

[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 1785

L

Unique Paper Code : 2342011201

Name of the Paper : Object-Oriented Programming with C++

Name of the Course : B.Sc. (H) Computer Science

Semester : II

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** (Question 1) is compulsory.
3. Attempt **any 4** questions from **Section B** (Questions 2 to 6).

SECTION A

1

2

- (a) Given: `int m = 1, n = 2, i = 1, j = 5;`

`i--&&(4* ++ m<=n)`

Evaluate the given expression and determine the final value of each variable after its execution.

- (b) Write any two ways to code constants in C++? Explain using examples.

2

- (c) (i) Give the output of the following code fragments:

2

```
int n = 5;
if (n == 0)
{
    cout << "n is zero";
}
else
{
    cout << "n is not zero";
}
cout << "The square of n is " << n * n;
```

- (ii) `int n, k = 5;`
`n = (100 % k ? k + 1 : k - 1);`
`cout << "n=" << n << " k=" << k;`

2

- (iii) `string x = "FROM:abcd@gmail.com";`
`int colonPos = x.find(':');`
`string prefix = x.substr(0, colonPos);`
`string suffix = x.substr(colonPos+1);`
`cout << "This message is from" << suffix << prefix;`

3

- (iv) `int list[] = {2,4,-5,6,7,0,-1,6};`
`bool pos = true;`
`int i = 0;`
`while (pos && list[i] != 0)`
`{`
 `pos = (list[i++] > 0);`
`}`
`cout << "The final output is:" << pos << " for i=" << i;`

3

(v) `#include <iostream>`
`using namespace std;`
`class A`
`{`
`public:`
`A()`
`{`
`cout<<"Constructor called\n";`
`}`
`~A()`
`{`
`cout<<"Destructor called\n";`
`}`
`};`
`int main(int argc, char const *argv[])`
`{`
`A *a = new A[5];`
`delete[] a;`
`return 0;`
`}`

(d) Write a function that replaces all vowels in a character array with an asterisk (*). 4

(e) Write the output of the following program : 4

```
#include <iostream>
using namespace std;
int main ()
{
    int x, y;
    x = 5;
    y = ++x * ++x;
    cout << x << " " << y << endl;
    x = 5;
    y = x++ * ++x;
    cout << x << " " << y << endl;
    return 0;
}
```

(f) Explain with an example, under what circumstances the following statements are used: 4

- (i) `throw;`
- (ii) `catch(...)`

SECTION B

2 (a) What are static members of a class? How are they accessed? Explain with an example. 4

(b) b) Which of the following are valid identifiers and which are invalid identifiers? 5
 Give reasons:

- i) `@count`
- ii) `_count`
- iii) `3count`
- iv) `#marks`
- v) `english_marks`

- (c) Consider the following class: 6
- ```
Class Circle
{
float radius;
};
```
- Add the following member function definitions to the above class:
- A default constructor
  - A parameterized constructor
  - A function that computes the area
- Write the main() function to show the use of the above constructors and the function.
- 3 (a) Write a function that prints the following pattern for a given integer n=3: 4
- ```
1
22
333
```
- (b) Write a C++ program that uses a for loop to compute the sum of the first n natural numbers. 5
- (c) Define a class Person with name as its data member. Inherit two classes, Student and Employee, from Person. The Student class should include additional attributes such as course, marks, and year, while the Employee class should include department and salary. Implement a display() method in all three classes to show their respective data members. Also, include the necessary mechanisms to demonstrate runtime polymorphism, such as using a base class pointer to invoke the overridden display() methods of the derived classes. 6
- 4 (a) What are inline functions? Rewrite the following code by changing it to an inline function: 4
- ```
#include <iostream>
using namespace std;
int square(int a);

int main()
{
 int c;
 c = square(4);
 return 0;
}

int square(int x)
{
 return x * x;
}
```
- (b) Rewrite the following for statement as an equivalent while statement: 5

```

for(i=0; i<max_length;i++)
{
 if(input_line[i]=='?')
 {
 quest_count++;
 }
}

```

- (c) Differentiate call-by-value and call-by-reference with an example. Write a program to swap two numbers using templates. 6
- 5 (a) Give the output of the following code snippet: 4
- ```

int x = 5;
int *p = &x;
int **q = &p;
cout << *p << " " << **q;

```
- (b) Find the error and correct the code: 5
- int m==40;
 - if (x = 10) cout << "Yes";
 - float arr[5] = {1, 2, 3, 4, 5, 6};
 - int *p = &x; delete p;
 - for (int i=0; i<10; i--);
- (c) Write a program in C++ to copy the contents of one text file to another file after removing all whitespaces. 6
- 6 (a) What is the difference between built-in data types and derived data types? Show with an example of each. 4
- (b) Create a class `Triangle` to represent a triangle with appropriate data members for its sides. Incorporate exception handling to ensure that all sides are greater than zero and that the sum of any two sides is greater than the third side; if these conditions are not satisfied, an appropriate exception should be raised. 5
- (c) Differentiate between the following terms : 6
- Data Abstraction and Data Encapsulation
 - Constructor and Destructor
 - ifstream and ofstream

[This question paper contains 8 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 1786

L

Unique Paper Code : 2342011204

Name of the Paper : Programming Using C++ (Adm. in 2024)

Name of the Course : B.Sc. (H) Computer Science

Semester : II

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section-A** is compulsory.
3. Attempt any **four** Questions from **Section B**.
3. Parts of question must be answered together.

Section A

Q1. a) What is a void pointer in C++? Explain with the help of a suitable code. (2)

b) Find the output of the following code: (3)

```
#include <iostream>
using namespace std;
int main() {
    int arr[3] = {8, 20, 30};
    int *p = arr;
    cout<<*p <<endl;
    cout<<*p+1 << "\t" << *(p+2) << "\t" <<endl;
    return 0;
}
```

c) What is an inline function? Write a C++ program using an inline (3)
function to calculate the cube of a number.

d) Write a C++ function to swap the values of two variables using call by (3)
reference.

e) Find the output of the following code(s): (3*1)

i.

```
#include <iostream>
using namespace std;
int main() {
    int a = 5;
    cout<< (a > 3 ? a + 2 : a - 2);
    return 0;
}
```

ii.

```
#include <iostream>
using namespace std;
int main() {
```

```
    char ch = 'A';
```

कालिन्दी महाविद्यालय पुस्तकालय
KALINDI COLLEGE LIBRARY

P.T.O.

```

        cout<<ch + 2;
        return 0;
    }

    iii.
#include <iostream>
using namespace std;
int main() {
    int x = 2;
    cout<< x * x++ << " " << ++x;
    return 0;
}

```

f) State whether true or false: (3*1)

- i. Destructors can be declared as private members of the class.
- ii. Constructors can have default arguments.
- iii. Constructors can be used to return values from the function.

g) Find the output of the following code: (3)

```

#include <iostream>
using namespace std;
int func(int x, int y = 2, int z = 3) {
    return x * y + z;
}
int main() {
    cout<<func(4) <<endl;
    cout<<func(4, 5) <<endl;
    cout<<func(4, 5, 6) <<endl;
    return 0;
}

```

h) Write a C++ function to find the length of a string without using any (3)
inbuilt string functions.

i) How does the visibility modes of inheritance differ in the following (3*1)
derived classes:

- i. class D1: private B
- ii. class D2: protected B
- iii. class D3: public B

j) Find the output of the following code: (2*2)

```

i.
#include <iostream>
using namespace std;
int main() {
    int x = 1, y = 2;

```

```

switch(x) {
    case 1:
        switch(y) {
            case 1: cout<< "Inner1"; break;
            case 2: cout<< "Inner2"; break;
        }
        break;
    default: cout<< "Outer";
}
return 0;}

```

ii.

```

#include <iostream>
using namespace std;
int main() {
    int i, j = 0;

    for(i = 1; i<= 3; i++) {
        j = i;
        while(j > 0) {
            cout<< j << " ";
            j--;
        }
    }
    return 0;
}

```

Section B

Q2. a) What is polymorphism? Explain early binding and late binding in the context of polymorphism. (5)

b) Find the output of the following code: (5)

```

#include <iostream>
using namespace std;
int x = 100;
class Test {
public:
    int x;
    Test() { x = 50; }
    void show() {
        int x = 10;
        cout<< x << " ";
    }
}

```



```

cout<< this->x << " ";
cout<< ::x << " ";
cout<< x + this->x + ::x <<endl;
    }
};

int main() {
    Test t;
    t.show();
    return 0;
}

```

- c) Create a base class Shape with one data member radius and a virtual function area(). Derive a class Circle using public inheritance and override the function area() to calculate the area of the circle. Also, create another derived class Square with data member side and override the function area(). Demonstrate how you can call the area functions of both derived classes using a single pointer to the base class. (5)

Q3. a) Find the output of the following code: (5)

```

#include <iostream>
using namespace std;
int main() {
    int a[] = {1, 2, 3, 4};
    int *p = a;
    cout<< *p++ << " ";
    cout<< ++*p << " ";
    cout<< *p;
    return 0;
}

```

- b) Write a C++ program to accept two integers a and b from the user and perform the division a / b. Before performing the division, appropriate conditions should be checked such that if the denominator b is zero, an integer exception is thrown, and if the numerator a is negative, a character exception is thrown. Use multiple catch blocks to handle each type of exception separately and display suitable error messages for each case. (5)
- c) Write a C++ program to initialize an integer array of size 10. Find the smallest number and print its position. (5)

Q4. a) Write a C++ program to implement following specifications: (10)

- i. Class Distance1
 - a. Private data member: meters.
 - b. Single parameter Constructor.
 - c. Public member function, init1() to initialize the value.

- ii. Declare a function void addDistance(Distance1) in the class that:
 - a. Accepts accept object of Distance1.
 - b. Adds their distances.
 - c. Displays the total distance.
- iii. In the main() function:
 - a. Create objects of class
 - b. Initialize them with values
 - c. Call the addDistance function to display the result.

- b) Write a function DiagonalMatrix() that accepts a square matrix A (5) and its order n. The function should convert matrix A into a diagonal matrix by assigning 0 to all non-diagonal elements.

- Q5. a) Write a program in C++ to print the following pattern. Take number of (5) rows as input from the user.

```
A
B C
D E F
G H I J
```

- b) Find the output of the following code: (5)

```
#include <iostream>
using namespace std;
class Base {
public:
    virtual void show() {
        cout<< "Base class show\n";
    }
};
class Derived1 : public Base {
public:
    void show() {
        cout<< "Derived1 class show\n";
    }
};
class Derived2 : public Base {
public:
    void show() {
        cout<< "Derived2 class show\n";
    }
};
int main() {
    Base *b;
    Derived1 d1;
```

```

    Derived2 d2;
    Base baseObj;
    b = &baseObj;
    b->show();
    b = &d1;
    b->show();
    b = &d2;
    b->show();
    b = &d1;
    b->show();
    b = &baseObj;
    b->show();
    return 0;
}

```

c) Write the function prototypes (declarations only) for the following (5) situations in C++:

- i. A function add that takes two integers and returns their sum.
- ii. A function display that takes a constant character pointer (string) and does not return any value.
- iii. A function power that takes two integers x and y, where y has a default value of 2, and returns integer.
- iv. A function swap that swaps two integers using call by reference.
- v. A function area that takes a float radius and returns the area of a circle, with a default value of radius as 1.0.

Q6. a) Write a C++ program to count the number of vowels in a file, data.txt. (5)

b) Find the output of the following code: (2*2.5)

```

i.
#include <iostream>
using namespace std;

template <class T>
class Sample {
    T x;
public:
    Sample(T a) {
        x = a;
    }
    void show() {
        cout << x * 2 << " ";
    }
};

```

```
int main() {  
    Sample<int> s1(5);  
    Sample<char> s2('A');  
    s1.show();  
    s2.show();  
    return 0;  
}
```

ii.

```
#include <iostream>  
using namespace std;  
int main() {  
    try {  
        int a = 5, b = 0;  
        if(b == 0)  
            throw b;  
        cout<< a / b;  
    }  
    catch(char e) {  
        cout<< "Character exception\n";  
    }  
    catch(int e) {  
        cout<< "Integer exception\n";  
    }  
    catch(...) {  
        cout<< "General exception\n";  
    }  
  
    try {  
        throw 'x';  
    }  
    catch(char e) {  
        cout<< "Character exception\n";  
    }  
    catch(int e) {  
        cout<< "Integer exception\n";  
    }  
    catch(...) {  
        cout<< "General exception\n";  
    }  
  
    try {  
        throw 3.14;  
    }  
    catch(char e) {  
        cout<< "Character exception\n";  
    }  
    catch(int e) {
```



```

        cout<< "Integer exception\n";
    }
    catch(...) {
        cout<< "General exception\n";
    }
    return 0;
}

```

- c) Write a C++ program to demonstrate constructor overloading using a class Student with the following specifications: (5)

- i. Data Members: Name, Roll Number, Marks
- ii. Implement the following three types of constructors:
 - a. Default Constructor: Initialize data members with default values
 - b. Parameterized Constructor: Initialize data members with user-defined values.
 - c. Copy Constructor: Initialize an object by copying values from another object
- iii. In the main() function, create objects using all three types of constructors.

- Q7. a) Write a C++ program for a library management system with the following specifications: (10)

- i. Class: LibraryMember
 - a. Private Data Members: Name of the member, Age
 - b. Public Data Members: Membership ID, Type of membership (Student/Faculty), Number of books issued
- ii. Member Functions (Inside the Class):
 - a. A function to assign initial values
 - b. A function to issue a book (increase book count)
- iii. Member Functions (Outside the Class):
 - a. A function to return a book after verifying that at least one book is issued
 - b. A function to display the member's name and number of books currently issued.
- iv. Create an object of the class and demonstrate the working of the member functions.

- b) Write a program to create a class Car with static data member, number. Display the number of objects created and destroyed. (5)

[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 2051

L

Unique Paper Code : 2342011202

Name of the Paper : Discrete Mathematical Structures

Name of the Course : B.Sc. (Hons.) Computer Science (NEP-UGCF-2022)

Semester : II

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Question No. 1 is compulsory.
3. Attempt any 4 of Questions Nos. 2 to 7.
4. Parts of a question must be answered together.
5. Use of a simple calculator is allowed.

SECTION A

1. (a) Let R be the relation on positive integer set Z^+ defined as:

5

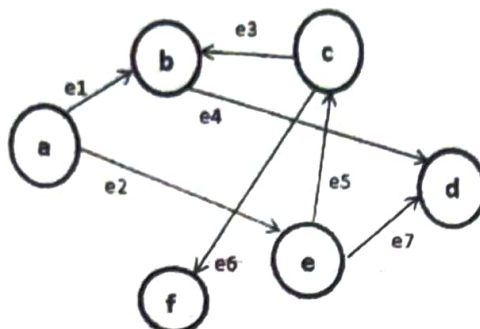
$$R = \{(x, y) \mid 2x + 6y = 24\}$$

- (i) Define R as a set of ordered pairs.
 - (ii) Is R a partial order relation? Justify.
- (b) Is $(p \rightarrow q) \rightarrow [(p \rightarrow q) \rightarrow q]$ a tautology? Justify your answer with the help of a truth table.

(c) For the directed graph given below:

5

- (i) Determine the outdegree of every vertex.
- (ii) Give the incidence matrix representation.



(d) Prove by contraposition that if n is an integer and $3n + 2$ is odd, then n is odd.

5

(e) A drawer contains a dozen black socks and a dozen white socks, all unmatched. A man takes socks out at random in the dark.

5

- (i) How many socks must he take out to be sure that he has at least two socks of the same color?
- (ii) How many socks must he take out to be sure that he has two black socks?

(f) Let a be a numeric function such that

5

$$a_r = \begin{cases} 3^r & 0 \leq r \leq 3 \\ 3^{-r} & r \geq 4 \end{cases}$$

Determine $S^2 a$ and $S^{-2} a$.

Section B

2. (a) Let $A = \{1, 2, 3, 4, 6, 8, 12, 24\}$ and let R be the divisibility relation on A .

7

- (i) Show that (A, R) is a POSET.
- (ii) Draw the Hasse diagram.
- (iii) Find the maximal and minimal elements.

(b) Show that the following argument is valid:

8

"If Priya attends the meeting, then she presents the report. If she presents the report, then the committee will be satisfied. If the committee is satisfied, then the project will be approved. Therefore, if Priya attends the meeting, the project will be approved."

3. (a) Using mathematical induction, prove that for every non-negative integer n : 7

$$1 + 3 + 3^2 + \dots + 3^n = (3^{n+1} - 1) / 2$$

- (b) Let $f : X \rightarrow Y$ where $X = \{1, 2, 3, 4, 5\}$ and $Y = \{a, b, c, d, e\}$ be defined by: 8

$$f(1) = b, f(2) = d, f(3) = a, f(4) = e, f(5) = c$$

- (i) Show that f is a bijection.
- (ii) Find f^{-1} and write it as a set of ordered pairs.
- (iii) Determine the image of the subset $S = \{1, 3, 5\}$ under f .

4. (a) (i) Define planar graphs. Is Q_3 planar? Justify. 7

- (ii) Suppose that a connected planar graph has four vertices of degree 3, two vertices of degree 4, and five vertices of degree 2. Into how many regions does this representation divide the plane?

- (b) In a college, 300 students are enrolled. Out of these, 90 have enrolled in Java, 150 in Python and 70 in C++. Further, 40 have enrolled in both Java and Python, 30 in both Python and C++, and 25 in both Java and C++. There are 230 students that have enrolled in at least one of these three subjects. 8

- (i) How many students have enrolled in a course in all three subjects?
- (ii) How many students have not enrolled in any of the three subjects?
- (iii) How many students opted for Python but not for Java?

