

AiTI PhD Entrance: Sample Questions for Online Written Examination

Unit 1: General Aptitude & Communication (Weightage: 15%)

Q1. If '→' denotes increasing order of intensity, then the meaning of the words [dry → arid → parched] is analogous to [diet → fast → _____]. Which one of the given options is appropriate to fill the blank?

- A) starve
- B) reject
- C) feast
- D) deny

Q2. The number of coins of Rs.1, Rs.5, and Rs.10 denominations that a person has are in the ratio 5: 3: 13. Of the total amount, the percentage of money in Rs.5 coins is

- A) 21%
- B) $14\frac{2}{7}\%$
- C) 10%
- D) 30%

Q3. For positive non-zero real variables p and q , if $\log(p^2 + q^2) = \log p + \log q + 2\log 3$, then the value of $\frac{p^4 + q^4}{p^2 q^2}$ is

- A) 79
- B) 81
- C) 9
- D) 83

Q4. In an engineering college of 10,000 students, 1,500 like neither their core branches nor other branches. The number of students who like their core branches is $\frac{1}{4}$ th of the number who like other branches. The number who like both is 500. The number of students who like their core branches is

- A) 1,800
- B) 3,500
- C) 1,600
- D) 1,500

Q5. In the sequence 6,9,14, x , 30,41, a possible value of x is

- A) 25

- B) 21
- C) 18
- D) 20

Unit 2: Research Aptitude, Methodology & Ethics (Weightage: 5%)

Q6. Which one of the following is an example of non-probability sampling?

- A) Simple random sampling
- B) Stratified sampling
- C) Convenience sampling
- D) Systematic sampling

Unit 3: Engineering Mathematics for AI (Weightage: 20%)

Q7. Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be a function such that $f(x) = \max\{x, x^3\}$, $x \in \mathbb{R}$. The set of all points where $f(x)$ is NOT differentiable is

- A) $\{-1, 1, 2\}$
- B) $\{-2, -1, 1\}$
- C) $\{0, 1\}$
- D) $\{-1, 0, 1\}$

Q8. The sum of the elements in each row of a matrix $A \in \mathbb{R}^{n \times n}$ is 1. If $B = A^3 - 2A^2 + A$, which one of the following statements is correct (for $x \in \mathbb{R}^n$)?

- A) The equation $Bx = 0$ has no solution
- B) The equation $Bx = 0$ has exactly two solutions
- C) The equation $Bx = 0$ has infinitely many solutions
- D) The equation $Bx = 0$ has a unique solution

Q9. Let $f(x) = \frac{e^x - e^{-x}}{2}$, $x \in \mathbb{R}$ (the hyperbolic sine function). Let $f^{(k)}(a)$ denote the k -th derivative of f evaluated at a . What is the value of $f^{(10)}(0)$?

- A) 0
- B) 1
- C) $1/10!$
- D) $2/10!$

Q10. Let p and q be any two propositions. Consider $S_1: p \rightarrow q$, $S_2: \neg p \wedge q$, $S_3: \neg p \vee q$, $S_4: \neg p \vee \neg q$. Which one of the following options is correct? ($\equiv$ denotes logical equivalence)

- A) $S_1 \equiv S_3$
- B) $S_2 \equiv S_3$
- C) $S_2 \equiv S_4$
- D) $S_1 \equiv S_4$

Q11. Let $A \in \mathbb{R}^{n \times n}$ be such that $A^3 = A$. Which one of the following statements is ALWAYS correct?

- A) A is invertible
- B) Determinant of A is 0
- C) The sum of the diagonal elements of A is 1
- D) A and A^2 have the same rank

Q12. Let A be the adjacency matrix of a simple undirected graph G . Suppose A is its own inverse ($A = A^{-1}$). Which one of the following statements is always TRUE?

- A) G is a cycle
- B) G is a perfect matching
- C) G is a complete graph
- D) There is no such graph G

Unit 4: Probability, Statistics & Data Interpretation (Weightage: 20%)

Q13. Suppose X and Y are random variables. The conditional expectation of X given Y is denoted by $E[X | Y]$. Then $E[E[X | Y]]$ equals

- A) $E[X | Y]$
- B) $E[X]/E[Y]$
- C) $E[X]$
- D) $E[Y]$

Q14. Let X be a continuous random variable whose CDF, for some t , is: $F(x) = 0$ for $x \leq t$; $F(x) = \frac{x-t}{4-t}$ for $t \leq x \leq 4$; $F(x) = 1$ for $x \geq 4$. If the median of X is 3, what is the value of t ?

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

Q15. Let $X = aZ + b$, where Z is a standard normal random variable and a, b are unknown constants. It is given that $E[X] = 1$, $E[(X - E[X])Z] = -2$, and $E[(X - E[X])^2] = 4$. The values of a, b are:

- A) $a = -2, b = 1$
- B) $a = 2, b = -1$
- C) $a = -2, b = -1$
- D) $a = 1, b = 1$

Q16. Consider the CDF of a random variable X : $F(x) = 0$ for $x \leq -1$; $F(x) = \frac{(x+1)^2}{4}$ for $-1 \leq x \leq 1$; $F(x) = 1$ for $x \geq 1$. The value of $P(X^2 \leq 0.25)$ is

- A) 0.625
- B) 0.25
- C) 0.5
- D) 0.5625

Q17. Let X be an exponentially distributed random variable. If $P(X > 5) = 0.35$, then the conditional probability $P(X > 10 \mid X > 5)$ is

