



Data Science and Artificial Intelligence (DA)

Note: The DA test paper is newly introduced in GATE 2024. This is a sample question paper for the candidates to practice the subject. We will not be able to answer any queries or provide answer keys to these questions.

Q.1 – Q.25 Carry ONE mark each.	
Q.1	Let b be the branching factor of a search tree. If the optimal goal is reached after d actions from the initial state, in the worst case, how many times will the initial state be expanded for iterative deepening depth first search (IDDFS) and iterative Deepening A* search (IDA*)?
(A)	IDDFS - d , IDA* - d .
(B)	IDDFS - d , IDA* - b^d
(C)	IDDFS - b^d , IDA* - d .
(D)	IDDFS - b^d , IDA* - b^d .
Q.2	Given 3 literals A , B , and C , how many models are there for the sentence $A \vee \neg B \vee C$?

Q.3	Which of the following first order logic sentence matches closest with the sentence “All students are not equal”?
(A)	$\forall x \exists y[student(x) \wedge student(y)] \Rightarrow \neg Equal(x, y)$
(B)	$\forall x \forall y[student(x) \wedge student(y)] \Rightarrow \neg Equal(x, y)$
(C)	$\forall x \exists y[student(x) \wedge student(y) \wedge \neg Equal(x, y)]$
(D)	$\forall x \forall y[student(x) \wedge student(y) \wedge \neg Equal(x, y)]$
Q.4	The mean of the observations of the first 50 observations of a process is 12. If the 51st observation is 18, then, the mean of the first 51 observations of the process is
(A)	12
(B)	12.12
(C)	12.36
(D)	18

Q.5	$\lim_{x \rightarrow 2} \frac{\sqrt{x} - \sqrt{2}}{x - 2}$
(A)	0
(B)	$\sqrt{2}$
(C)	$\frac{1}{2\sqrt{2}}$
(D)	$\frac{1}{\sqrt{2}}$
Q.6	Which among the following may help to reduce overfitting demonstrated by a model i) Change the loss function. ii) Reduce model complexity. iii) Increase the training data. iv) Increase the number of optimization routine steps.
(A)	ii and i
(B)	ii and iii
(C)	i, ii, and iii
(D)	i, ii, iii, and iv

Q.7	A fair coin is flipped twice and it is known that at least one tail is observed. The probability of getting two tails is
(A)	$\frac{1}{2}$
(B)	$\frac{1}{3}$
(C)	$\frac{2}{3}$
(D)	$\frac{1}{4}$
Q.8	Given n indistinguishable particles and $m (> n)$ distinguishable boxes, we place at random each particle in one of the boxes. The probability that in n preselected boxes, one and only one particle will be found is
(A)	$\frac{n!}{m^n}$
(B)	$\frac{(m-1)! n!}{(m+n-1)!}$
(C)	$\frac{1}{m^n}$
(D)	$\frac{1}{m}$

Q.9	For two events A and B, $B \subset A$ Which of the following statement is correct?
(A)	$P(B A) \geq P(B)$
(B)	$P(B A) \leq P(B)$
(C)	$P(A B) < 1$
(D)	$P(A B) = 0$
Q.10	X is a uniform distribution random variable with support in $[-2, 2] \cup [99.5, 100.5]$. The mean of X is _____
Q. 11	You are reviewing four papers submitted to a conference on machine learning for medical expert systems. All the four papers validate their superiority on a standard benchmarking cancer dataset, which has only 5% of positive cancer cases. Which of the experimental setting is acceptable to you? paper i) We evaluated the performance of our model through a 5-fold cross validation process and report an accuracy of 93%. paper ii) The area under the ROC curve on a single left out test set of our model is around 0.8, which is the highest among all the different approaches. paper iii) We computed the average area under the ROC curve through 5-fold cross validation and found it to be around 0.75 – the highest among all the approaches.

	paper iv) The accuracy on a single left out test set of our model is 95%, which is the highest among all the different approaches.
(A)	paper i
(B)	paper i and paper iv
(C)	paper ii and paper iv
(D)	paper iii
Q.12	Increasing the regularizing coefficient value for a ridge regressor will i) Increase or maintain model bias. ii) Decrease model bias. iii) Increase or maintain model variance. iv) Decrease model variance.
(A)	i and iii
(B)	i and iv
(C)	ii and iii
(D)	ii and iv
Q.13	A decision tree classifier learned from a fixed training set achieves 100% accuracy. Which of the following models trained using the same training set will also achieve 100% accuracy? i) Logistic regressor. ii) A polynomial of degree one kernel SVM.