5. (a) In how many ways can 20 different books be distributed among seven students so that: 7

- (i) Five students would have 2 books each and the other two students would have 5 books each.
- (ii) Three particular students would get exactly 1 book each.
- (iii) One particular student does not get any book.

- (b) Evaluate $3^{29} \text{ mod } 191$ using the Fast Modular Exponentiation algorithm. 8

6. (a) Prove that a tree with n vertices has exactly $n - 1$ edges. 7

- (b) (i) Show that $\bar{p}$ is tautologically implied by $(\bar{p} \wedge \bar{q}), \bar{q} \vee r, \bar{r}$ using equivalence rules. 8

- (ii) State the converse, contrapositive, and inverse of the following statement:
"If it snows today, I will ski tomorrow"

7. (a) Given the numeric functions a_r and b_r ,

7

$$a_r = \begin{cases} 5 & 0 \leq r \leq 1 \\ 0 & r \geq 2 \end{cases}$$

$$b_r = \begin{cases} 2 & 0 \leq r \leq 4 \\ 0 & r \geq 5 \end{cases}$$

Determine the closed form of:

- (i) $a + b$
 (ii) $a * b$
- (b) Consider the graph with $V = \{a, b, c, d, e, f, g\}$ and edge set
 $E = \{ab, bc, cd, de, ef, fa, ag, bg, cg, dg, eg, fg\}$. 8
- (i) Check whether the graph has a Euler Circuit. Give such a circuit, if it exists.
 (ii) Check whether the graph has a Hamiltonian Circuit. Give such a circuit, if it exists.
 (iii) Identify all the cut vertices and cut edges, if any.

[This question paper contains 4 printed pages.]

Your Roll No.....

L

Sr. No. of Question Paper : 2070

Unique Paper Code : 2342011203

Name of the Paper : Probability for Computing

Name of the Course : B.Sc. (H) Computer Science (NEP)

Semester : II / DSC

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **First** question is compulsory and attempt any **four** questions from remaining.
3. Parts of the questions to be attempted together.
4. Attempt all questions from **Section A**.
5. Attempt **any 4** questions from **Section B**.
6. Use of non-programmable scientific calculator is allowed.

Section -A

1. (a) Prove that $Var(\sum_{i=1}^n X_i) = \sum_{i=1}^n Var(X_i)$ 4

- (b) Given the cumulative distribution function: 4

$$F(x) = \begin{cases} 0, & \text{if } x < -1, \\ (x + 1)/2, & \text{if } -1 \leq x < 1, \\ 1, & \text{if } 1 \leq x. \end{cases}$$

Compute:

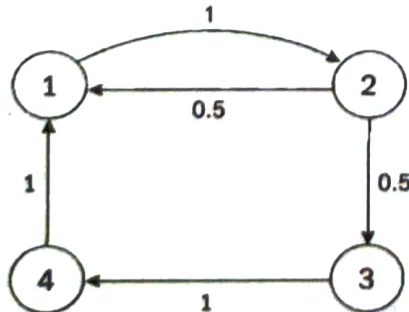
i) $p(-\frac{1}{2} < X \leq \frac{1}{2})$

ii) $p(X \geq 3/4)$

- (c) Suppose that X is a random variable with mean 10 and variance 15. 4

- i) Calculate the bound on probability $P\{5 \leq X \leq 15\}$ using Chebyshev's inequality?
- ii) What can we say about $P\{X \geq 20\}$ using Markov's inequality?

- (d) What is meant by the periodicity of a class in Markov chain. Identify the periodicity of states in the following Markov chain: 4



- (e) Suppose that a trainee soldier shoots a target in an independent fashion. If the probability that the target is hit in any shot is 0.7. What is the probability that it takes him fewer than four shots to hit the target? 4

- (f) The joint probability mass function of X and Y , $p(x,y)$, is given by: 5

$$\begin{array}{lll} p(1, 1) = \frac{1}{9} & p(2, 1) = \frac{1}{3} & p(3, 1) = \frac{1}{9} \\ p(1, 2) = \frac{1}{9} & p(2, 2) = 0 & p(3, 2) = \frac{1}{18} \\ p(1, 3) = 0 & p(2, 3) = \frac{1}{6} & p(3, 3) = \frac{1}{9} \end{array}$$

Compute $E[X|Y = i]$ for $i = 1, 2, 3$.

- (g) Assume the experiment of throwing two dice and answer the following: 5
- Give the number of outcomes in the sample space for the experiment.
 - Give the favorable outcomes for the event that the number on the first die is greater than 5.
 - Find the probability of getting a sum of the numbers on two dice greater than or equal to 10, given that the number on the first die is 4

Section- B

2. (a) Derive the expression for the expected value of a geometric distribution. 5
- (b) Let $\{X_n\}_{n \geq 0}$ be a discrete-time Markov chain on the state space $\{1, 2, 3\}$ with one-step transition probability matrix: 5

$$\begin{vmatrix} 0.4 & 0.3 & 0.3 \\ 0.5 & 0.2 & 0.3 \\ 0.2 & 0.4 & 0.4 \end{vmatrix}$$

And the initial distribution:

$$P(X_0 = 1) = 0.5, P(X_0 = 2) = 0.2, P(X_0 = 3) = 0.3.$$

Then calculate the value of $P(X_1 = 2, X_2 = 3, X_3 = 1)$ (rounded off to three decimal places).

- (c) Differentiate between mutually independent and pairwise independent events. 5
- Let the equally likely outcomes of an experiment be one of the four points in the three-dimensional space with Cartesian coordinates $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, and $(1, 1, 1)$. Let us define the events A , B and C as:

A: X-coordinate value is one.

B: Y-coordinate value is one.

C: Z-coordinate value is one.

Verify if these events are mutually-independent or not?

3. (a) A random sample of size 100 is taken from a population whose mean is 60 & variance is 400. Using central limit theorem, with what probability can we assert that the mean of the sample will not differ from mean by more than 4? Given that $\phi_z(2) = 0.9772$ 7

- (b) For a continuous random variable, write the expression for condition probability density function of Random variable Y given that X is defined. 8
 Let $f(x, y) = x + y$ for $0 \leq x \leq 1$ and $0 \leq y \leq 1$. Find $f_{Y/X}(y/x)$.
4. (a) Give the probability mass function for the Poisson distribution. Let X & Y be independent Poisson random variables with respective means λ_1 & λ_2 . Calculate the distribution of X+Y. 7
- (b) Consider a Markov chain with state space $S = \{1, 2, 3, 4, 5\}$ 8
 The transition probability matrix is given by:
- $$P = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 & 0 \\ 0.4 & 0.6 & 0 & 0 & 0 \\ 0 & 0 & 0.3 & 0.3 & 0.4 \\ 0 & 0 & 0.2 & 0.8 & 0 \\ 0 & 0 & 0 & 0.6 & 0.5 \end{bmatrix}$$
- Construct the corresponding state transition diagram for the above Markov chain.
 - Identify the communicating classes in the diagram
 - Classify each state as: Transient or Recurrent
 - Determine whether the chain is Ergodic or not?
- Q5 (a) The amount of time a person must wait for a train to arrive in a certain town is uniformly distributed between 0 and 40 minutes.
- Determine the probability density function $f(x)$ for the distribution. 1
 - Determine the cumulative distribution function $F(x)$. 1
 - Calculate the mean. 1
 - What is the probability that a person must wait more than 30 minutes? 2
 - Calculate $P(10 < X < 26)$. 2
- (b) i. Give three differences between covariance and correlation. 4 + 4
 ii. Prove that $Cov(\bar{X}, X_i - \bar{X}) = 0, i=1, 2, 3 \dots n$
- Q6. (a) Mani wants to watch either a suspense-thriller movie or comedy movie. If the number of advertisements in a suspense-thriller movie is Poisson distributed with mean 2 and if the number of advertisements in a comedy movie is Poisson distributed with mean 5, then assuming Mani is equally likely to watch either movie, what is the expected number of advertisements that Mani will come across? 7

- (b) A subscription service is analyzing how users switch between two plans: basic plan and premium plan, which can be considered as Markov chain with each plan representing a state. From historical data, 30% of Basic users upgrade to Premium each month with probability = 0.3, while 20% of Premium users downgrade to Basic each month with probability = 0.2. From the above description: 8
- Form the transition matrix for this system.
 - Find the **steady-state probabilities** of users in each plan.
 - Interpret the result in business terms.
- Q7. (a) An NGO receives donations from n independent donors. Let the amount donated by the i^{th} donor be represented by a random variable X_i , for $i = 1, 2, \dots, n$. Assume that all X_i are independent and identically distributed. If $n = 100$ and the donation amount from each donor follows a uniform distribution between ₹100 and ₹500, compute the expected total donation. 4
- (b) A streaming platform tracks user behavior across three states on any given day as an active user (watches content), Idle user (logs in but doesn't watch) and Churned user (does not return). Based on historical data, the active user may turn into an idle user with probability 0.3 and may churn with probability 0.1, An idle user may convert to an active state or remain idle with probabilities 0.2 and 0.3 respectively. A churned user will never convert to either active or idle state. Convert above into a Markov chain with three states and construct its equivalent transition probability matrix. 4
- (c) What are random numbers, and why is the ability to generate random numbers considered to be important? Generates a sequence of five random numbers using the linear congruential generation algorithm with $a = 3$, $m = 7$, and initial seed $X_0 = 2$. 7

[This question paper contains 8 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 231

L

Unique Paper Code : 2343010035

Name of the Paper : Data Analysis and Visualization using Python

Name of the Course : B.Sc. (H) CS (DSE)

Semester : IV

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Section A is compulsory.
3. Attempt any four questions from Section B.
4. Parts of a question must be answered together.

Assume numpy has been imported as np and pandas has been imported as pd.

**Section- A
(Compulsory)**

1. a) Give the output that will be produced by the following given code: (3)

```
s1 = pd.Series(['Red', 'Blue', 'Green'],
               index=[0,3,6])
s2 = s1.reindex(range(8), method= 'ffill')
print(s2)
```

- b) Write the Python command to create a 3-dimensional NumPy array a1 of shape (4,2,3) filled with random integers between 1 and 100. (3)

- c) Give the output that will be produced by the following Python code: (3)

```
arr1 = np.array([[4, 1, 7],
                 [3, 2, 9],
                 [2, 3, 11]])
arr2 = np.swapaxes(arr1, 0, 1)
print(arr2)
```


- d) Consider a DataFrame df given below:

```
data = pd.Series(np.random.uniform(size=9),
index=[['a', 'a', 'a', 'b', 'b', 'c', 'c', 'd',
'd'], [1, 2, 3, 1, 3, 1, 2, 2, 3]])
```

Write the Python command to retrieve following data from df:

- i. All rows with index 'b' and 'c'.
 - ii. All rows with second index level as '2'.
- e) Define unimodal and bimodal distribution. Find the mode(s) of the data given below:

9, 3, 1, 5, 6, 9, 5, 9, 3, 5, 7, 9, 6, 5

- f) Identify and correct error(s) in the Python code given below: (3)

```
data = {'Name': ['Joe', 'Bob', 'John', 'Robin',
'Oscar'], 'Age': [22, 18, 21, 19, 20]}
f1 = pd.dataframe(data, column = ['Name', 'Age',
'Stream', 'Percentage'])
print(f1.iloc["Name"])
```

- g) Consider a Matplotlib plot. Write the function used for each of the following tasks: (3)

- i. To add title 'Graph' to the figure.
- ii. To label x-axis as 'Index' and y-axis as 'Values'.

- h) What is a pivot table? Give a suitable example. (3)

- i) Write the Python code to read a file named "file1.csv" into a DataFrame f1. (3)
The file "file1.csv" contains 4 columns and has no header row. Assign column labels as ["a", "b", "c", "d"] while reading the file into f1.

j) Give the command for the following operations on a DataFrame df: (3)

- i. To drop NaN values .
- ii. To replace NaN values with 0 .
- iii. To check for NaN values.

Section B

2 a) Write NumPy commands to perform the following operations on array arr1: (5)

- i. Create a 1D array arr1 of size 20 containing values from 1 to 20.
- ii. To modify elements >10 to their respective negative values.
- iii. To retrieve elements at index 5, 9, 13, 18.
- iv. To reshape it to array of size 4 x 5 and store it in arr2.
- v. To find transpose of arr2 created in part iv.

b) Explain different types of sampling techniques. (5)

c) Consider a Pandas series `s1 = pd.Series([1,4,9,16,25])`. Write the Python code to plot a line graph of s1. Set the x limit to [0,5] and y limit to [0,25]. Set the style of line graph to 'dash' pattern and marker to 'circle' shape. (5)

3 a) Explain the concept of data wrangling. Also, mention any three common issues encountered during the data wrangling process (5)

b) Consider two DataFrames f1 and f2 given below: (5)

```
f1 = pd.DataFrame({'key': ['a', 'b', 'c', 'a', 'b', 'b'], 'data1': range(6)})  
f2 = pd.DataFrame({'key': ['a', 'b', 'd'], 'data2': range(3)})
```

Give the output of the following merge functions, invoked on DataFrames f1 and

f2 :

- i. `pd.merge(f1, f2, how='inner')`
- ii. `pd.merge(f1, f2, how='outer')`

c) Write NumPy commands to create the following arrays: (5)

- i. 2D array 'a1' of shape (2, 3) with ones on the main diagonal and zeros elsewhere.
- ii. Array 'a2' filled with 15s with the same shape and type as the array 'a1'.
- iii. Array 'a3' of shape (5, 4) filled with value 10.
- iv. 2D array 'a4' which is the transpose of array 'a3'.
- v. 1D array 'a5' of ten random floating-point numbers in the range 0 to 1.

4 a) Explain the split-apply-combine strategy used in group operations in Pandas. (5)
Illustrate, with an example, how the `agg()` function can be used to apply custom aggregation functions.

b) Give the output that will be produced by following Python code segment? (5)

```
a1 = np.array([[6, 2, 8], [4, 1, 5], [7, 1, 9]])
print(a1[0, 2])
print(a1[:2])
print(a1 + 2)
a1[:2, :1] = 0
print(a1)
print(a1.dtype)
```

c) Differentiate between bar chart and histogram. (5)

Consider the following list:

```
A = [10, 20, 30, 10, 20, 10, 30, 10, 20, 30, 40, 20, 30, 40, 20]
```

Write Python commands to plot a histogram titled 'Histogram of Data A' for the data in A with 5 bins. Import appropriate library.

5 a) Write Python commands to perform the following tasks: (5)

- i. Create a DataFrame `df` with 3 rows and 4 columns containing random integers in the range 1 to 30. The column labels should be: 'One', 'Two', 'Three', 'Four'. The row indices should be: 'A', 'B', 'C'.
- ii. Calculate the row-wise sum of values in `df`.
- iii. Find the maximum value in each column of `df` and its corresponding row index.

b) Consider the following Python code: (5)

```
x = [ 2, 5, 6, 9, 8, 7, 10, 3, 5]
y = [9, 8, 7, 8, 10, 6, 10, 7, 9]
z = [4, 7, 5, 6, 9, 5, 8, 6, 7]
df = pd.DataFrame({'x': x, 'y': y, 'z': z})
```

Import necessary libraries and write Python commands to perform the following tasks:

- i. Create a pair plot using the variables `x`, `y`, and `z`.
 - ii. Generate a heatmap to display the correlation between all three data variables in `df`.
 - iii. Plot a regression plot to show the relationship between `y` and `z`.
- c) Consider a DataFrame `df1` having 10 rows and 4 columns named `ItemID`, `Name`, `Price`, `Size`. Write Python code to: (5)
- i. Find the covariance of `Size` with `Price`.
 - ii. Find mean of `Price`.
 - iii. Display `ItemID` and `Name`.
 - iv. Count number of items with different `Size`.
 - v. Find unique values of `Price`.

6 a) Consider the following data:

(5)

```
marks = [35, 48, 52, 67, 73, 85, 91, 60, 77, 40]
df = pd.DataFrame({'Marks': marks})
```

Write Python commands to perform the following tasks:

i. Divide the marks into three bins:

- 0-50 → 'Low'
- 50-75 → 'Medium'
- 75-100 → 'High'

Ensure that the bins are right-closed left-open intervals and appropriate labels are assigned.

ii. Modify the above code to make the bins left-closed right-open.

iii. Display the frequency of observations in each category created.

b) Consider the following DataFrame:

(5)

```
df = pd.DataFrame({
    'A': [3, np.nan, 5, np.nan],
    'B': [6, np.nan, np.nan, 7],
    'C': [np.nan, np.nan, 8, np.nan]
})
```

Write Python code to:

- i. Delete all rows containing at least two missing values.
- ii. Delete columns containing at least one missing value.
- iii. Delete columns with all missing values.

c) Consider the Series given below:

(5)

```
sdata = {"Ohio": 35000, "Texas": 71000, "Oregon":
         16000, "Utah": 5000}
obj3 = pd.Series(sdata)
```

Give the output of the following:

- i. `print(obj3.index)`
- ii. `print(obj3.values)`
- iii. `print(pd.isna(obj3))`
- iv. `obj3 = obj3.reindex(["Ohio", "Texas", "California", "Utah"])`
`print(obj3)`

7 a) Consider the following DataFrame df:

(5)

```
data = {
    'Dept': ['HR', 'IT', 'HR', 'IT', 'Finance', 'HR',
            'Finance'],
    'Emp' : ['A', 'B', 'C', 'D', 'E', 'F', 'G'],
    'Bonus': [4000, 6000, 4500, 6500, 7000, 4000, 7200],
    'Experience': ['Junior', 'Senior', 'Junior', 'Senior',
                  'Senior', 'Junior', 'Senior']
}
df = pd.DataFrame(data)
```

Write Python commands to perform the following tasks:

- i. Group the data by 'Dept' and calculate the total bonus for each department.
 - ii. Group the data by 'Experience' and calculate the average 'Bonus' for each experience level.
 - iii. Group the data by both 'Dept' and 'Experience' and determine the number of employees in each group.
- b) What is the difference between `stack()` and `unstack()` functions? Give an example of each. (5)

e) Define the following terms:

(5)

- i. Independent and Dependent variables
- ii. Descriptive and Inferential statistics

[This question paper contains 4 printed pages.]