- A) 0.35
- B) 0.1225
- C) 0.70
- D) 0.175

Q18. A bag contains 10 red balls and 15 blue balls. Two balls are drawn randomly without replacement. Given that the first ball drawn is red, the probability that both balls drawn are red is

- A) 0.375
- B) 0.400
- C) 0.417
- D) 0.360

Unit 5: Programming, Algorithms & Data Management (Weightage: 20%)

Q19. Given an integer array of size N , we want to check if the array is sorted (in either ascending or descending order). An algorithm solves this problem by making a single pass through the array and comparing each element only with its adjacent elements. The worst-case time complexity of this algorithm is

- A) both $O(N)$ and $\Omega(N)$
- B) $O(N)$ but not $\Omega(N)$
- C) $\Omega(N)$ but not $O(N)$
- D) neither $O(N)$ nor $\Omega(N)$

Q20. Consider the following C program:

```
#include <stdio.h>
int main(){
    int a = 6;
    int b = 0;
    while(a < 10) {
        a = a / 12 + 1;
        a += b;
    }
    printf("%d", a);
}
```

```
    return 0;
}
```

Which one of the following statements is CORRECT?

- A) The program prints 9 as output
- B) The program prints 10 as output
- C) The program gets stuck in an infinite loop
- D) The program prints 6 as output

Q21. In a B+ tree, the requirement of at least half-full (50%) node occupancy is relaxed for which one of the following cases?

- A) Only the root node
- B) All leaf nodes
- C) All internal nodes
- D) Only the leftmost leaf node

Q22. Consider the following C program. Assume parameters to a function are evaluated from right to left.

```
int g(int p) { printf("%d", p); return p; }
int h(int q) { printf("%d", q); return q; }
void f(int x, int y) { g(x); h(y); }
int main() { f(g(10), h(20)); }
```

Which one of the following options is the CORRECT output of the above C program?

- A) 20101020
- B) 10202010
- C) 20102010
- D) 10201020

Q23. Once the DBMS informs the user that a transaction has been successfully completed, its effect should persist even if the system crashes before all its changes are reflected on disk. This property is called

- A) durability
- B) atomicity
- C) consistency
- D) isolation

Q24. Consider the following disk queue with requests for I/O to block on cylinders: 45,81,185,33,175,99,150,77 (the disk head starts from location 58). What will be the total head movement of cylinders for the SSTF (Shortest Seek Time First) disk scheduling technique?

- A) 160
- B) 124

- C) 177
- D) 240

Unit 6: AI & Machine Learning Foundations (Weightage: 20%)

Q25. A linear classifier $y = \text{sign}(f(x; w, b))$, $f(x; w, b) = w^T x + b$, is trained on a labelled dataset with $y_i \in \{+1, -1\}$. Whenever the classifier wrongly predicts a sample (x_n, y_n) , the parameters are updated by the mistake-driven rule: $w_{\text{new}} = w_{\text{old}} + y_n x_n$, $b_{\text{new}} = b_{\text{old}} + y_n$. Consider a mistake on $(x_n, +1)$, i.e., $f(x_n; w_{\text{old}}, b_{\text{old}}) < 0$. Then

- A) $f(x_n; w_{\text{new}}, b_{\text{new}}) > f(x_n; w_{\text{old}}, b_{\text{old}})$
- B) $f(x_n; w_{\text{new}}, b_{\text{new}}) < f(x_n; w_{\text{old}}, b_{\text{old}})$
- C) $f(x_n; w_{\text{new}}, b_{\text{new}}) = f(x_n; w_{\text{old}}, b_{\text{old}})$
- D) $y_n f(x_n; w_{\text{old}}, b_{\text{old}}) > 1$

Q26. For a classification problem, Principal Component Analysis (PCA) has been used to reduce the dimensionality of a feature space from 100 to 10. Which of the following options is true about the angle θ between the first and the tenth principal components?

- A) $\theta = 0^\circ$
- B) $\theta = 90^\circ$
- C) $90^\circ < \theta \leq 180^\circ$
- D) $0^\circ < \theta < 90^\circ$

Q27. Consider training a classifier for a 10-class classification problem where each input is a 512-dimensional vector. There are 1000 samples, of which the first 100 are held out for testing. Leave-One-Out Cross-Validation (LOOCV) is used for model selection before testing. What is the correct number of validation splits generated?

- A) 10
- B) 512
- C) 900
- D) 1000

Q28. Which of the following algorithms is NOT an example of uninformed search?

- A) Breadth First Search
- B) Depth First Search
- C) A* Search
- D) Depth-limited Search

Q29. Which of the following statements is true for Ridge Regression?

- A) The regularizer in the objective function of Ridge Regression is used to guard against scenarios where the model works well for the test data, but poorly for the training data.
- B) The regularizer of Ridge Regression uses the L_1 norm.

- C) Ridge Regression aims to reduce the number of parameters that have negative values.
- D) The regularizer of Ridge Regression may increase the bias of the model, but it helps in reducing the variance in predictions.

Q30. 20 stories by Author X and 10 stories by Author Y are mixed together (without author names) and a classifier is asked to predict the author of each story. Of X's 20 stories, 6 are classified as Y. Of Y's 10 stories, 2 are classified as X. Treating X and Y as the two classes, what is the overall accuracy of the classifier?

- A) $11/15$
- B) $14/15$
- C) $7/15$
- D) $4/5$