string: `x = "hello"` has characters h, e, l, l, o.

Step 2 – apply the reverse slice: Reading the characters from last to first gives o, l, l, e, h.

Step 3 – join into the output string: `"olleh"`

Final Answer: output is `olleh` $\Rightarrow$ $\boxed{\text{C}}$

$\boxed{\textbf{Answer: (C)}}$ **Go Back to Q20**

Q21.

Solution

Concept – `dict.get` with a default: `d.get(k, default)` returns the value stored at key `k` if present, otherwise `default`.

Step 1 – evaluate `d.get('a', 0)`: Key `'a'` exists with value 1, so `d.get('a', 0)` returns 1 (the default 0 is ignored).

Step 2 – add 10**:** $1 + 10 = 11$

Step 3 – assign back to `d['a']`: `d['a']` now holds 11.

Step 4 – print: `print(d['a'])` outputs 11.

Final Answer: the value printed is $11 \Rightarrow \boxed{\text{A}}$

$\boxed{\textbf{Answer: (A)}}$ **Go Back to Q21**

Q22.

Solution

Concept – queue is First-In-First-Out (FIFO): `dequeue` removes the element that has been in the queue the longest.

Step 1 – after `enqueue(5)`, `enqueue(8)`: Queue (front $\to$ rear): $[5,\ 8]$.

Step 2 – first `dequeue()`: Removes the front element 5; queue becomes $[8]$.

Step 3 – `enqueue(12)`: Queue becomes $[8,\ 12]$.

Step 4 – second (last) `dequeue()`: Removes the current front element, which is 8.

Final Answer: the last `dequeue()` returns $8 \Rightarrow \boxed{\text{C}}$

$\boxed{\textbf{Answer: (C)}}$ **Go Back to Q22**

Q23.

Solution

Concept – BST height with sorted insertion: Inserting keys already in increasing order gives a right-skewed chain; the height is the number of edges on the longest path.

Step 1 – insert 10: 10 becomes the root.

Step 2 – insert 20: $20 > 10$, so it goes to the right of 10.

Step 3 – insert 30 and 40: Each new key is larger than all before it, so it keeps attaching to the right: $10 \to 20 \to 30 \to 40$.

Step 4 – count the edges on the longest path: Root-to-deepest-leaf path is $10 \to 20 \to 30 \to 40$, which has 3 edges.

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

Answer: (A) Go Back to Q23

Q24.

Solution

Concept – linear probing: On a collision at index i, try $i+1$, $i+2$, $\dots$ (modulo the table size) until an empty slot is found.

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

Step 2 – insert 17: $h(17) = 17 \bmod 7 = 3$; slot 3 is taken, so probe index 4, which is empty $\Rightarrow$ 17 at index 4.

Step 3 – insert 24: $h(24) = 24 \bmod 7 = 3$; slot 3 taken, probe 4 taken, probe 5 empty $\Rightarrow$ 24 at index 5.

Step 4 – read off the required index: Key 24 occupies index 5.

Final Answer: final index of 24 is $5 \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q24

Q25.

Solution

Concept – halving to a single element: Repeated halving of n elements can be done about $\log_2 n$ times.

Step 1 – start with the size: $n = 64$.

Step 2 – halve step by step: $64 \to 32 \to 16 \to 8 \to 4 \to 2 \to 1$

Step 3 – count the halving steps: $64 \to 32$ (1), $32 \to 16$ (2), $16 \to 8$ (3), $8 \to 4$ (4), $4 \to 2$ (5), $2 \to 1$ (6): 6 halvings.

Step 4 – confirm with logarithm: $\log_2 64 = 6$.

Final Answer: number of halvings $= 6 \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q25

Q26.

Solution

Concept – insertion sort best case: When the array is already sorted, each new element needs only one comparison and no shifts.

Step 1 – count comparisons per element: For an already-sorted array, inserting the i-th element requires exactly 1 comparison (it stays in place).

Step 2 – total over $n-1$ insertions: Total comparisons $= (n-1) \times 1 = n-1$.