Your Roll No.....

L

Sr. No. of Question Paper : 232

Unique Paper Code : 2343012005

Name of the Paper : Data Mining – I

Name of the Course : B.Sc. (Hons.) Computer Science

Semester : IV

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** (Question No. 1) is compulsory.
3. Attempt any **four** questions from **Section B** (Questions 2 to 7).
4. The use of scientific calculator is allowed.
5. Parts of the question must be answered together.

Section A

- Q1 (a) Data Mining datasets involve non-traditional types of data and data distributions. What are non-traditional types of data in Data Mining? Give any two examples. (2)
- (b) Why is the Naïve Bayes classifier considered an early learner? Briefly justify your answer. (2)
- (c) Identify whether or not the given activities are data mining tasks. Justify your choice. (3)
- i. Extracting all records of students scoring above 90% from a database.
 - ii. Discovering groups of customers with similar buying behavior.
 - iii. Identifying fraudulent credit card transactions from transaction history.
- (d) A researcher applies Leave-One-Out Cross-Validation (LOOCV) on a dataset containing n instances. Determine the number of iterations involved. What is the composition of training and test set in each iteration? (3)
- (e) Is regression considered a supervised learning task or an unsupervised learning task? Justify your answer. (3)
- (f) In Naïve Bayes classifier, if certain combination of attributes and class labels are never observed, it results in a zero conditional probability. What problems may arise in this scenario? Which alternate estimates can be used in such a scenario? (4)
- (g) Consider an attribute Purchased Products with possible values {Laptop, Mobile, Tablet, Headphones, Smartwatch}. Convert this attribute to asymmetric binary attribute. (4)
- (h) Define entropy in the context of decision tree learning. How is it used to measure impurity and guide the selection of the most appropriate attribute for splitting the data? (4)

- (i) A data analyst is using the DBSCAN algorithm to identify clusters in a spatial dataset of customer locations. During the analysis, the analyst needs to determine which points form dense regions and which are considered outliers. (5)

Explain how the parameters *Eps* and *MinPts* are used in this context, and describe how points are classified as core, border, or noise points with suitable reasoning.

Section B

- Q2. (a) A classifier is used to predict whether an email is Spam or Not Spam. The following results were obtained on a test dataset:

Email ID	Actual Class	Predicted Class
E1	Not Spam	Not Spam
E2	Not Spam	Not Spam
E3	Not Spam	Not Spam
E4	Not Spam	Spam
E5	Not Spam	Not Spam
E6	Not Spam	Not Spam
E7	Spam	Spam
E8	Spam	Not Spam
E9	Spam	Not Spam
E10	Spam	Not Spam

- Construct the confusion matrix. (2)
 - Compute accuracy, precision, recall and F1-score. (6)
 - Identify whether the data shows class imbalance. Is accuracy alone sufficient to evaluate the model? Give proper reasoning. (4)
- (b) Differentiate between noise and outliers with one suitable example for each. (3)

- Q3. Consider the following market basket transactions:

TID	List of Items
T1	{Atta, Sugar, Tea}
T2	{Sugar, Biscuits}
T3	{Sugar, Ghee}
T4	{Atta, Sugar, Biscuits}
T5	{Atta, Ghee}
T6	{Sugar, Ghee}
T7	{Atta, Ghee}
T8	{Atta, Sugar, Ghee, Tea}
T9	{Atta, Sugar, Ghee}

- (a) Apply the Apriori algorithm on the given transactional dataset to compute the candidate and frequent itemsets for each dataset scan. Assume a support threshold of 23%. (8)
- (b) Enumerate all association rules generated from the largest frequent itemsets. Compute the confidence of each generated rule. Assuming that the minimum confidence threshold is 50%, find all the strong association rules. (7)

Q4. Consider the following dataset:

Weekend	Weather	Parents	Income	Decision
W1	Sunny	Yes	Rich	Cinema
W2	Sunny	No	Rich	Tennis
W3	Windy	Yes	Rich	Cinema
W4	Rainy	Yes	Poor	Cinema
W5	Rainy	No	Rich	Shopping
W6	Rainy	Yes	Poor	Cinema
W7	Windy	No	Poor	Cinema
W8	Windy	No	Rich	Shopping
W9	Windy	Yes	Rich	Cinema
W10	Sunny	No	Rich	Tennis
W11	Rainy	Yes	Poor	Tennis
W12	Sunny	No	Rich	Shopping

- (a) Compute the Gini Index of Weather, Parents, and Income attributes. Given that you construct a decision tree using the Gini Index as the splitting criteria, which of these three attributes would you choose at the root? Justify your choice. (12)
- (b) Compute the Gini Index of 'Weekend' attribute. Is it a good splitting attribute for constructing a decision tree? Justify your answer. (3)

- Q5. (a) Why is it important to standardize data before performing distance-based computations? (7)

Given the dataset, show computations for normalizing the values of Height and Income attributes such that the transformed values lie in the range of 0 to 1.

Name	Height (in cms)	Income
Manas	170	40,000
Shreya	160	30,000
Surbhi	180	50,000
Hemant	175	45,000
Paarth	165	55,000

- (b) Why is K-Nearest Neighbor algorithm considered as an instance-based learner? (8)

Consider the two-dimensional labelled data set with attributes Brightness and Saturation and class label Color given below:

Brightness	Saturation	Color
40	20	Red
50	50	Blue
60	90	Blue
10	25	Red
70	70	Blue
60	10	Red
25	80	Blue

Classify the data point where Brightness = 20 and Saturation = 35 according to the 5-nearest neighbor algorithm. Use the majority voting scheme and Euclidean distance.

- Q6. (a) What is the need for sampling in data mining? How can one determine an appropriate sample size? Explain the suggested approach. (5)
- (b) Consider the following dataset of customers visiting a fitness centre:

Emp. Code	Name	Membership Type	Date of Joining	Monthly Fee
M1	Riya	Premium	12-03-2020	₹5000
M2	Arjun	Standard	25-07-2021	₹3000
M3	Kavya	Premium	10-01-2019	₹5200
M4	Rohit	Basic	18-09-2022	₹2000
M1	Riya	Premium	12-03-2020	₹5000

- Identify the type of attributes as nominal, ordinal, interval, or ratio. Give justification for each. (5)
 - What is de-duplication of data? Why is it necessary? (2)
 - State the advantages of aggregation. Suggest a method to aggregate Monthly Fee attribute. (3)
- Q7. (a) Given two initial centroids C1(2.0, 3.0) and C2(9.0, 8.0), show clearly the clusters and new cluster centres obtained after each iteration of the K-means algorithm for two iterations using the following dataset. Use Euclidean distance as the distance metric. (12)

Points	X	Y
P1	2.0	3.0
P2	3.0	4.0
P3	4.0	5.0
P4	6.0	6.0
P5	7.0	7.0
P6	9.0	8.0
P7	5.5	5.0

- (b) Explain the following types of clusters by giving suitable examples: (3)
- Centre-based clusters
 - Density-based clusters



[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 1806

L

Unique Paper Code : 2342012401

Name of the Paper : Design and Analysis of Algorithms

Name of the Course : B.Sc. (H) Computer Science

Semester : IV

Duration : 3 Hours

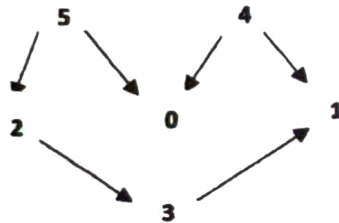
Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** is Compulsory.
3. Attempt any **four** questions from **Section B**.
4. Parts of a question must be answered together.
5. Write all the intermediate steps, wherever required to get full credits.

Section A

- Q1. (a) List all the topological orderings of the Directed Acyclic Graph (DAG) G given below.



3

- (b) Consider the following procedure SUM-ARRAY which computes the sum of n numbers in array $A[1:n]$.

SUM-ARRAY (A, n)

```
1  sum = 0
2  for i = 1 to n
3    sum = sum + A[i]
4  return sum
```

State a loop invariant for this procedure, and use its initialization, maintenance, and termination properties to show that the SUM-ARRAY procedure returns the sum of the numbers in $A[1:n]$.

3

- (c) Quicksort partitions an array into two subarrays based on a pivot element. The running time depends on how balanced the partition is. Suppose at each step, the pivot divides the array into two parts in the ratio 1:3. Formulate the recurrence relation for the running time of Quicksort in this case. Solve the recurrence and determine the time complexity.

3

- (d) Consider a hash table of size 11 with the hash function $h(k) = k \bmod 11$. Insert the keys 21, 32, 43, 54, 65, and 76 into the table. Handle any collisions using linear probing. 3
- (e) A matrix multiplication technique reduces the number of recursive multiplications by dividing matrices into submatrices and combining results using fewer multiplications than the conventional method. Identify the algorithm and compute the product of the following matrices using the same algorithm.

Given two matrices:

$$A = \begin{bmatrix} 1 & 7 \\ 3 & 5 \end{bmatrix} \quad B = \begin{bmatrix} 6 & 4 \\ 2 & 8 \end{bmatrix}$$
4

- (f) Consider the interval scheduling problem wherein we are given a resource and set of requests each having a start time and finish time. The goal is to maximize the number of requests scheduled. Show that the greedy strategy of selecting the request with the fewest incompatible requests does not always yield an optimal solution for this problem. 4
- (g) Let $S \subset V$ be a subset of vertices in a connected, weighted, undirected graph, and let $V-S$ denote the remaining vertices. Suppose an edge (u,v) has the minimum weight among all edges connecting S and $V-S$. Determine whether this edge should be included in a Minimum Spanning Tree (MST). Justify your answer. 5
- (h) A delivery company has packages with weights $\{2, 5, 7, 3, 9\}$ and truck capacity 15 kg. Design a greedy algorithm to maximize the number of packages loaded. Explain its time complexity. 5

Section B

- Q2. (a) Which sorting technique is most efficient if the range of input data is not significantly greater than the number of elements to be sorted and runs in $\Theta(n+k)$ time, where n is the size of the array and k is the range of the non-negative key values. Sort the array $A = \{5, 9, 0, 1, 5, 4, 2, 0, 6, 9\}$ using the same algorithm. 7
- (b) Consider the following two algorithms used for sorting a list of numbers:
- Algorithm A preserves the relative order of equal elements and follows a divide-and-conquer approach, using extra memory.
 - Algorithm B uses a tree-based structure, requires minimal extra space, but does not preserve relative order.
- i. Identify Algorithm A and Algorithm B.
 - ii. Identify the property that ensures preservation of the relative order of equal elements.

- iii. State the time complexity of Algorithm B.
 iv. Analyse the running time complexity of the algorithm A by formulating and solving a recurrence relation for the same.

8

Q3.

- (a) The Quick sort algorithm produces highly unbalanced partitions for already sorted inputs. Analyze this situation by formulating and solving a recurrence relation to determine its time complexity. Suggest an improvement to the pivot selection strategy to reduce the likelihood of this behaviour. Justify your answer.

7

- (b) A city's emergency response system is modelled as a graph, where locations such as hospitals, fire stations, and intersections are represented as vertices, and roads connecting them are represented as edges. All roads are assumed to have equal travel time. Suggest an algorithm for the control centre to determine the minimum number of roads required to reach every location from a central emergency hub A in the following graph. Find out the path determined by the algorithm from A to all other locations.

8

The graph is defined as follows:

Vertices: {A, B, C, D, E, F, G}

Edges: {A-B, A-C, B-D, B-E, C-F, E-F, D-G, F-G}

Q4.

- (a) The subset-sum problem involves a given set $X = \{x_1, x_2, x_3, \dots, x_n\}$ and a positive integer target sum S. Solve the problem for the set $X = \{2, 3, 7, 8, 10\}$, $S = 11$ and determine its running time.

7

- (b) Consider a graph with n vertices and m edges, where all edge weights are non-negative. Analyze the time complexity of Dijkstra's algorithm when implemented using a priority queue (min-heap). Explain the role of the priority queue in improving the efficiency of the algorithm.

8

Q5.

- (a) A digital media company schedules video programs on a streaming platform. Each program has a start time, end time, and associated revenue, and only one program can be shown at a time. Which algorithm will maximize the total revenue? Also, determine the optimal set of programs that maximizes the revenue. Explain why a greedy approach may not yield the optimal solution.

Program	Start	End	Revenue
P1	1	3	5
P2	2	5	6
P3	4	6	5
P4	6	7	4
P5	5	8	11
P6	7	9	2

7

- (b) Write a recursive algorithm to find the i^{th} Fibonacci number. How can the algorithm be improved using dynamic programming? Compute the 6th Fibonacci number using the dynamic programming.

8

Q6.

- (a) "Bubble sort introduces a flag variable which is set to true if no exchange is made for an entire pass over the array. It reduces the number of passes." Explain with an example, how this variable helps reducing the number of passes.

7

- (b) A traveller has a backpack with a maximum capacity of 10 kg. There are 5 items available:

8

Item	Weight (kg)	Value
1	2	6
2	3	10
3	5	12
4	4	8
5	1	3

Each item can either be taken completely or left. Find the maximum total value the traveller can carry without exceeding the capacity and list the selected items. Design a dynamic programming table to solve the problem and state its time complexity.

Q7.

- (a) Write the best-case and worst-case time complexities of insertion and deletion operations in hashing when collisions are resolved using the chaining method. Explain it with an example.
- (b) What is a bipartite graph? Describe how Breadth-First Search (BFS) can be used to check whether a given graph is bipartite. State its time complexity.

7

8

[This question paper contains 8 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 1925

L

Unique Paper Code : 2342012402

Name of the Paper : Database Management Systems

Name of the Course : B.Sc. (H) Computer Science (NEP-UGCF)

Semester : IV

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. The paper has **two** sections A and B. Section A is compulsory.
3. Attempt any **four** questions from Section B. Each question is of **15** marks.
4. Answer **all** parts of a question together.

Section A

- 1 (a) Differentiate between DDL and DML commands. Give one example for each type of command. (3)
- (b) Define the following: Primary Key, Candidate Key and Super Key. (3)
- (c) Consider the following relations in a database: (3)
 - COURSE (CourseID, CourseName, Credits)
 - STUDENT (StudentID, Name)
 - REGISTRATION (StudentID, CourseID, Grade)

Foreign Keys in relation REGISTRATION:

 - StudentID references STUDENT (StudentID)
 - CourseID references COURSE (CourseID)

कालिन्दी महाविद्यालय पुस्तकालय
KALINDI COLLEGE LIBRARY

P.T.O.

Identify the type of relational model constraint violated (if any) in each of the following cases:

- i. Inserting a tuple into REGISTRATION with a StudentID that does not exist in STUDENT.
- ii. Updating a tuple in STUDENT to set StudentID as NULL.
- iii. Deleting a tuple from COURSE that is referenced in REGISTRATION.

- (d) Identify and explain the problem associated with the concurrent execution of transactions T1 and T2 in the following schedule. (3)

TIME	T1	T2
1	read(W)	
2	$W = W - 40$	
3		read(W)
4		$W = W - 10$
5	write(W)	
6	read(X)	
7		write(W)
8	$X = X + X$	
9	write(X)	

- (e) State any three characteristics of the database approach. (3)
- (f) List the ACID properties of a database transaction. How are Atomicity and Consistency related? (3)
- (g) Design an entity EMPLOYEE with the following attributes using ER Diagram notation: (3)
- EmpID
 - Name
 - Address (Street, City, Zip)
 - Multiple Phone Numbers
 - Age derived from DOB

(h) Consider the following relations:

(3)

- EMPLOYEE (SSN, Name, Salary, SUPRSSN, DNO)
- DEPARTMENT (DNO, DName)

In the EMPLOYEE relation:

- SUPRSSN is a foreign key referencing SSN in EMPLOYEE (supervisor relationship).
- DNO is a foreign key referencing DNO in DEPARTMENT.

Write the SQL CREATE TABLE statement for EMPLOYEE, including appropriate FOREIGN KEY constraints and clauses so that referential integrity is maintained when a referenced department (DNO) is deleted or updated.

(i) Differentiate between UNION and OUTER UNION with the help of an example.

(3)

(j) Consider the following relation:

(3)

A	B	C
a1	b1	c1
a1	b1	c2
a2	b1	c1
a2	b1	c3

For the given relation state, do the following functional dependencies hold (justify your answer):

- $A \rightarrow B$
- $A \rightarrow C$
- $B \rightarrow A$

Section B

2 (a) A database administrator changes the internal storage structure to improve performance, while users continue to access the database without any changes to their applications.

(6)

- Define Data Independence.
- Which concept of data independence is illustrated in the above scenario?
- How does the three-schema architecture support Data Independence?

- (b) Suppose we have an ordered data file with $r = 50,000$ records, stored on a disk (9)
with block size $B = 8,192$ bytes. The file records are of fixed size and unspanned,
with record length $R = 128$ bytes.

A primary index is built on the file. The index has:

- Ordering key field size $V = 12$ bytes
- Block pointer size $P = 8$ bytes

Compute the following:

- i. Blocking factor of the data file (b_{fr})
- ii. Number of blocks in the data file (b)
- iii. Record size of the index file (R_i)
- iv. Blocking factor of the index file (b_{fr_i})
- v. Number of blocks in the index file (b_i)
- vi. Total number of block accesses required for binary search on the index file.