(A)	i
(B)	i and ii
(C)	all of the above
(D)	none of the above
Q.14	Consider two relations $R(x, y)$ and $S(x, z)$. Relation R has 100 records, and relation S has 200 records. What will be the number of attributes and records of the following query? SELECT * from R CROSS JOIN S;
(A)	3 attributes, 20000 records
(B)	4 attributes, 20000 records
(C)	3 attributes, 200 records
(D)	4 attributes, 200 records

Q.15	Consider two relations $R(x, y)$ and $S(y)$, and perform the following operation $R(x,y) \text{ DIVIDE } S(Y)$ If X is the relation returns by the above operation, which of the following option(s) is/are always TRUE?
(A)	$ X \leq R $
(B)	$ X \leq S $
(C)	$ X \leq R \text{ AND } X \leq S $
(D)	All of the Above
Q.16	Which of the following statements is/are TRUE?
(A)	Every relation with two attributes is also in BCNF.
(B)	Every relation in BCNF is also in 3NF.
(C)	No relation can be in both BCNF and 3NF.
(D)	None of Above

Q.17	For matrix $H = \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix}$, one of the eigenvalues is 5. Then, the other eigenvalue is
(A)	12
(B)	10
(C)	8
(D)	6
Q. 18	Two non-zero vectors $\mathbf{x}$ and $\mathbf{y}$ are perpendicular if
(A)	$\mathbf{x}^T \mathbf{y} = 0$
(B)	$\mathbf{x}^T \mathbf{y} > 0$
(C)	$\mathbf{x}^T \mathbf{y} < 0$
(D)	None of the above.

Q. 19	The function $f(x) = 1 + x + x^2$ has a												
(A)	Minima at $x = -0.5$												
(B)	Maxima at $x = -0.5$												
(C)	Saddle point at $x = -0.5$												
(D)	None of the above.												
Q. 20	The Pearson's correlation coefficient between x and y rounded to the first decimal point for the given data in below table is <table border="1"> <thead> <tr> <th>x</th><th>y</th></tr> </thead> <tbody> <tr> <td>-6</td><td>6.4</td></tr> <tr> <td>2</td><td>4.7</td></tr> <tr> <td>0.2</td><td>8</td></tr> <tr> <td>7</td><td>2</td></tr> <tr> <td>-4</td><td>3.4</td></tr> </tbody> </table>	x	y	-6	6.4	2	4.7	0.2	8	7	2	-4	3.4
x	y												
-6	6.4												
2	4.7												
0.2	8												
7	2												
-4	3.4												
(A)	-0.5												
(B)	0.5												
(C)	0.3												
(D)	-0.3												

Q.21	The worst-case running times of Insertion sort, Merge sort and Quick sort respectively are
(A)	$\Theta(n \log n)$, $\Theta(n^2)$, $\Theta(n^2)$
(B)	$\Theta(n^2)$, $\Theta(n \log n)$, $\Theta(n \log n)$
(C)	$\Theta(n^2)$, $\Theta(n \log n)$, $\Theta(n^2)$
(D)	$\Theta(n^2)$, $\Theta(n^2)$, $\Theta(n \log n)$
Q.22	Consider the following program. <pre> int func(int n) { if(n <= 1) return n; else return 3*func(n-3) - 3*func(n-2); end } </pre> The running time of the above function is
(A)	$\Theta(n)$
(B)	$\Theta(n^2)$
(C)	$\Theta(3^n)$
(D)	$\Theta(2^n)$
Q.23	Which of the following correctly describes the recurrence relation for the standard binary search algorithm on a sorted array of n numbers where c is a constant.

(A)	$T(n) = 2 * T(n/2) + c$
(B)	$T(n) = T(n/2)$
(C)	$T(n) = T(n-1) + c$
(D)	$T(n) = T(n/2) + c$
Q.24	Consider the following C program <pre> int func(int A[], int n, int m) { int s = A[0]; for(int i = 1; i <= n - 1; i++) total = m*s + A[i]; return m; } </pre> Let Z be an array of 10 elements with $Z[i] = 2$ for all i such that $0 \leq i \leq 9$; The value returned by <code>func(Z,10,2)</code> is _____
Q.25	Two eigenvalues of 3×3 matrix X are $(1 + i)$ and 2. The determinant of the matrix X is _____

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：

<https://d.book118.com/876014200053010121>