Step 3 – express asymptotically: $n - 1 = O(n)$.

Step 4 – contrast with the worst case: A reverse-sorted array needs $O(n^2)$; the best case here is the linear $O(n)$.

Final Answer: best-case time $= O(n) \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q26

Q27.

Solution

Concept – depth-first search (DFS): DFS goes as deep as possible along one branch before backtracking, choosing unvisited neighbours in alphabetical order here.

Step 1 – start at A: Visit A. Neighbours of A: B, C; pick B (alphabetical).

Step 2 – go deep from B: Visit B. Neighbours: A (visited), D; go to D.

Step 3 – continue from D: Visit D. Neighbours: B (visited), C, E; pick C (alphabetical).

Step 4 – from C then backtrack: Visit C. Neighbours A, D both visited, so backtrack to D and visit its remaining neighbour E.

Step 5 – assemble the order: $A,\ B,\ D,\ C,\ E$.

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

Answer: (B) Go Back to Q27

Q28.

Solution

Concept – Master theorem: For $T(n) = a\,T(n/b) + f(n)$, compare $f(n)$ with $n^{\log_b a}$.

Step 1 – identify the parameters: $a = 4$, $b = 2$, $f(n) = n$.

Step 2 – compute the critical exponent: $\log_b a = \log_2 4 = 2$, so $n^{\log_b a} = n^2$.

Step 3 – compare $f(n) = n$ with n^2: $n = O(n^{2-\varepsilon})$ for $\varepsilon = 1$, so we are in Case 1 (the recursion cost dominates).

Step 4 – write the result: $T(n) = \Theta(n^{\log_b a}) = \Theta(n^2)$.

Final Answer: $T(n) = \Theta(n^2) \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q28

Q29.

Solution

Concept – degree of a natural join: The natural join keeps each common attribute once, so its degree is (degree of R) + (degree of S) − (number of common attributes).

Step 1 – list the degrees: Degree of $R = 5$, degree of $S = 3$, common attributes $= 1$.

Step 2 – apply the formula: degree$(R \bowtie S) = 5 + 3 - 1$

Step 3 – compute: $= 7$

Step 4 – note the contrast: A *Cartesian product* would keep all $5+3 = 8$ attributes; the natural join drops the one duplicated attribute, giving 7.

Final Answer: degree $= 7 \Rightarrow \boxed{\text{A}}$

Answer: (A) Go Back to Q29

Q30.

Solution

Concept – `COUNT(DISTINCT col)`: This counts how many *different* non-null values the column holds.

Step 1 – list the values: `HR`, `HR`, `IT`, `IT`, `IT`, `Finance`.

Step 2 – keep only distinct values: The distinct values are `HR`, `IT`, `Finance`.

Step 3 – count them: 3 distinct values.

Step 4 – contrast with `COUNT(dept)`: Plain `COUNT(dept)` would return 6 (all rows); DISTINCT collapses duplicates to 3.

Final Answer: the query returns $3 \Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q30

Q31.

Solution

Concept – BCNF vs 3NF: BCNF requires the left side of every non-trivial FD to be a superkey. 3NF also allows an FD $X \to A$ when A is a *prime* attribute (part of some candidate key).

Step 1 – note the candidate keys: AB and AC; hence the prime attributes are A, B and C.

Step 2 – check $AB \to C$ for BCNF: AB is a candidate key (a superkey), so this FD satisfies BCNF.

Step 3 – check $C \to B$ for BCNF: C is *not* a superkey, so $C \to B$ violates BCNF.

Step 4 – check $C \to B$ for 3NF: The right-hand side B *is* a prime attribute (it lies in candidate key AB), so $C \to B$ is allowed under 3NF.

Step 5 – conclude: R meets 3NF but not BCNF, so the highest normal form is 3NF.

Final Answer: highest normal form $=$ 3NF $\Rightarrow \boxed{\text{B}}$

Answer: (B) Go Back to Q31

Q32.

Solution

Concept – natural join matches on the common attribute: Each tuple of R pairs with every tuple of S that has the same B-value.