- 3 (a) Consider the following relation. (6)

$R = \{A, B, C, D, E\}$

and the set of functional dependencies:

$F = \{AB \rightarrow C, A \rightarrow D, B \rightarrow E\}$

Answer the following:

- i. Find the primary key of R .
 - ii. Explain why R is not in Second Normal Form (2NF). Decompose R into corresponding 2NF relations.
- (b) Specify the following queries in Relational Algebra on the given database schema: (9)

PRODUCT (ProdNo, Pname, Category)

SALES (SaleNo, Sdate, ProdNo, Amount)

DELIVERY (SaleNo, StoreNo, DelDate)

- i. List the SaleNo and DelDate for all sales delivered from StoreNo 'S1'.

- ii. List all SaleNo that were delivered either from StoreNo 'S1' or StoreNo 'S2'.
- iii. List all Pname for which exactly two sales have been delivered.

- 4 (a) Consider the following relations T1 and T2 as follows:

(6)

T1

A	B	C
1	10	x
2	20	y
3	30	z
4	40	x

T2

P	Q	R
2	20	y
3	10	x
5	50	z
4	40	w

Assuming that the relations T1 and T2 are union compatible, perform the following operations and show the resulting relations:

- i. $T1 \cap T2$
- ii. $T1 \bowtie_{T1.A = T2.P} T2$
- iii. $T1 - T2$

- (b) Consider a MOVIE database that stores information about the movie industry. Using the given data requirements, identify the primary keys, relationships, cardinality ratios, and participation constraints, and draw the corresponding Entity-Relationship (ER) diagram. (9)

- Each movie is identified by a title and a year of release. Each movie has a length in minutes. Each has a production company, and each is classified under one or more genres (such as horror, action, drama, and so forth). Each movie has one or more directors, and one or more actors appear in it. Each movie also has a plot outline.
- Actors are identified by name and date of birth and appear in one or more movies. Each actor has a role in the movie.
- Directors are also identified by name and date of birth and direct one or more movies. It is possible for a director to act in a movie (including one that he or she may also direct).
- Production companies are identified by name, and each has an address. A production company produces one or more movies.

5 (a) Differentiate between:

- i. Entity Type and Entity Set
- ii. Database Administrator and Database Designer
- iii. Natural Join and Equi-Join

(9)

(b) Consider the following schema:

STUDENT (Sid, Sname)
COURSE (Cid, Cname)
DEPARTMENT (Did, Dname)
ENROLLMENT (Sid, Cid, Did)

The database records information about students, courses, departments, and includes a ternary relationship ENROLLMENT among students, courses, and departments.

Write SQL queries to perform the following:

- i. List the courses that have at least one student enrolled.
- ii. Produce a listing of department ids along with department names and the total number of students enrolled in each department.
- iii. Retrieve the department names, student names, and course names for those records where the student's name starts with the letter 'A'.

6 (a) Consider the entities SUPPLIER, PART, and PROJECT.

(6)

- A Supplier supplies certain parts.
- A Project requires specific parts.
- A supplier may supply different parts to different projects, and the same part may be supplied by different suppliers for different projects.

The relationship captures the fact that a supplier supplies a particular part for a given project.

- i. Draw an ER Diagram showing a ternary relationship amongst the above entities.
- ii. Explain why this relationship cannot be replaced by three binary relationships (SUPPLIER-PART, PART-PROJECT, SUPPLIER-PROJECT).

- iii. Under what conditions can a ternary relationship be represented using binary relationships?

- (b) Consider the following relations for a bookstore database:

(9)

BOOK

ISBN	Title	Publisher	Price
B101	DBMS Concepts	McGraw Hill	500
B102	Data Structures	Pearson	450

CUSTOMER

CustID	Cname	City
C1	Amit	Delhi
C2	Neha	Mumbai

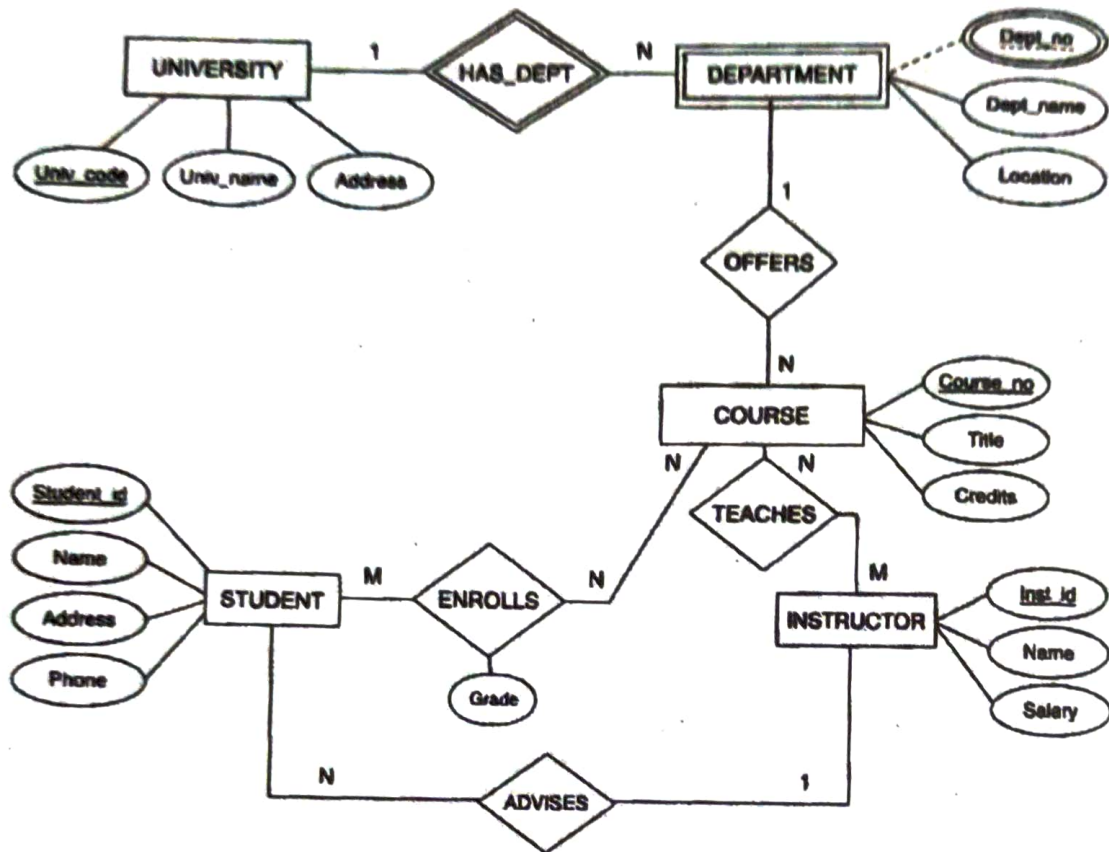
SALE

SaleID	ISBN	CustID	Sale_date	Sale_price
S1	B101	C1	2026-04-01	400
S2	B102	C2	2026-04-02	420

- Identify the primary key and foreign keys in the SALE relation.
- Give one example of an update operation on SALE that will violate referential integrity. Suggest appropriate action to handle this situation.
- Write an SQL statement to create the SALE table with appropriate primary key, foreign key constraints, and a condition that Sale_price must be between 200 and 5000.

- 7 (a) Explain the temporary update problem and the incorrect summary problem that may arise in the absence of concurrency control mechanisms. Illustrate each case with a suitable transaction schedule. (6)

- (b) Consider the ER diagram shown below for part of a University database. Each university can have multiple departments, and each department can offer multiple courses. Each course can be enrolled in by multiple students and taught by multiple instructors. Each student is advised by one instructor. (9)



Map the above Entity Relationship (ER) diagram to its corresponding relational schema. Specify all primary and foreign keys clearly.

9

[This question paper contains 4 printed pages.]

Your Roll No.....

L

Sr. No. of Question Paper : 2032
Unique Paper Code : 2342012403
Name of the Paper : Computer Networks
Name of the Course : B.Sc. (H) Computer Science – DSC (NEP-UGCF-2022)
Semester : IV
Duration : 3 Hours
Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. The paper has **two** sections. **All** questions in **Section A** are compulsory.
3. Attempt any **four** questions from **Section B**.
4. Parts of a question must be attempted together.

Section A

- 1 (a) What is the sliding window protocol? How does it improve efficiency over Stop-and-Wait protocol? If the sequence number field is 3 bits long, what is the maximum window size for Selective Repeat protocol? 3
- (b) A synchronous TDM system transmits 4000 frames per second. Each frame consists of 12 slots, and each slot carries 8 bits. Additionally, 2 extra bits are added to each frame for synchronization. 3
 - (i) What is the size of each frame (in bits)?
 - (ii) Give the transmission rate (in bps).
 - (iii) What is the duration of each output frame?
- (c) Explain the significance of "signal levels" in the Nyquist formula. How does increasing the number of levels affect the data rate? 3
A channel uses four signal levels and has a bandwidth of 150 KHz.
What is the maximum data rate achievable using the Nyquist formula?
- (d) In a bus topology, the total bandwidth is 10 Mbps. 3
Due to collisions and retransmissions, only 70% of the bandwidth is effectively utilized. Calculate the effective data rate of the network.
If 4 devices share the channel equally, what is the average data rate per device?
Explain why collisions are common in bus topology.
- (e) A satellite communication system is used to provide connectivity across remote and geographically dispersed regions. 3
 - (i) Explain how satellite communication enables wide-area or global coverage.
 - (ii) Write any two limitations associated with satellite communication.
- (f) A TCP connection is established over a communication link with a bandwidth of 5 Mbps, a round trip time (RTT) of 40 ms, and a TCP window size of 20,000 bits. 3
Calculate the actual throughput of the connection and comment on whether the link is fully utilized or not based on the result.

- (g) A noiseless channel has bandwidth 4 kHz and uses 8 signal levels. A noisy channel has the same bandwidth with SNR = 20 dB. Which channel has a higher maximum data rate? Justify. 3
- (h) A communication channel has a bandwidth of 10 Mbps between Device A and Device B. 3
- In Half-Duplex mode, A and B alternately transmit for 2 seconds each. Calculate the effective data rate experienced by each device over time.
 - In Full-Duplex mode, both devices transmit simultaneously using the same 10 Mbps link. Explain how the bandwidth is utilized and determine the effective data rate for each device?
- (i) Modern computer networks rely on different devices operating at various layers of the OSI model. 3
- Compare Hub, Switch and Router in terms of:
- Layer of operation
 - Functionality
- (j) In a computer network, a host needs to send data to another host. It knows the destination IP but not the MAC address. Write the name of the protocol and explain its working. 3

Section B

- 2 a) TCP is considered a reliable protocol, while UDP is called an unreliable and connectionless protocol. Justify the above statement by comparing the mechanisms such as handshaking, error control, reliability, ordering of segments and congestion control. Provide examples of applications that benefit from each protocol. 7
- b) Answer the following questions with respect to the Stop and Wait Protocol: 8
- How does the Stop-and-Wait protocol handle duplicate frames?
 - Why is a timer necessary at the sender side in the Stop-and-Wait protocol? What could happen if there's no timer?
 - What is the significance of sequence numbers being only 1 bit in Stop-and-Wait ARQ?
 - What happens if the sender sends a frame and the receiver receives it correctly, but the ACK frame is lost? Illustrate the complete flow with the help of a sequence diagram.
- 3 a) Consider a network of six routers labelled A, B, C, D, E, and F. The cost of the links between the routers is as follows: 8
- A-B: 3
A-C: 1
B-C: 7
B-D: 5
C-D: 2

D-E: 7

C-F: 8

E-F: 2

Using Dijkstra's algorithm, determine the shortest path from router A to router E and show all intermediate steps.

Also, State the Optimality Principle.

- b) Discuss the advantages and disadvantages of allowing browsers to be extended using plug-ins and helper applications. 7
 Can a single-threaded web server handle multiple client requests at the same time? Justify your answer with reasons related to server architecture and performance. Describe the role of multithreading in improving Web server throughput.
- 4 a) Why must a station in CSMA/CD transmit for a minimum duration of 2τ before it can assume the channel has been seized? Illustrate using a diagram. 7
 In a CSMA/CD network, a station experiences a collision after transmitting a frame. If this is the third collision, calculate the maximum wait time before the next transmission attempt. Assume a slot time of $50 \mu s$.
 What is the significance of the "backoff time" in CSMA/CD?
- b) An organization is assigned the IP address 192.168.10.0/24. The network administrator needs to create two subnets to accommodate 120 and 30 hosts respectively. 8
 For each subnet:
 (i) Provide the subnet mask in dotted-decimal notation
 (ii) List the first and last valid host IPs
 (iii) Indicate the broadcast address
- 5 a) Explain the purpose of the following ICMP messages: 7
 (i) Destination Unreachable
 (ii) Time Exceeded
 (iii) Echo and Echo Reply
 (iv) Timestamp Request and Timestamp Reply
 (v) Source Quench
 (vi) Redirect
 (vii) Router Advertisement/ Solicitation
- b) What is Cyclic Redundancy Check (CRC)? A bit stream 10011101 is transmitted using the standard CRC method. The generator polynomial is $x^3 + 1$. Show the actual bit string transmitted. Suppose that the third bit from the left is inverted during transmission. Show that this error is detected at the receiver's end. 8

- 6 a) A data frame of 600 bits is transmitted using both bit stuffing and byte stuffing. 8
In bit stuffing, there are 50 occurrences of five consecutive 1s.
In byte stuffing, 10 flag bytes appear in data.
- i. Calculate the overhead in both techniques. Which technique is better for this case?
 - ii. Consider that during transmission over a noisy channel, 2 stuffed bits are flipped in the bit-stuffed frame.
Explain how this affects frame synchronization. Can the receiver correctly recover the original data? Suggest a method to detect such errors.
- b) What are the different categories of UTP cables defined by the EIA? Explain each with its specifications and applications. 7
- 7 a) Compress the 128 bit (16 bytes) IPv6 address written as eight groups of four hexadecimal digits 8000:0000:0000:0000:0123:4567:89AB:CDEF. 3
- b) A file transfer application requires reliable delivery, while a live video streaming application prioritizes speed over accuracy. Explain the types of communication services suitable for each case. 4
- c) A TCP connection has a bandwidth of 10 Mbps and RTT of 20 ms. 4
- (i) Calculate the minimum window size required to fully utilize the bandwidth.
 - (ii) Explain the significance of this result in terms of network performance.
- d) Discuss how propagation characteristics in terms of reflection, refraction and diffraction affect the performance of radio wave and microwave communication. 4

10

[This question paper contains 2 printed pages.]

Your Roll No.....

L

Sr. No. of Question Paper : 278

Unique Paper Code : 2353010015

Name of the Paper : Research Methodology

Name of the Course : Bachelor of Science (Honours Course) / Bachelor of Arts / Bachelor of Science (Programme) – DSE

Semester : VI

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
 2. Attempt all the six questions, selecting any two parts from each question.
 3. Parts of the question are to be attempted together.
 4. All the six questions carry equal marks.
-
1. (a) List and elaborate the general principles of mathematical writing.
(b) What are the various ways for practicing good mathematical judgement? Explain in detail.
(c) What is the format of a mathematical research paper? Suppose that there is a journal called 'XYZ', which publishes research papers in Mathematics. What is the unique international identifier, associated with this journal 'XYZ' called?
 2. (a) What are the three questions that should be considered, when formulating a definition in a mathematics research paper? Explain the importance of each of these questions.
(b) State the important points that should be kept in mind, when citing papers/ references in a research paper.
(c) What are the key points that should be taken into consideration, before and after submission of a manuscript to a journal, for possible publication?
 3. (a) Explain the use of each of the following LaTeX commands: \label, \ref, and \cite. Give examples for each of the above commands.

काठिन्दी महाविद्यालय पुस्तकालय
KALINDI COLLEGE LIBRARY

P.T.O.

- (b) Write a code in Latex to obtain the following output :

Example of a Theorem Environment

We can now state and use our custom theorem environment.

Theorem 1 (Pythagoras Theorem). *Let a and b be the legs of a triangle of right angles and c be the hypotenuse. Then*

$$c^2 = a^2 + b^2.$$

We can also give a proof using the built-in proof environment:

- (c) What are the *equation* and the *align* environments in LaTeX? Explain the use of each of the two environments. State when, each one should be used.
4. (a) Write PSTricks code to draw a triangle and label its vertices A, B, and C.
- (b) Discuss the important features of a good poster, including content organization, visual design, and readability.
- (c) Make a Beamer presentation on the topic 'Importance of Mathematical Writing' with exactly five slides including a title slide, three content slides, and a thank you slide.
5. (a) List any four features of each of the following web resources for Mathematics: AMS and MAA.
- (b) What is meant by JCR? How is the metric 'impact factor' of journal, defined as per JCR? Give the purpose and challenges of this journal metric.
- (c) Write a short note on Google Scholar metrics. What are the two main indicators on the basis of which, publications are ranked by Google Scholar metrics. Explain these indicators in brief.
6. (a) What is Scopus? What are the key features of Scopus? What is the importance of Scopus, with respect to the development of research in Mathematics?
- (b) State the research ethics as standardized by COPE, concerning all parties involved in the conduct of research.
- (c) List out the steps/instructions for using the plagiarism checker 'Turnitin' software, for finding out the percentage of similarity of a given research paper.



[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 1726

L

Unique Paper Code : 2342013601

Name of the Paper : Artificial Intelligence

Name of the Course : B.Sc. (H) Computer Science

Semester : VI

Duration : 3 Hours

Maximum Marks : 90

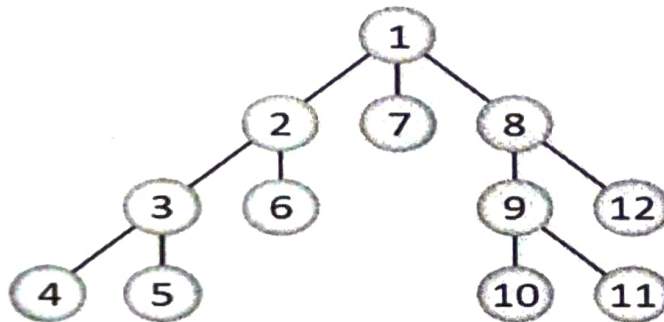
Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. The paper has **two** sections A and B. **Section A** is compulsory.
3. Attempt any **four** questions from **Section B**.
4. All parts of a question must be answered together.

Section A

(Compulsory)

- 1 (a) Explain the architecture of a learning agent with its components. (5)
- (b) Mention any two differences between Breadth First Search (BFS) and Depth First Search (DFS). Illustrate the result of BFS and DFS on the following tree. (5)



- (c) Consider the following predicates: (5)
 $P(f(x), g(x))$ and $P(y, y)$
 - i. Determine whether the above expressions can be unified.
 - ii. If unification is possible, find the most general unifier (MGU). If not, clearly explain the reason.
- (d) What is the Chomsky Hierarchy of Grammars? Write the properties of Type 0, Type 1, Type 2, and Type 3 grammars with examples. (5)
- (e) How do the cut (!) operator and fail predicate affect backtracking? Write a Prolog program to implement fact (N, F) that computes the factorial of a number N, storing the result in F. (5)

- (f) What are Conceptual Dependency (CD) primitives? Explain any two of the following: (5)
ATRANS, PTRANS, MTRANS, PROPEL, INGEST. Represent the following sentences using appropriate CD primitives:

- i. The boy threw the ball.
- ii. She ate an apple.

Section B

- 2 (a) Convert the following into clausal form: (7)
A1: If x is a student, then x studies.
A2: Rahul is a student.
A3: If someone studies, they pass.

Express the following sentences as conceptual dependency structures:

- i. Charlie drove the pickup fast.
- ii. Ram is a programmer.

- (b) A company is developing an intelligent self-driving car system. The system uses sensors to detect the environment, processes the information, and takes actions such as braking, accelerating, and steering. (8)

- i. Identify the components of an intelligent agent in this system by specifying:

- Sensors
- Actuators
- Environment
- Performance

- ii. Draw a diagram of the agent structure showing the interaction between the agent and its environment.

- iii. Is the self-driving car an example of Weak AI, Strong AI, or Narrow AI? Justify your answer.

- 3 (a) Prepare a script for "Online shopping" using appropriate script elements (roles, props, places, scenes, entry conditions and post conditions etc.). (7)

- (b) Explain Turing Test approach to AI. How is Turing Test approach different from the rational Agent approach? (8)

- 4 (a) Differentiate between the following: (8)

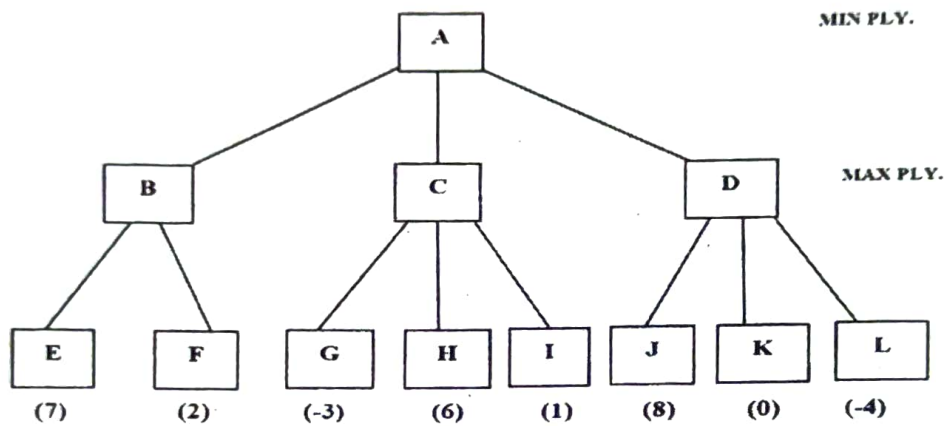
- i. Fully observable vs. partially observable environment
- ii. Deterministic vs. stochastic environment
- iii. Model-based agent and Goal-based agent
- iv. Goal-based agent and Utility-based agent

(b) Write a Prolog program to perform the following list operations: (7)

1. Check whether an element is a member of a list.
2. Find the length of a list using recursion.
3. Concatenate two lists.
4. Find the sum of elements in a list.

5 (a) Define Heuristic Search technique. Why should a heuristic function of A* underestimate? (7)

(b) Consider the following game tree with ply depth 2 where the indicated scores are from the MIN player's point of view. Which move A should choose and why? Which nodes will be pruned according to the α - β pruning procedure? Give justifications of each. (8)



6 (a) Using the *Constraint Satisfaction approach*, solve the following crypt-arithmetic problem: (7)

$$\begin{array}{r}
 S E N D \\
 + M O R E \\
 \hline
 M O N E Y
 \end{array}$$

Show all the steps.

(b) What are the differences between Recursive Transition Network (RTN) and Augmented Transition Network (ATN)? (8)

Given a grammar $G = (V_T, V_N, S, P)$

where:

V_T = Set of terminal symbols,

V_N = Set of non-terminal symbols,

S = Start symbol

P = Production rules

$S \rightarrow NP VP$
 $NP \rightarrow DET N$
 $VP \rightarrow V PP$
 $PP \rightarrow PREP NP$
 $N \rightarrow \text{bird} \mid \text{tree}$
 $V \rightarrow \text{flew}$
 $DET \rightarrow \text{The} \mid \text{a}$
 $PREP \rightarrow \text{over}$

Construct the parse tree for "The bird flew over the tree."

7

Consider the following axioms:

(15)

- i. **A1:** Sneha likes all easy subjects.
- ii. **A2:** Mathematics and Physics are easy subjects.
- iii. **A3:** Anything that anyone studies and passes without extra help is an easy subject.
- iv. **A4:** Rohan studies Chemistry and passes without extra help.
- v. **A5:** Anyone who takes extra help does not pass without extra help.
- vi. **A6:** Meera studies every subject Rohan studies.

Express the above axioms into First Order Predicate Logic (FOPL) statements and convert them into clausal form. Using resolution principle, prove the statement "Sneha likes Chemistry."

12

[This question paper contains 8 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 1745

L

Unique Paper Code : 2342013602

Name of the Paper : Machine Learning

Name of the Course : B.Sc. (H) Computer Science

Semester : VI

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Attempt **all** questions from **Section-A**.
3. Attempt any **four** questions from **Section-B**.
4. Attempt all parts of the question together.
5. Use of Scientific Calculator is allowed.

SECTION A

1. (a) A student trains a logistic regression model on a medical dataset to predict cancer (positive class = 1). The model achieves 98% accuracy, but the doctor rejects it immediately. Explain precisely why is the doctor right to reject it, what metric should be used instead, and give its mathematical formula. (3)
- (b) For each of the following three real-world applications (A1, A2, A3) of machine learning, identify the primary type of machine learning involved (Supervised, Unsupervised, or Reinforcement) and provide a one-sentence reason for your choice. (3)

- A1: A spam filter learns to identify emails as “spam” or “not spam” using thousands of previously labelled emails.
 - A2: A self-driving car learns to navigate a highway by receiving a reward for staying in its lane and a penalty for swerving, without being told every correct steering angle.
 - A3: An online news website groups articles into topics (e.g., sports, politics, tech) based only on the words they contain, without any pre-assigned topic labels.
- (c) What is the bias-variance trade-off? Is it feasible to minimize both bias and variance at the same time? Explain briefly. (3)
- (d) Normalize the given data using standardization method (3)
- Data (X): [2, 4, 11, 23, 27, 29]
- (e) A classification model performs with a high accuracy on training data, but generalizes poorly to new instances. Identify the problem and briefly explain one method to overcome this problem. (3)
- (f) Describe and illustrate the Sigmoid function. Determine its derivative as well. (3)
- (g) Show that the XOR function is not linearly separable, and hence cannot be solved by a neural network with no hidden layer. (3)
- (h) A dataset at a node of a decision tree has the following class distribution:
- Node X: 12 samples of “Yes”, 0 samples of “No”
 - Node Y: 6 samples of “Yes”, 6 samples of “No” (3)
- (i) Compute the entropy H for both Node X and Node Y. Show all the calculations clearly.
- (ii) Which node is pure and which is maximally impure based on the computed values?
- (i) Explain "why liner regression is not suitable for classification problems". (2)

- (j) A machine learning-based classifier was tested on 440 patients, yielding the following results: 100 patients were correctly diagnosed with the disease, 300 were correctly identified as healthy, 15 healthy patients were incorrectly classified as diseased, and 25 diseased patients were incorrectly classified as healthy.

Construct the confusion matrix for the given data. Also, calculate the accuracy, precision, and F1-score. (4)

SECTION B

- 2 (a) Find the regression line for the given dataset using the Ordinary Least Square method. Show the computation at each step. Predict the cost of a house if number of bedrooms are 8. Note that the Property ID column is an identification label (metadata) and should not be used as a numerical feature in the regression calculations. (5)

Property ID	Bedrooms	Cost (lakhs)
H-1	2	12
H-2	1	6
H-3	3	14
H-4	4	16
H-5	5	20

- (b) Describe the Best Subset Selection algorithm for feature selection. Also discuss its complexity. (5)

- (c) Consider a logistic regression model $\hat{p}(X) = \frac{e^{-6+0.005X}}{1+e^{-6+0.005X}}$ to classify whether a customer is a defaulter or not. Evaluate the model for two customers, each with a balance of Rs. 8000:

- Customer A actually Defaulted ($y=1$).
- Customer B actually Did Not Default ($y=0$).

Calculate the predicted probability and the total likelihood.

3. (a) Consider a fully connected feed-forward neural network with 3 input units (x_1, x_2, x_3) each set to 1.0, a single hidden layer containing 2 units, and 1 output unit, where all connection weights are 1.0, bias at hidden layer is -1.0, and the output layer unit has a bias of +1.0. All neurons utilize the ReLU activation function. Given a ground truth value of $y=1.0$, answer the following :

- Draw the schematic diagram for this network, clearly labelling all weights on the connections and the biases at each node.
- Calculate the total number of trainable parameters (total weights and total biases) in the network.

Perform a step-by-step forward propagation to determine the final predicted output $\hat{y}$. (5)

- (b) What is regularization? When would you choose Ridge Regression? What is the effect of the following on the model :

(i) The regularization parameter $\lambda \rightarrow 0$

(ii) The regularization parameter $\lambda \rightarrow \infty$ (5)

- (c) Suppose we have a classification dataset consisting of 2 classes A and B with 100 and 50 samples respectively. We use stratified sampling to split the data into train and test sets. Which of the following train-test splits would be appropriate? Justify your answer. (5)

ST. Train- {A : 80 samples, B : 30 samples}, Test- {A : 20 samples, B : 20 samples}

S2: Train- {A : 20 samples, B : 20 samples}, Test- {A : 80 samples, B : 30 samples}

S3: Train- {A : 80 samples, B : 40 samples}, Test- {A : 20 samples, B : 10 samples}

S4: Train- {A : 20 samples, B : 10 samples}, Test- {A : 80 samples, B : 40 samples}

4. (a) A medical diagnostic laboratory uses K-Nearest Neighbors (KNN) to classify whether a patient has a "High Risk" or "Low Risk" of a specific metabolic condition based on two features: Glucose Level (x_1) and BMI (x_2). The training dataset of 5 patients is given below: (5)

Patient Id	Glucose (x_1)	BMI (x_2)	Risk Level (y)
P1	100	18	Low
P2	110	22	Low
P3	105	25	Low
P4	135	32	High
P5	150	35	High

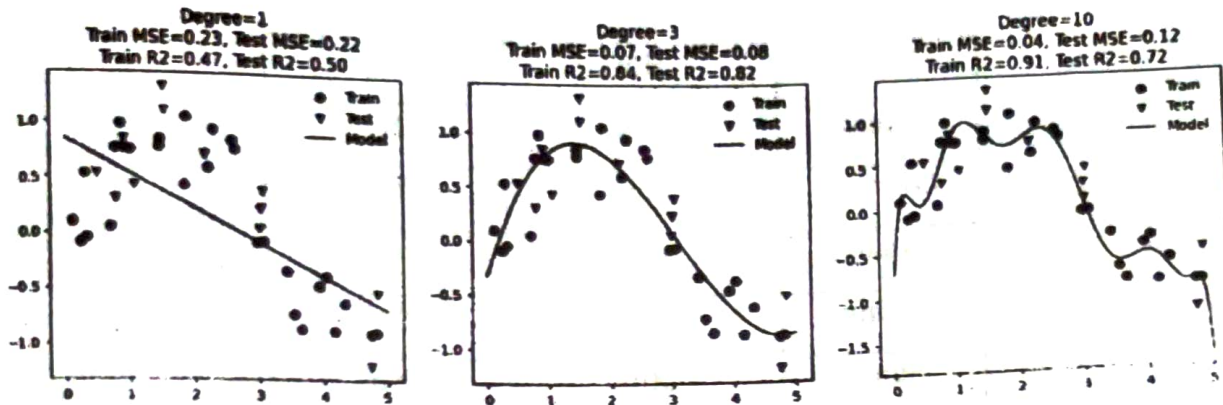
A new patient arrives with Glucose (x_1) = 130 and BMI (x_2) = 28. Using the Euclidean Distance metric and ($K=3$), determine the Risk Level for the new patient. Show all the steps.

- (b) Distinguish between hard margin and soft margin in SVM. Write the constraint that must be satisfied to identify support vectors. Also analyse the role of C and ξ_i in controlling margin violations. (5)
- (c) Why is it important to perform standardization before applying Principal Component Analysis (PCA) method? Give the formal equation for computing first principal component and its interpretation in PCA method. (5)
- 5 (a) What is the key assumption of a Naive Bayes Classifier? Using Naive Bayes classification method for the following data, predict whether a working individual with Age = Old and Income = High will buy a computer or not. (5)

Id	Age	Income	Working	Buy Computer
1	Young	High	No	No
2	Young	High	No	No
3	Middle	High	No	Yes
4	Old	Medium	No	Yes
5	Old	Low	Yes	Yes
6	Old	Low	Yes	No
7	Middle	Low	Yes	Yes
8	Young	Medium	No	No

- (b) You are using a One-vs-All (OVA) approach for a 10-class problem. You notice the confidence scores are inconsistent. Logically, what is the primary weakness of OVA compared to One-vs-One (OVO)? (5)
- (c) Given a dataset of five points : A(1, 2), B(2, 1), C(4, 4), D(5, 4), and E(2, 5), perform hierarchical clustering using the single linkage method.
- Compute the initial Euclidean distance matrix.
 - Show the stepwise merging of clusters and the updated distance calculations.
 - Draw the final dendrogram, clearly labelling the merge distances on the vertical axis. (5)
6. (a) A neuron has two inputs $x_1=0.5$ and $x_2=0.8$ with weights $w_1=-0.2$ and $w_2=0.5$. If the bias $b=-0.1$, calculate the weighted sum(z) and find output using a Sigmoid activation function (output 1 if threshold ≥ 0.5 else 0). (5)

(b)



Consider polynomial regression models of degree 1, 3, and 10. Based on the given plots and their associated performance metrics (training and test error/R²), answer the following :

- Which model is the most suitable for deployment?
- Explain the primary mathematical evidence (from the metrics above the plots) that justifies this selection over the other two options. (5)

(c) Given a dataset with four points: A(1,1), B(2,1), C(4,3), and D(5,4).

The points are grouped into two clusters:

- Cluster 1: {A, B}
- Cluster 2: {C, D}

Calculate the centroid for each cluster and the total SSE for this clustering scheme. (5)

7. (a) Using the ID3 Decision Tree algorithm, determine the root node based on information gain from the given dataset considering 'Play Game' as the target variable: (5)

Weather	Temperature	Play Game
Sunny	Hot	No
Sunny	Hot	No
Overcast	Hot	Yes
Rainy	Mild	Yes
Rainy	Cool	Yes
Sunny	Mild	No

- (b) A linear regression model is used to predict house prices. The following table shows the Actual prices (y) and the Predicted prices ($\hat{y}$) for 4 houses (in lakhs) :

(5)

House No.	Actual prices (y)	Predicted prices ($\hat{y}$)
1	10	11
2	20	18
3	30	31
4	40	40

Compute the following Regression Evaluation Metrics using above table data

- Mean Squared Error (MSE).
- R^2 Score.

- (c) What happens to the training process if the learning rate is extremely high? Explain how it affects loss convergence and model stability with a suitable diagram.

(5)

13
[This question paper contains 2 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 1765

L

Unique Paper Code : 2342013604

Name of the Paper : Cloud Computing

Name of the Course : B.Sc. (Hons) Computer Science (DSC)

Semester : VI

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Question No. 1 (Section A) is compulsory.
3. Attempt any four (04) questions from Section B.
4. Parts of a question should be attempted together.