Step 1 – group S by its B-value: $B = x$ has 2 tuples $\{(x, 10), (x, 20)\}$; $B = y$ has 1; $B = z$ has 1.

Step 2 – match R's tuple $(1, x)$: $B = x$ matches 2 tuples of $S \Rightarrow 2$ output rows.

Step 3 – match R's tuple $(2, x)$: $B = x$ again matches 2 tuples $\Rightarrow 2$ output rows.

Step 4 – match R's tuple $(3, y)$: $B = y$ matches 1 tuple $\Rightarrow 1$ output row. (The S tuple with $B = z$ has no partner in R.)

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

Final Answer: number of tuples $= 5 \Rightarrow \boxed{\text{D}}$

$\boxed{\text{Answer: (D)}}$ Go Back to Q32

Q33.

Solution

Concept – OLAP slice: A *slice* fixes one dimension at a single value, cutting a lower-dimensional sub-cube out of the data cube.

Step 1 – interpret the operation: Fixing `year = 2024` selects one value along the time dimension.

Step 2 – match to the OLAP vocabulary: Selecting one value of a single dimension is exactly a slice.

Step 3 – rule out the others: Roll-up aggregates to a coarser level, drill-down goes finer, and pivot rotates the axes; none fixes a single dimension value.

Final Answer: the operation is a slice $\Rightarrow \boxed{\text{B}}$

$\boxed{\text{Answer: (B)}}$ Go Back to Q33

Q34.

Solution

Concept – minimum fan-out of a B$^+$ tree node: A non-root internal node of order m must have at least $\lceil m/2 \rceil$ child pointers.

Step 1 – substitute the order: $m = 4$, so the minimum is $\lceil 4/2 \rceil$.

Step 2 – evaluate the ceiling: $4/2 = 2$, and $\lceil 2 \rceil = 2$.

Step 3 – state the minimum: A non-root internal node must have at least 2 child pointers.

Step 4 – sanity check against the maximum: The maximum is $m = 4$ pointers (and 3 keys); the minimum 2 is within this range.

Final Answer: minimum child pointers $= 2 \Rightarrow \boxed{\text{D}}$

$\boxed{\text{Answer: (D)}}$ 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 A itself: $A^+ = \{A\}$.

Step 2 – apply $A \to B$: A determines B, so add B: $A^+ = \{A, B\}$.

Step 3 – apply $B \to C$: B is now in the closure and determines C, so add C: $A^+ = \{A, B, C\}$.

Step 4 – check for more: No FD has a left side made only of $\{A, B, C\}$ that adds D or E, so the closure stops.

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

Answer: (C) Go Back to Q35

Q36.

Solution

Concept – logistic regression forward pass: Compute $z = w^\top x + b$, then apply $\sigma(z) = \dfrac{1}{1 + e^{-z}}$.

Step 1 – dot product $w^\top x$: $w^\top x = (2)(1) + (-1)(2)$

$= 2 - 2$

$= 0$

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: $= \dfrac{1}{2} = 0.50$

Final Answer: $\sigma(z) = 0.50 \Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q36

Q37.

Solution

Concept – bias–variance for underfitting: High error on both training and test data means the model is too simple to capture the pattern.

Step 1 – read the symptom: Training error is high *and* test error is high (and close to each other).

Step 2 – link high training error to bias: Being unable to fit even the training data indicates *high bias*.

Step 3 – link the stability across sets to variance: Because the two errors are similar and the model is simple, it is not sensitive to the particular sample, i.e. *low variance*.

Step 4 – name the regime: High bias with low variance is classic underfitting.

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

Answer: (A) Go Back to Q37

Q38.

Solution

Concept – 1-nearest-neighbour: Classify the query with the label of its single closest training point (compare squared distances to avoid roots).