Section A

- Q1 (i). Define the term utility-oriented computing and provide one specific (3)
example of an application that is well-suited for this model.
- (ii). Define virtualization. State any two characteristics of a virtualized (3)
environment.
- (iii). List any three candidate sectors suitable for Community Cloud (3)
deployment.
- (iv). Distinguish between client responsibility and vendor responsibility in (3)
cloud security.
- (v). Differentiate between Capital Expenditure (CapEx) and Operating (3)
Expenditure (OpEx) in the context of cloud computing.
- (vi). How is parallel computing different from distributed computing? (3)
Explain using examples.
- (vii). Define OAuth and explain any two advantages of using this protocol (3)
over the traditional method of sharing a username and password with a
third-party application.
- (viii). What is data-intensive computing? Give any two application areas (3)
where it is used.
- (ix). Describe the fundamental architecture of the Google File System (GFS) (3)
and explain the specific roles played by the Master and Chunkservers
- (x). Which type of hypervisor would you choose if you just need to test a (3)
new Linux distro or run a piece of legacy software on your personal
computer without wiping your hard drive? Justify

Section B

- Q2 (i.) Explain the concept of Cloud Computing. List two key benefits and two significant limitations with a suitable example to support your answer. (9)
- (ii.) Explain the limitations of using traditional RDBMS for large-scale data-intensive applications. How does NoSQL address these limitations? (6)
- Q3 (i.) Explain the importance of hypervisors in the context of cloud computing. Differentiate between the types of hypervisors using suitable examples. (9)
- (ii.) For the following statements, state if true or false. Justify your answer. (6)
- a) Horizontal scaling is suitable for parallel computing.
- b) A private cloud is suitable for a collaboration between multiple hospitals that share a research database for cancer treatment.
- Q4 (i.) Explain the advantages of hardware virtualization. Describe the major hardware virtualization techniques in detail. (9)
- (ii.) Discuss the security service boundary in IaaS, PaaS, and SaaS. (6)
- Q5 (i.) Explain the three major milestones that led to the development of cloud computing. (9)
- (ii.) What are Service Level Agreements (SLAs)? Discuss their importance in cloud computing with key parameters. (6)
- Q6 (i.) What are the components of a distributed system? How do they form a unified system? (9)
- (ii.) A startup wants to build, deploy, and scale its own web application with low initial investment and fast setup. Which cloud service model and deployment type would you recommend? Explain with reasons. (6)
- Q7 (i.) Explain the classification of Cloud computing services and compare the different service models in terms of their characteristics, product types, and vendor offerings with suitable examples. (9)
- (ii.) What is the MapReduce programming model? Explain the specific roles of the Map and Reduce functions, and describe how the Shuffle and Sort phase bridges them. (6)

[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 2556

L

Unique Paper Code : 2353010015

Name of the Paper : Research Methodology

Name of the Course : Bachelor of Science (Honours Course) / Bachelor of Arts / Bachelor of Science (Programme) – DSE

Semester : VI

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Attempt all the six questions, selecting any two parts from each question.
3. Parts of the question are to be attempted together.
4. All the six questions carry equal marks.

1. (a) Complete the following sentences, that describe the rules that must be observed for avoiding errors and for the betterment of a mathematical writing.

(i) Begin sentences with words, not _____.

(ii) Adopt a _____ notation.

(iii) Use a balance of _____ and _____ while writing mathematical sentences or statements.

(iv) Write words fully, without _____.

कालिन्दी महाविद्यालय पुस्तकालय
KALINDI COLLEGE LIBRARY

P.T.O.

- (v) Write sentences in _____ style, not _____ style.
- (vi) Use math symbols within math expressions, not as _____ between math expressions and words.
- (b) List out the basic techniques that successful mathematicians have found useful in discovering mathematics and that are central part of the mathematical thought process.
- (c) What is the general sequence of components of a research/review paper. Explain each component in detail.
2. (a) Write short notes on the following terms, with reference to mathematical writing, giving an example in each case:
- (i) Parallelism
 - (ii) Otiose Symbols
 - (iii) Different kinds of ellipsis
- (b) What points should be kept in mind, while writing the 'Introduction' section of a research paper?
- (c) Explain the key factors that a researcher should consider, to choose an appropriate journal for publication of his/her research paper?
3. (a) Describe how mathematical expressions are written in LaTeX, explaining the LaTeX environments for producing, an in-text formula and a displayed formula. Give one example each, for a fraction, an integral, and matrix, to show how each of these is written using LaTeX commands.

- (b) Write a code in LaTeX software to obtain the following output :

1 Introduction

1.1 Importance of LaTeX

LaTeX is widely used for professional documents.

1. High-quality typesetting
2. Easy equations
3. Auto-numbering

$$x + y = 10 \quad (1)$$

$$x - y = 4 \quad (2)$$

- (c) What is cross-referencing in LaTeX? Explain the steps to add a figure, give it a label, and refer to it in the text
4. (a) What is the advantage of giving a presentation of the research work? Write a short note on poster presentation. Give the general format of a poster.
- (b) Make a Beamer presentation on the topic 'The Mathematical Research' with at least one bullet list and one equation.
- (c) Explain the use of the PSTricks commands `\begin{pspicture}`, `\pscircle`, and `\psline`, and describe how coordinates are used in PSTricks.
5. (a) Describe the following web resources for mathematics, with reference to research in Mathematics:
- (i) MAA
 - (ii) MathsSciNet
 - (iii) ArXiv

- (b) What is Impact Factor of a journal? Who introduced Impact Factor and when? Which database can be used to find the impact factor of a journal?
- (c) Explain in detail, the journal metrics SJR and MCQ.
6. (a) Describe research ethics concerning :
- (i) the participants
 - (ii) the researchers.
- (b) Explain the key features of the databases 'MathSciNet and 'Scopus'. Write the official websites for accessing these databases.
- (c) What do you mean by plagiarism with reference to a mathematical writing? Name any plagiarism checker software. What can be the possible penalties, that a researcher might face in case, the publisher finds out the researcher's work to be plagiarized?

13

[This question paper contains 8 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 1707

L

Unique Paper Code : 2342014801

Name of the Paper : Information Security

Name of the Course : B.Sc. (H) CS

Semester : VIII

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. The paper has **two** sections.
3. **Section A** is compulsory.
4. Attempt any **four** questions from **Section B**.
5. Parts of a question must be answered together.

SECTION A

1. (a) Briefly explain the three core objectives of computer security and why are they essential?
(3)
- (b) How is cryptanalysis different from brute-force attack?
(3)
- (c) Compare Symmetric Block Ciphers and stream ciphers.
(3)
- (d) Define three classes of subject in an access control system.
(3)
- (e) Discuss the SQLi in band attack types with suitable examples.
(3)
- (f) What are the typical phases of operation of a virus?
(3)

- (g) Define a Denial-of-Service (DoS) attack. What are the three primary system resources typically targeted to restrict or refine a service? (3)
- (h) List and explain the key elements of the Public Key Infrastructure X.509 (PKIX). (3)
- (i) What are the key factors contributing to higher security risk of wireless networks compared to wired networks? (3)
- (j) What is cross-site scripting (XSS)? How does it differ from SQL injection? (3)

SECTION B

2. (a) List and briefly define the kind of threat consequences and the types of threat actions which cause these consequences. (8)
- (b) Explain the step involved in a typical SQLi attack with a suitable diagram. (7)

3. (a) For the access matrix shown below make access control lists and capability lists. (8)

			OBJECTS		
		File 1	File 2	File 3	File 4
	User X		Own Read Write	Own Read Write	
SUBJECTS	User Y	Own Read Write	Read	Read	Write
	User Z	Read Write		Own Read Write	Read

- (b) Draw a step-by-step diagram of how a sender creates a digital signature and how the receiver verifies it. Discuss how Public-Key Certificates are used to distribute symmetric keys securely between two parties. (7)

4. (a) Define Role-Based Access Control (RBAC) Reference Models. (8)

Assume that a bank defines access based on Branch (20 branches), Employee Rank (5 levels), and Employment Type (Full-time, Part-time, Contract).

- (i) Calculate the number of roles required for RBAC.
 - (ii) If the bank adds a "Day of the Week" (7 values) constraint to the access policy, how many total roles will exist now?
 - (iii) Comment on the feasibility of this growth.
- (b) Describe the different types of attributes in the Attribute-Based Access Control (ABAC) model.

Assume that an application requires access control policies based on the applicant's age and the type

of funding to be provided. Using an ABAC approach, write policy rules for each of the following scenarios:

(i) If the applicant's age is more than 40, only "Research Grants (RG)" can be provided.

(ii) If the applicant's age is less than or equal to 40, both "RG and Travel Grants (TG)" can be provided. (7)

5. (a) What is a logic bomb? Explain the differences between a zombie, bot, and botnet. How are botnets used to launch DDoS attacks? (8)

(b) (i) Explain the fundamental mechanism of Stack Buffer Overflow attacks.

(ii) Draw a detailed diagram of a Stack Frame before and after a successful overflow attack. Contrast compile-time defenses with run-time defenses with suitable examples. (7)

6. (a) Explain the role of the sensors, analyzers, and user interface in a standard Intrusion Detection System (IDS) architecture. What is a Honeypot? Distinguish between low-interaction and high-interaction honeypots in terms of risk and information gathering. (8)

- (b) Explain the working of a packet filtering firewall.

What is the difference between application-level gateway and circuit-level gateway? Describe the situations where application-level gateway can be used and where circuit-level gateways can be used. (7)

7. (a) (i) Define the five major actors in terms of the roles and responsibilities in the cloud computing reference architecture.

- (ii) Describe Internet of Things (IoT) security environment and secure IoT framework.

(8)

- (b) (i) What are the security threats to wireless network? What are the techniques recommended for wireless network security?
- (ii) Describe the principle elements of a mobile device security strategy. (7)

16

[This question paper contains 8 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 7292

L

Unique Paper Code : 2343010029

Name of the Paper : Natural Language Processing

**Name of the Course : B.Sc. (H) / B.A. (P) DSE /
BSc. (P)**

Semester : VIII

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Question No. 1 (Section A) is compulsory.
3. Attempt any 4 (four) questions from Section B.
4. Parts of a question should be attempted together.

**कालिन्दी महाविद्यालय पुस्तकालय
KALINDI COLLEGE LIBRARY**

P.T.O.

Section A

1. (a) Explain the three different editing operations allowed in Minimum Edit Distance Calculation. (3)
- (b) How is a Recurrent Neural Network different from a Feed Forward Neural Network? (3)
- (c) Explain how Word2Vec model is used for computing embeddings? (3)
- (d) Describe the purpose of any two of the following methods offered by the 're module': (4)
 - (i) re.match()
 - (ii) re.sub()
 - (iii) re.finditer ()
- (e) Differentiate between Static and Contextual embeddings. (4)

(f) Why is lemmatization more accurate in NLP than stemming? Justify using an example. (4)

(g) Write regex patterns for the following: (4)

(i) 10-digit phone number

(ii) Only alphabets

(h) Apply Laplace Smoothing to compute the smoothed unigram probabilities for the vocabulary words given for the following data: (5)

Vocabulary = {'Mouse', 'Keyboard', 'computer'},

Total training words in the vocabulary is 5,

Counts: 'Mouse' = 3, 'Keyboard' = 2, 'computer' = 0.

What happens if smoothing is not used in n-gram language modelling?

Section B

2. (a) Name the three main architectures used in Transformer-based language models and briefly describe each. (5)

- (b) Apply Byte Pair Encoding (BPE) on the word: (5)

“low low lower”

Assume initial symbols are split into characters and $\langle /w \rangle$ marks end of word:

l o w $\langle /w \rangle$

l o w $\langle /w \rangle$

l o w e r $\langle /w \rangle$

- (c) What is prompting in large language models? Explain any two prompt engineering techniques with examples. (5)

4. (

3. (a) A document D has 100 words. Find which term is more important according to its TF-IDF.

Term	Count of terms in document D	Number of document containing term (Out of 10)
A	10	5
B	5	1

What happens if a term appears in all documents?

Justify your answer. (5)

(b) What is the purpose of an activation function?

Explain the following activation functions using their formulae : (10)

- Sigmoid
- Tanh
- ReLU
- Softmax

4. (a) What is the n-gram language modelling technique in NLP? Generate Unigrams, Bigrams and trigrams for the sentence 'I love my Country India'.

(5)

- (b) Why are word embeddings needed in NLP? Differentiate between one-hot encoding vectors and word embedding vectors. (5)
- (c) Why is cosine similarity used in NLP? Compute cosine similarity for two vectors.
 $A = [1, 2, 3]$, $B = [2, 3, 4]$ (5)
5. (a) Give advantages and disadvantages of the Bag of Words model for representing text documents. (5)
- (b) What is the utility of residual stream in transformer model? (5)
- (c) Briefly explain any two ethical and any three safety issues related to modern language models. (5)
6. (a) What is temperature sampling? How is it different from greedy sampling? (5)
- (b) Explain any five preprocessing steps in NLP. (5)

- (c) What is the role of attention in the transformer model? Explain the three matrices using which the attention head computes the attention vector. (5)

7. (a) Compare : (5)

- (i) Parallelized vs Sequential attention computation
- (ii) Autoregressive models and Bidirectional models

- (b) What is the XOR problem? Why can it not be solved using a single-layer perceptron? Explain the role of non-linear activation functions in solving the XOR problem. (5)

- (c) (i) Given the sentence: (5)

The quick brown fox jumps over the lazy dog."

Perform Part of Speech POS tagging for each word.

(ii) Given the sentence:

“Steve Jobs founded Apple in California.”

Identify and classify named entities in this sentence.

17

[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 7293

L

Unique Paper Code : 2343010030

Name of the Paper : Blockchain and its Applications

Name of the Course : **B.Sc. (H)**

Semester : VIII

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. The paper has **two** sections. **All** questions in “**Section A**” are compulsory.
3. Attempt any **four** questions from “**Section B**”.
4. Parts of a question must be answered together.

Section A

1. (a) How blockchain is helpful in e-commerce. (2)
(b) State the any two-scalability issue in blockchain. (2)

कालिन्दी महाविद्यालय पुस्तकालय
KALINDI COLLEGE LIBRARY

P.T.O.

- (c) What is a genesis block? Explain. (3)
 - (d) Explain the concept of “Immutability” in blockchain. (3)
 - (e) Describe Bitcoin address with the help of suitable example. (3)
 - (f) What is a digital wallet? Explain. (3)
 - (g) Compare Bitcoin System and Bank System. (3)
 - (h) List any three advantages of Cryptocurrency. (3)
 - (i) What is ‘Gas’ in the context of the Ethereum platform? (3)
 - (j) Describe the role of Solidity in blockchain with the help of an example. (5)
2. (a) Explain the concept of blockchain technology as a distributed ledger system. Compare blockchain with traditional databases. Highlight at least four differences. (10)
- (b) Describe the five key characteristics of blockchain. (5)

3. (a) Describe the importance of Merkle Trees in blockchain. Construct a Merkle Tree for 4 transactions T1, T2, T3 and T4. Explain its working. (8)
- (b) Illustrate the step-by-step working of a blockchain transaction from initiation to validation. (7)
4. (a) Describe the Bitcoin transaction lifecycle and analyze its advantages and limitations in modern financial systems. (10)
- (b) Explain the concept of mining and its importance. (5)
5. (a) Explain the architecture of a permissioned blockchain system and its use in enterprise applications. (5)
- (b) Explain SHA-256 hashing process in blockchain. (5)
- (c) Explain how transactions are validated and recorded in blockchain. (5)
6. (a) Explain the concept of Smart Contracts in blockchain. How do they work? Describe the advantages and limitations of Smart Contracts in blockchain. (10)

P.T.O.

- (b) Explain the role of Ethereum as a platform for smart contracts. (5)
7. (a) Analyze the applications of blockchain technology in the domains given bellow :
- Banking and Finance
 - Supply Chain Management (10)
- (b) Differentiate between public and private blockchain with help of suitable example. (5)

[This question paper contains 8 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 7970

L

Unique Paper Code : 2343010029

Name of the Paper : Natural Language Processing

Name of the Course : **B.Sc. (H)/B.A (P)/B.Sc. (P) –
DSE**

Semester : VIII

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** is compulsory.
3. Attempt any **FOUR** questions from **Section B**.
4. Parts of a question should be attempted together.

Section A

1. (a) What are stopwords? Why are stopwords removed in text processing? (3)

- (b) What is tokenization in natural language processing and why is it required? (3)
- (c) State the Markov assumption that underlies bigram language models and write the bigram probability formula it produces. Give one limitation of the bigram language model. (3)
- (d) Explain the Bag-of-Words (BoW) model and its limitations. How does Word2Vec address these limitations? (3)
- (e) Why are word embeddings important in NLP, and how do they differ from one-hot encoding? (3)
- (f) State one application for each of the three types of Transformer-based models. (3)
- (g) What is pooling in neural NLP models, and how do mean pooling and max pooling work? (3)

- (h) Explain perplexity as a metric for language model evaluation. If a bigram model assigns probability 0.001 to a test sentence of 10 words, compute its perplexity. (3)
- (i) Why is Minimum Edit Distance computed using dynamic programming instead of a brute-force approach, and what are the edit operations involved in calculating it? (3)
- (j) Define Term Frequency-Inverse Document Frequency (TF-IDF). Explain its significance in text representation. (3)

Section B

2. (a) Given the regular expression $r'\backslash b[A-Z][a-z]^+\backslash b'$ and the text *"Alice went to Paris in January"*.

I. List all the words matched by the pattern.

II. Explain the purpose of each part of the pattern $(\backslash b, [A-Z], [a-z]^+)$.

III. Modify the pattern so it also matches words written entirely in uppercase, and apply it to the same text. (5)

(b) Explain the differences between lemmatisation and stemming with one example. (5)

(c) Define Natural Language Processing (NLP). Explain any four real-world applications with suitable examples. (5)

3. (a) Perform two iterations of the Byte Pair Encoding (BPE) algorithm on the given corpus: 'l o w </w>' ($\times 5$), 'l o w e r </w>' ($\times 2$), 'n e w e s t </w>' ($\times 6$). Compute the frequency of all adjacent symbol pairs before each merge, identify the most frequent pair, update the corpus after each merge, and list the resulting vocabulary. (5)

- (b) Compute the Minimum Edit Distance (MED) between "Saturday" and "sunday" using dynamic programming with unit costs for insertion, deletion, and substitution. (7)

I. Write the recurrence relation used.

II. Construct the Dynamic Programming/ word edit distance matrix.

III. Show one optimal sequence of edit operations.

- (c) Explain the zero-probability problem in n-grams language models. Using Laplace (Add-1) smoothing, compute the probability $P(\text{cat}|\text{the})$ given $\text{Count}(\text{the cat})=0$, $\text{Count}(\text{the})=15$, and vocabulary size $V=500$. (3)

4. (a) What is Vector Semantics? Explain the intuition behind representing word meanings as vectors. What is the difference between static embeddings and contextual embeddings? (6)

- (b) Compute the TF, IDF, and TF-IDF score for the term “*cat*” in document D1 given the corpus: D1 = “*the cat sat*”, D2 = “*the cat ate*”, D3 = “*the dog sat*”.
- (c) Compute the cosine similarity between the vectors $A=[1,2,3]$ and $B=[2,1,2]$. Show the dot product and magnitude calculations, and interpret the result.
- (4)
5. (a) What is Part-of-Speech (POS) tagging? Perform for the sentence “*The quick brown fox jumps over the lazy dog.*” Assign an appropriate tag to each word.
- (5)
- (b) Apply NER tagging to the sentence “*Sundar Pichai joined Google in Mountain View on Monday.*” Identify all named entities and assign appropriate BIO tags along with their entity types.
- (5)

- (c) Explain the XOR problem in neural networks. Why is it not linearly separable, and how does a multi-layer perceptron overcome this limitation? (5)

6. (a) Write short notes on **Any two** of the following:

I. Feedforward Neural Networks

II. Recurrent Neural Networks

III. Activation Functions (6)

- (b) What is self-attention in Transformer models, and why is it important for language modeling? Also, explain how the Query, Key, and Value matrices are used to compute attention. (6)

- (c) What is prompting in large language models? Explain prompt engineering with one example. (3)

7. (a) Explain how text is generated in language models. Describe any one sampling methods used during generation. (4)

- (b) What are the major ethical issues in Large Language Models? Explain any four with real-world implications. (5)
- (c) Explain the three types of Transformer-based language model architectures with suitable diagram. (6)

[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 8647

L

Unique Paper Code : 2343010029

Name of the Paper : Natural Language Processing

Name of the Course : B.Sc. (H) / B.A. (P) / B.Sc. (P) – DSE

Semester : VIII

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** is compulsory.
3. Attempt any **FOUR** questions from **Section B**.
4. Parts of a question should be attempted together.

Section A

- Q1 a) What is a feedforward network? Name the three layers of the simple 2-layer feedforward network. (3)
- b) What is the purpose of the softmax function? Also, give its formula. (3)
- c) Given the bigram count $C(\text{like}, \text{NLP}) = 3$ and unigram count $C(\text{like}) = 6$ and size of the vocabulary $|V| = 10$. Compute $P(\text{NLP} | \text{like})$ using Laplace smoothing. (3)
- d) What is part-of-speech tagging? Perform part-of-speech tagging for the following sentence: "*The cat chased the mouse*". (3)

- e) What do you understand by query, key and value with respect to attention in Transformer? (3)
- f) List all trigrams from: "I am studying Python." (3)
- g) Explain how word2vec embeddings are different from count-based embeddings. (3)
- h) What are activation functions? Explain their purpose. Name *any two* activation functions. (3)
- i) Briefly explain the *three* stages of training a language model. (3)
- j) Briefly explain *any three* applications of Natural Language Processing. (3)

Section B

- Q2 a) Write regular expressions to extract: (5)
- I. Indian mobile numbers,
 - II. dates in DD/MM/YYYY format,
 - III. and email addresses.
- Also give one example for each.
- b) Explain stemming and lemmatisation. Justify which technique is better with suitable examples. (5)
- c) Explain the NLP pipeline. List and briefly describe *any four* stages involved in processing text. (5)

- Q3 a) Given the corpus: “I love NLP”, “I love AI”, and “I enjoy NLP”, find the unigram and bigram counts. Then calculate $P(\text{love} \mid I)$, $P(\text{NLP} \mid \text{love})$, and $P(I \text{ love NLP})$ using a bigram model. (5)
- b) Explain BPE (Byte Pair Encoding) subword tokenisation. Using the given training data: (5)
- l o w </w> (x6), l o w e r </w> (x4), n e w e s t </w> (x3)
- Write the initial vocabulary.
 - Perform two merge steps and show the updated vocabulary.
 - Explain how the trained BPE tokeniser handles a new word like “lowest”.
- c) Define Minimum Edit Distance (MED) between two strings. What are the three editing operations used? Explain the Levenshtein distance metric and give an example. (5)
- Q4 a) Given 3 documents: (5)
- D1: “cat sat on mat”
 - D2: “cat sat”
 - D3: “dog barked”
- Find the TF-IDF of the term “cat” in D1. What does the resultant value signify?
- b) What is the utility of word2Vec embeddings in NLP? Mention any one advantage and any one limitation of this approach. (5)
- c) Find the cosine similarity between two vectors (5)
- $A=[1,2,3]$ and $B=[2,3,4]$

- Q5 a) How is ambiguity resolved using POS tagging? Give an example to justify your answer. (5)
- b) Identify named entities in the given sentence "*Narendra Modi visited Delhi on 15th August.*" List any two tasks where named entity recognition (NER) is useful. (5)
- c) Justify why the attention mechanism is a core component of the Transformer architecture used in Natural Language Processing. (5)
- Q6 a) Describe the architecture of a feedforward neural language model. (4)
- b) Distinguish between the "*pooling*" and "*concatenation*" methods used for passing embeddings in language models. (5)
- c) Describe the steps involved in computing the attention score in the Transformer model. (6)
- Q7 a) Discuss *any five* major ethical and safety risks associated with the deployment of Large Language Models. (5)
- b) Define a prompt in the context of NLP. What is prompt engineering? Give two examples. (5)
- c) Explain the Transformer decoder architecture and describe how each of its component functions. (5)

[This question paper contains 8 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 2347

L

Unique Paper Code : 2342571201

Name of the Paper : Data Structures

**Name of the Course : B.A. (P) With CS as Minor/
Non-Major, B.Sc (Phy. Sc.)/
SOL**

Semester : II

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** is compulsory.
3. Attempt any **four** questions from **Section B**.
4. Parts of a question must be answered together.

**कालिन्दी महाविद्यालय पुस्तकालय
KALINDI COLLEGE LIBRARY**

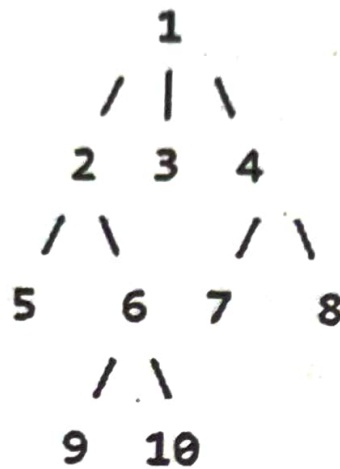
P.T.O.

Section A

1. (a) Define Big omega notation. (2)
- (b) Mention any two applications of Queue. (2)
- (c) Arrange the following Asymptotic Functions in increasing order: (2)
 $n!$, 2^n , n^3 , $n \log n$, $\log^2 n$
- (d) Convert the following infix expression into postfix expression. Show all the intermediate steps : (3)
 $(A - B / C) * (D + E)$
- (e) Depict the scenario when a recursive function has no base condition? (3)
- (f) Differentiate between Binary Heap and Binary Search Tree in term of complexity. (3)
- (g) Which pointer updates are required when inserting a node at the beginning of a Singly Linked List? Write the process step by step to insert a new node at the end of Singly Linked List. (4)

(h) Consider the following tree :

(5)



Based on the above tree, answer the following :

- (i) Identify the siblings of node 6.
 - (ii) List all ancestors of node 9.
 - (iii) List all descendants of node 2.
 - (iv) Find the degree of root node.
 - (v) Find the height of the tree.
- (i) Compare the time complexity of searching an element in :

(5)

- (i) Array
- (ii) Linked List
- (iii) Binary Search Tree
- (iv) Balanced Binary Search Tree

Which data structure is most efficient for searching and why?

Section B

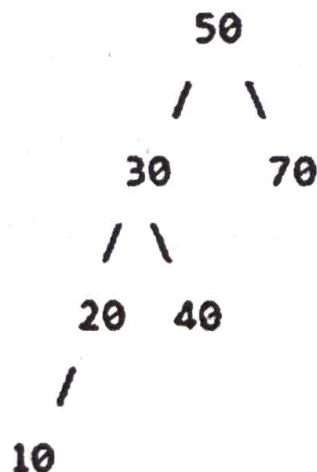
2. (a) Solve the following recurrence relation using Master's Theorem (5)

$$T(n) = 8T(n/2) + n^2.$$

- (b) Explain the binary recursion with the help of a suitable example. (5)

- (c) Explain the significance of balancing a Tree. Is the following binary tree an AVL Tree? (5)

If not, identify the node(s) violating AVL property:



3. (a) Compare Singly Linked List and Doubly Linked List with the help of suitable diagrams. (5)
- (b) Write a recursive code to compute the factorial of a number. Trace the execution for $n=5$. (5)

- (c) Evaluate the following prefix expression using stack. Show the stack status after each step:

$$* + 4 5 2 / 8 4 \quad (5)$$

4. (a) Determine the number of nodes in a full binary tree of height 4. Also calculate the number of leaf nodes. (5)

- (b) Show that $4n^3 + 3n^2 = O(n^3)$ and find the value of c and n_0 . (5)

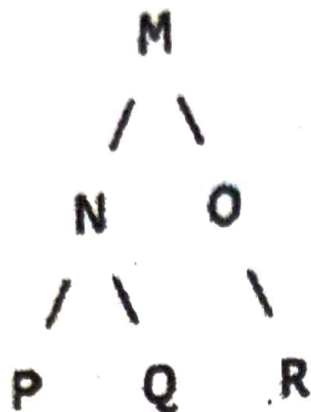
- (c) Find the output of the following recursive function for fun (4) after execution : (5)

```
int fun(int n)
{
    if(n==0)
        return 0;
    return n + fun(n-1);
}
```

5. (a) A circular queue of size 5 is initially empty. Perform the following operations and show Front/Rear after each step : (5)

ENQUEUE(11), ENQUEUE(22), ENQUEUE(33),
DEQUEUE(), ENQUEUE(44), ENQUEUE(55)

- (b) Express the following Binary Tree using Preorder, Inorder, and Postorder traversal : (5)



- (c) A stack is implemented with a maximum size of 5. Perform the following operations: (5)

push(15), push(13), pop(), push(17), push(72),
pop(), push(12), size(), peek()

Explain what happens at each step, the final contents of the stack, and how overflow or underflow conditions are handled.

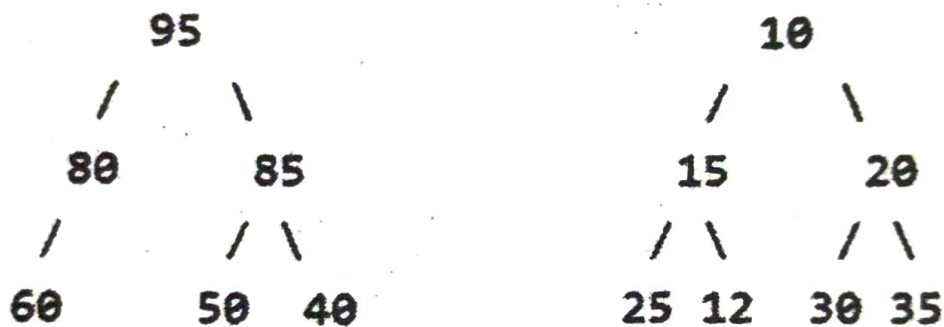
6. (a) Perform the following operations on an Array of size 5. (5)

- (i) Insert the elements: 40, 25, 30, 60, 15.
- (ii) Delete an element 30 from the list.
- (iii) Display the current list of elements after the insertions and deletions.

(b) (i) Write the steps to delete the Root node in a Max Heap.

(ii) Two Binary Trees are given below. Determine whether each tree is a Binary Heap or not.

Justify your answer with the help of suitable reason. (5)



(c) Construct the binary tree from the following given Inorder and Postorder traversal. (5)

Inorder : 4 2 5 1 6 3 7

Postorder : 4 5 2 6 7 3 1

(a) Define height of a binary tree. Write a C++ function to compute it. (5)

- (b) Insert the following elements one by one into an empty Max Heap and show the heap after each insertion :

25, 40, 15, 60, 10, 50 (5)

- (c) Construct a Binary Search Tree using the following data elements.

45, 25, 65, 15, 35, 55, 75, 10, 20

Delete node 25 and show the final BST. (5)

/(This question paper contains 8 printed pages.)

Your Roll No.....

Sr. No. of Question Paper : 2663

L

Unique Paper Code : 2342571201

Name of the Paper : Data Structures

**Name of the Course : B.A. (P) With CS as Minor/
Non-Major, B.Sc. (Phy. Sc.)/
SOL**

Semester : II

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** is compulsory.
3. Attempt any **four** questions from **Section B**.
4. Parts of a question must be answered together.

कालिन्दी महाविद्यालय पुस्तकालय
KALINDI COLLEGE LIBRARY

P.T.O.

Section A

1. (a) Find the Asymptotic notation (Big O) for the following nested loop: (2)

```
for (int i = 0; i < m; i++) {  
    for (int j = 0; j < n; j++) {  
        int a=5;  
    }  
}
```

- (b) Write any two real-life applications of queues.

(2)

- (c) Is an array that is in sorted order a Min-Heap?

Explain using a suitable example. (3)

- (d) What is the role of Stack in the implementation of recursion with the help of a suitable example?

(3)

(e) Differentiate between Linear and Non-Linear data structures. (3)

(f) Draw a Binary Search Tree using the following key values : (3)

18, 9, 24, 23, 16, 17.

(g) Given the values of $a=3$, $b=2$, $c=4$, $d=2$; evaluate the following given postfix expression :

$$a \ b \ * \ c \ d \ / \ + \ a \ d \ ^ \ -$$

Show the stack representation for the calculations. (4)

(h) Write a C++ program to insert an element at the front of a Singly Linked List. (5)

(i) Perform Insertion sort on the array [8, 2, 1, 9, 3], and show the steps after each iteration. Also, report the number of comparisons. (5)

Section B

2. (a) Write a program for generating first 10 Fibonacci numbers using a recursive function. (5)
- (b) Write a program to implement a queue using arrays. Write code for insert and delete operations. (5)
- (c) Solve the following recurrence relation using Master's Theorem: (5)

$$T(n) = 3T(n/2) + n$$

3. (a) A binary tree T has 9 nodes. The Inorder and Postorder traversals yield the following: (5)

Inorder : E A C K F H D B G

Postorder : E C K A H B G D F

Draw the Binary Tree and write the Preorder Traversal.

(b) Write a code snippet to delete a node from a Circular Linked List. Explain the steps using a suitable example. (5)

(c) What is an AVL tree? Why is the AVL tree called a self-balancing tree? Explain LL rotation used for balancing. (5)

4. (a) Construct a max heap using the following array:

[5, 3, 17, 10, 84, 19, 6, 22, 9]

(i) Build a heap using the bottom-up method

(ii) Delete root and reheapify Show all the steps. (5)

(b) Determine the number of nodes in a full Binary Tree of height 4. Calculate the number of leaf nodes. (5)

(c) Define Binary Recursion. Explain with the help of an example and report the number of recursive calls made. (5)

5. (a) Insert the following elements in an AVL tree:

10, 20, 30, 25

Show step-by-step insertion and imbalance, and also perform the required rotation. (5)

- (b) What is a Deque? Write a code snippet to delete an element from the front and rear of the Deque. (5)

- (c) Write short notes on : (5)

(i) Binary Tree

(ii) Big(O)

6. (a) Define a Doubly Linked List. Explain its structure with a neat diagram. Also mention two advantages and disadvantages. (5)

- (b) Perform the following stack operations (size = 5):

push(10), push(20), pop(), push(5), push(15),
push(25), pop(), push(30)

Show the stack contents after each operation.

(5)

(c) Show that : (5)

$7n^3+10n+50=0$ (n^3) and find the value of c and n_0 .

7. (a) Given the following Binary Search Tree: (5)

50, 30, 70, 20, 40, 60

Perform the following deletion operations :

(i) Node with one child (70)

(ii) Node with two children (50)

Explain each case with the help of diagrams.

(b) Explain the time complexity of the following basic operations in an array : (5)

(i) Access

(ii) Search

(iii) Insertion

(iv) Deletion

List one advantage and one disadvantage of using arrays.

- (c) Write a program to implement a Stack using a Linked List. Write the code for push and pop operations. (5)

[This question paper contains 8 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 2934

L

Unique Paper Code : 2342571201

Name of the Paper : Data Structures

**Name of the Course : B.A. (P) With CS as Minor/
Non-Major, B.Sc (Phy. Sc.)/
SOL**

Semester : II

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** is compulsory.
3. Attempt any **four** questions from **Section B**.
4. Parts of a question must be answered together.

**कालिन्दी महाविद्यालय पुस्तकालय
KALINDI COLLEGE LIBRARY**

P.T.O.

Section A

1. (a) Differentiate between an Array and a Linked List. (2)

- (b) Show that the Run Time Complexity of the following function: (3)

$$3n\log_{10}n + 4n + 5\log_{10}n$$

is $\theta(n\log_{10}n)$

- (c) Sort the following given array: (3)

$A = [6, 4, 12, 0, 1, 7, 9, 4]$ using Insertion Sort.

Show the steps after each iteration.