Step 1 – squared distance to $P_1(1,1)$: $(2-1)^2+(2-1)^2 = 1+1 = 2$

Step 2 – squared distance to $P_2(2,1)$: $(2-2)^2+(2-1)^2 = 0+1 = 1$

Step 3 – squared distances to $P_3(5,4)$ and $P_4(6,5)$: P_3: $(2-5)^2+(2-4)^2 = 9+4 = 13$; P_4: $(2-6)^2+(2-5)^2 = 16+9 = 25$.

Step 4 – pick the smallest: The minimum is 1, at P_2, whose label is Class 0.

Final Answer: Q is classified as Class $0 \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q38

Q39.

Solution

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

Step 1 – compute the proportions: $p = \frac{4}{8} = 0.5$ and $q = \frac{4}{8} = 0.5$.

Step 2 – substitute into the entropy formula: $H = -0.5\log_2 0.5 - 0.5\log_2 0.5$

Step 3 – use $\log_2 0.5 = -1$: $H = -0.5(-1) - 0.5(-1)$

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

Step 5 – interpret: A perfectly balanced node has maximum entropy of 1 bit.

Final Answer: $H = 1.0$ bit $\Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q39

Q40.

Solution

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

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

Step 2 – square each proportion: $0.75^2 = 0.5625$ and $0.25^2 = 0.0625$.

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

Step 4 – subtract from 1: $\text{Gini} = 1 - 0.625 = 0.375$

Final Answer: $\text{Gini} = 0.375 \Rightarrow \boxed{\text{C}}$

Answer: (C) Go Back to Q40

Q41.

Solution

Concept – distance from a point to a hyperplane: $d = \dfrac{|w^\top x + b|}{\|w\|}$.

Step 1 – evaluate the numerator $w^\top x + b$: $(1)(1) + (1)(1) + (-4) = 1 + 1 - 4 = -2$

Step 2 – take the absolute value: $|-2| = 2$

Step 3 – compute $\|w\|$: $\|(1, 1)\| = \sqrt{1^2 + 1^2} = \sqrt{2}$

Step 4 – divide: $d = \dfrac{2}{\sqrt{2}} = \sqrt{2} \approx 1.414$

Final Answer: $d \approx 1.414 \Rightarrow \boxed{\text{A}}$

Answer: (A) 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: $\bar{x} = \dfrac{0+2+0+2}{4} = \dfrac{4}{4} = 1$

Step 2 – average the y-coordinates: $\bar{y} = \dfrac{0+0+2+2}{4} = \dfrac{4}{4} = 1$

Step 3 – assemble the centroid: $(\bar{x}, \bar{y}) = (1,\ 1)$

Final Answer: centroid $= (1, 1) \Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q42

Q43.

Solution

Concept – variance explained in PCA: The first principal component explains a fraction equal to its eigenvalue divided by the sum of all eigenvalues.

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

Step 2 – variance of the first component: The largest eigenvalue is 6.

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

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

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

Answer: (C) Go Back to Q43

Q44.

Solution

Concept – perceptron with a step activation: Compute the weighted sum plus bias, then output 1 if it is positive, else 0.

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

$= 1 + 1$

$= 2$

Step 2 – add the bias: $z = 2 + (-1.5) = 0.5$

Step 3 – apply the step function: $z = 0.5 > 0$, so $\text{step}(0.5) = 1$.

Step 4 – interpret: This is the logical AND behaviour: both inputs 1 give output 1.

Final Answer: output $= 1 \Rightarrow \boxed{\text{D}}$

Answer: (D) Go Back to Q44

Q45.

Solution

Concept – greedy best-first search: It expands the frontier node with the smallest heuristic h, ignoring the path cost g.

Step 1 – list the heuristic values: $h(A) = 3$, $h(B) = 5$, $h(C) = 4$.

Step 2 – pick the smallest h: $\min(3, 5, 4) = 3$, which belongs to node A.

Step 3 – select the node: Greedy best-first expands A.

Step 4 – contrast with A*: A* would use $f = g + h$; here only h matters, so A (smallest h) is chosen.

Final Answer: the search expands $A \Rightarrow \boxed{\text{A}}$

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: $\min(6,\ 8) = 6$

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

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

Answer: (B) Go Back to Q46

Answer Key

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