- (d) Write a recursive function in C++ to find the factorial of a number. (3)

- (e) Explain the significance of balancing the Binary Search Tree. Explain with the help of a suitable example. (3)

- (f) Write a function in C++ to insert an element at the end of a Doubly Linked List. (3)
- (g) Suppose a Circular Queue of capacity n elements is implemented with an array of n elements. Assume that the insertion and deletion operations are carried out using FIRST and LAST as array index variables respectively. Initially $FIRST = LAST = -1$. Write the conditions to detect if the queue is full and empty. (3)
- (h) What are the different operations that can be performed on a Deque. Explain using a suitable example. (5)
- (i) Describe the output of the following series of Stack operations:

push(7), push(2), pop(), push(5), pop(), top(), pop(), pop(), push(8), pop().

Show the contents of the Stack after each operation. (5)

Section B

2. (a) Write a code snippet for Linked List representation of Stack. Write the necessary class definitions and define the following functions: Push(), Pop(), Display(), Empty(), Full(). (6)
- (b) Write the function to reverse the elements of an Array. (4)
- (c) Solve the following recurrence relation using Master Theorem: (5)
- $$T(n) = 2T(n/2) + n \log_2 n$$
3. (a) Differentiate between a Linked List and Circular Linked List with the help of a suitable example. Also, write a function to add a node at the end of a Circular Linked List. (6)
- (b) Draw all possible Binary Search Trees for the three elements A, B and C. (4)

- (c) Calculate the Run Time Complexity for the following given recurrence relation using the recursion tree method : (5)

$$T(n) = 2T(n/2) + cn$$

4. Consider the following sequence of elements: (15)

P, H, J, I, C, F, G, L, T, S, W, Y

- (i) Construct the Binary Search Tree by inserting the elements one at a time in the order given above.
- (ii) Write the Inorder traversal of the constructed tree.
- (iii) Give the Height of the above tree. Also write its Postorder traversal.
- (iv) Delete node with value 'H' from the tree and draw the resultant tree.

- (v) Is the resultant tree a Complete Binary Tree?
Justify your answer.

5. (a) Write a C++ program to create a class DList implementing a Doubly Linked List with appropriate data members. Declare and define the following functions :

(i) Insert values in the list.

(ii) Delete a value from the i^{th} position in the list. (6)

(b) Write a function to search an element x in a Singly Linked List. (4)

(c) Consider the following arithmetic expression:

$$[P/Q - (X * Y) + (R \% S)]$$

Use stack to check if the above expression is valid i.e. it contains nested parenthesis and brackets. Show the contents of the stack at each step. (5)

6. (a) Give the recursive functions for Inorder and Preorder traversal in a Binary search Tree.

(6)

- (b) Give the time complexity of the following basic operations in a Queue implemented using Arrays:

(i) Enqueue()

(ii) Dequeue()

(iii) Empty()

(iv) Full() (4)

- (c) What is an AVL tree? Why is the AVL tree called a self-balancing tree? Explain LR rotation used for balancing. (5)

7. (a) Explain the concept of Tail recursion and write a Tail recursive function. (6)

- (b) Write a function to find the length of the linked list. (4)

(c) Construct a max heap for the given array A of elements,

$$A = [2, 4, 8, 1, 12, 11, 15, 6, 3].$$

Also, report the number of comparisons. (5)

[This question paper contains 2 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 2388
Unique Paper Code : 2342572401
Name of the Paper : Operating Systems
Name of the Course : B.Sc. (Programme)
Semester : IV
Duration : 3 Hours

L

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.

SECTION A

- Q1a. Explain Operating System as Resource Manager. 3
b. Discuss reasons for process termination. 3
c. Why is dual-mode operation essential for OS security? 2
d. Why is context switching considered overhead? 2
e. How does multithreading overcome the limitations of single-threaded execution? 2
f. Can deadlock occur with only one process? Explain logically. 2
g. What is the role of the Memory Management Unit (MMU)? 2
h. How does swapping improve CPU utilization? 2
i. Name system call that can be used by a parent process to determine the termination of child process? 3
j. Under what circumstances does page fault occur? 3
k. Differentiate between: 6
(i) multi-tasking and multi-user system
(ii) user mode and kernel mode
(iii) time sharing and multiprogramming operating system

SECTION B

- Q2 a. Explain the structure of a disk with the help of a diagram. 4
b. Explain the contiguous memory allocation and non-contiguous memory allocation with suitable diagram. Why is contiguous allocation faster than non-contiguous allocation? 5
c. Disk requests come to the disk driver for cylinders 4 15, 25, 10, 6, 45, 6 15, 37 in that order. 6
A seek operation takes 10 milliseconds to move the disk arm to a cylinder.
How much seek time is needed for :
(1) First Come First Serve
(2) SSTF
(3) Elevator algorithm. (Moving upward)
Note: In all the cases the arm is initially at cylinder 25
- Q3 a. What is Dead lock? Explain four necessary conditions for dead lock to occur with suitable example. Describe the different methods for 10

prevention and avoidance of dead lock. Consider a system having two processes and three resources. Each process needs a maximum of two resources. Draw resource allocation graph for the same. Is a deadlock possible?

- b. What do you understand by demand paging? Explain its working. 5
- Q4 a. Apply first-fit, best-fit, worst-fit for: 6
 Memory blocks: 100KB, 500KB, 200KB, 300KB, 600KB
 Processes: 212KB, 417KB, 112KB, 426KB
 Calculate remaining memory after allocation using each strategy.
- b. Explain File system in Operating System. State and explain any two file allocation methods. 6
- c. Consider a machine with 64 MB physical memory and a 32bit virtual address space. If the page size is 4 KB, what is the approximate size of the page table? 3
- Q5 a. Consider the following: 9
- | Process | Arrival Time (ms) | Burst Time (ms) |
|---------|-------------------|-----------------|
| P1 | 0 | 10 |
| P2 | 1 | 4 |
| P3 | 2 | 5 |
- Find the average waiting time and the turn-around time for the following process scheduling algorithms:
- FCFS
 - SJF
 - Round Robin. (Time Quantum = 5 ms)
- b. (i) Explain the concept of paging with a Translation Lookaside Buffer (TLB) and illustrate it with a suitable diagram. 6
- Q6 a. Consider the following page reference string how many page faults will occur using LRU and optimal page replacement scheme. Compare the results. 1, 2, 3, 4, 5, 3, 4, 1, 6, 7, 8, 7, 8, 9, 7, 8, 9, 5, 4, 5, 4, 2. The number of page frames = 3. 6
- b. Describe the critical section problem with suitable example. What are the requirements of solution the critical-section problem? 5
- c. Explain the performance criteria for process scheduling algorithms. 4
- Q7 a. Explain the Layered System Structure of Operating System, with one advantage and one disadvantage. 4
- b. Explain why context switching is necessary in a multitasking system. 3
- c. Write short notes on the following (any four): 8
- Process State
 - Process Control Block
 - Swapping
 - Dispatch latency
 - File operations

[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 2433

L

Unique Paper Code : 2342572401

Name of the Paper : Operating Systems

Name of the Course : B.Sc. (P) Computer Science

Semester : IV – NEP

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. The first question in Section A is compulsory.
3. Attempt any four questions from Section B.
4. Parts of a question must be answered together.

SECTION A

1. (a) How does the distinction between kernel mode and user mode function as a rudimentary form of protection (security) system? (3)
- (b) List the three components of an I/O subsystem of an operating system. (3)
- (c) What is the purpose of the command interpreter? Why is it usually separate from the kernel? (3)
- (d) Which of the following instructions should be privileged? Justify.
i. Set value of timer. ii. Rename a file. iii. Issue a trap instruction. (3)
- (e) Describe three general methods for passing parameters to the operating system. (3)
- (f) What is the purpose of system calls? Explain with a suitable example. (3)
- (g) What is the main advantage of the microkernel approach to system design? What are the disadvantages of using the microkernel approach? (3)
- (h) Provide three examples in which multithreading provides better performance than a single-threaded solution. (3)
- (j) Differentiate between the following:
a) Thread and Process
b) Pre-emptive and non-pre-emptive Kernel
c) Message passing mode and Shared memory model (2 x 3)

SECTION B

2. (a) Suppose that a disk drive has 5,000 cylinders, numbered 0 to 4,999. The drive is currently serving a request at cylinder 2,150, and the previous request was at cylinder 1805. The queue of pending requests, in FIFO order, is:
2069, 1212, 2296, 2800, 544, 1618, 356, 1523, 4965, 3681
Starting from the current head position, what is the total distance (in cylinders) that the disk arm moves to satisfy all the pending requests, for each of the following disk-scheduling algorithms-
- (i) SSTF (6)
 - (ii) C-SCAN (Direction is towards lower track numbers) (6)
- (b) Assume that you have a page-reference string for a process with m frames (initially all empty). The page-reference string has length p , and n distinct page numbers occur in it. Answer these questions for any page-replacement algorithms:
- (i) What is a lower bound on the number of page faults? (5)
 - (ii) What is an upper bound on the number of page faults? (5)
- (c) What resources are used when a thread is created? How do they differ from those used when a process is created? (4)
3. (a) Compare the memory organization schemes of contiguous memory allocation, pure segmentation, and pure paging with respect to the following issues:
- (i) External fragmentation (4)
 - (ii) Internal fragmentation (4)
- (b) A computer system uses paging for memory management with the logical address space as 32 KB, page size as 1 KB and time to access memory as 100 ns. Also, the page table is stored in memory. Answer the following:
- (i) How many bits are there in the logical address space?
 - (ii) How many bits are required for the page number and page offset?
 - (iii) What is the effective memory access time (EMAT) for a paged system without a TLB? (8)
 - (iv) If a TLB is introduced with a hit ratio of 80% and TLB access time of 20 ns, calculate the new EMAT. (8)
- (c) Differentiate between paging and segmentation. (3)

4. (a) A solution to the critical-section problem must satisfy three requirements. Explain each of them with example. (7)

(b) Illustrate how the memory management unit implements dynamic relocation using a relocation register. (8)

5. (a) Illustrate the physical structure of a magnetic disk with a neat diagram. (7)

(b) Differentiate between data parallelism and task parallelism with a suitable example. (8)

6. (a) Consider the following page reference string:
1, 2, 3, 4, 2, 1, 5, 6, 2, 1, 2, 3, 7, 6, 3, 2, 1, 2, 3, 6.
How many page faults would occur for the following replacement algorithms, assuming four frames? Remember that all frames are initially empty, so your first unique pages will cost one fault each.

(i) LRU Replacement

(ii) FIFO Replacement

(6)

(b) Consider the following set of processes with the length of the CPU burst given in milliseconds, along with their arrival time and priority:

Process	Burst Time
P1	9
P2	6
P3	5
P4	3
P5	5

The processes are assumed to have arrived in the order P1, P2, P3, P4, P5 all at time 0.

(i) Draw Gantt charts that illustrate the execution of these processes using the following scheduling algorithms: pre-emptive SJF and FCFS.

(ii) What is the turnaround time of each process for each of the scheduling algorithms?

(iii) Which of the algorithms results in the minimum average waiting time (overall processes)? (4+2+3)

7. (a) Write short notes on any three of the following:

- (i) Sequential file access method
- (ii) Two-level directory structure
- (iii) File system mounting
- (iv) File attribute

(6)

(b) Elaborate briefly on each of the three methods for handling deadlocks. List any three necessary conditions for a deadlock to occur.

(6+3)

(300)

[This question paper contains 4 printed pages.]

Your Roll No.....

L

Sr. No. of Question Paper : 2625

Unique Paper Code : 2342202403

Name of the Paper : Artificial Intelligence

Name of the Course : B.A. (P)

Semester : IV

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** is compulsory.
3. Attempt **any 4** questions from **Section B**.
4. Parts of the question must be answered together.

Section A

- Q1. a) Define the term Rational Agent. 2
- b) Write the condition for admissible heuristic of A* algorithm. 2
- c) Explain using a diagram how an agent interacts with the environment through sensors and actuators? 4
- d) Determine whether the sentence is valid, contradictory or satisfiable. 4
- $(P \rightarrow Q) \vee (Q \rightarrow R)$
- e) Write the PEAS for an automated taxi driver. 4
- f) Write a Prolog program to reverse a list L. 4

g) Give the conceptual dependency representation for the following:

4

Seema gave a book to Raghav

h) State the difference between:

6

- I. Model-Based Agent and Goal-Based Agent
- II. Weak AI and Strong AI
- III. Fully Observable and Partially Observable Environment

Section B

Q2 a) Given the following grammar:

4

1. $S \rightarrow NP VP$
2. $NP \rightarrow Det N \mid NP PP$
3. $VP \rightarrow V NP \mid V NP PP$
4. $PP \rightarrow P NP$
5. $Det \rightarrow The \mid a$
6. $N \rightarrow girl \mid cat \mid mat$
7. $V \rightarrow placed$
8. $P \rightarrow on$

Derive the parse tree for the sentence "The girl placed the cat on a mat" using above grammar.

b) Write a script for going to a movie.

5

c) Compare and Contrast between Breadth First Search and Best First Search.

6

Q3. a) What is cut and fail facility in Prolog? Explain using an example.

4

b) Explain the Hill climbing search. What are its limitations?

5

c) Transform the following term into Clausal form

6

$$\exists x \forall y \forall z (P(f(x), y, z) \rightarrow (\exists u Q(x, u) \wedge \exists v R(y, v)))$$

Q4. a) Which of the following is not logically equivalent and why?

4

i) $((p \wedge q) \vee (\neg p \wedge \neg q))$ and $p \leftrightarrow q$

ii) $((p \wedge q) \rightarrow r)$ and $((p \rightarrow r) \wedge (q \rightarrow r))$

b) Draw an associative network for the following sentence:

5

A university called XYZ University has three departments: Computer Science, Electronics, and Mechanical Engineering. Dr. Bajaj is the Head of the Computer Science department. Ram and Neha are faculty members in the Computer Science department. Neha is married to Rohan, who works as a journalist for The Times of India. Neha and Rohan have two children and live on Park Road. Neha wears spectacles and is five feet six inches tall. Represent the network using labelled directed edges.

c) a) Transform the following FOPL into the corresponding CNF:

6

i. $(\neg A \wedge B) \vee (A \wedge \neg B) \wedge C$

ii. $P \vee (\neg P \wedge Q \wedge R)$

iii. $(\neg P \vee Q) \rightarrow (R \wedge \neg S)$

Q5. a) What is a Recursive Transition Network? Explain using a suitable example.

4

b) Create a frame for terrestrial motor vehicles (cars, trucks, motorcycles) and give one complete frame in detail for cars which includes the slots for the main component parts, their attributes, and relations between parts.

5

c) List the four types in the Chomsky Hierarchy of grammars and briefly describe each one.

6

Q6. a) Express the following sentences as FOPL:

4

I. All Romans were either loyal to Caesar or hated him.

II. Some students studying Artificial Intelligence are genius and hardworking.

- b) What is Modus Ponens? Give an example. Give two limitations of Propositional logic. 5
- c) Solve the cryptarithmic problem using : 6

$$\begin{array}{r} \text{YOUR} \\ + \text{YOU} \\ \hline \text{HEART} \end{array}$$

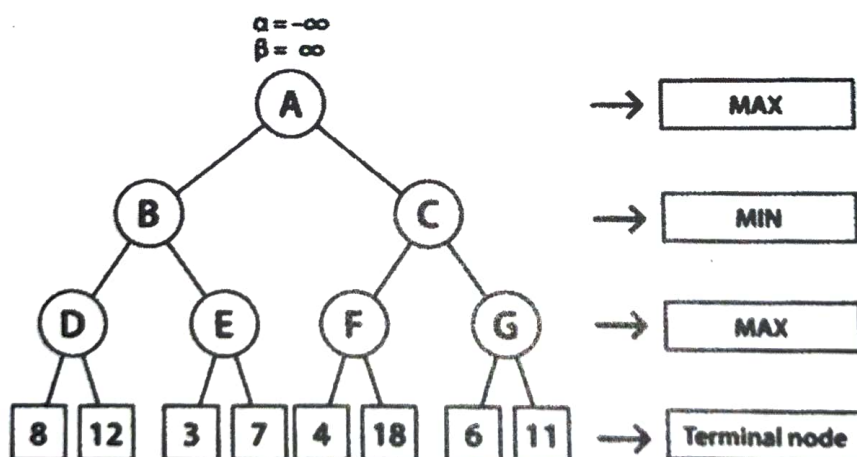
Q7 a) Define the following: 4

- I. Fairness in AI
- II. Biases in AI

- b) Define unification. Find whether the following sets are unifiable or not? If they are unifiable, find most general unifier (m.g.u.) else give justification why they are not unifiable. 5

- I. $\{S(x, \text{Hari}), S(y, \text{Geeta})\}$
- II. $\{P(x, y), P(f(x), z), P(z, x)\}$

- c) Consider the following game tree and find the value of the maximum at the root node. Which nodes will be pruned according to the alpha-beta pruning procedure? 6



[This question paper contains 2 printed pages.]

Your Roll No.....
L

Sr. No. of Question Paper : 2940
Unique Paper Code : 2342572401/232202401
Name of the Paper : Operating Systems
Name of the Course : B.A. Prog. DSC
Semester : IV
Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** is compulsory. Attempt any **FOUR** questions from **Section B**. Parts of a question should be attempted together.

Section A

1.
 - (a) What is the convoy effect. Give an example. (3)
 - (b) What is virtual memory. Describe its purpose. (3)
 - (c) Write any three major functions of an operating system with regard to memory management? (3)
 - (d) Write the actions taken by a kernel to context-switch between processes. (3)
 - (e) What is many-to-many multithreading model. Draw the diagram also. (3)
 - (f) Name any two operations which can be performed on a directory. List any one disadvantage of Two-level directory structure. (3)
 - (g) List any three advantages of multiprocessor systems. (3)
 - (h) Consider a logical address space of 256 pages with a 4-KB page size, mapped onto a physical memory of 64 frames:
 - i. How many bits are required in logical address? (3)
 - ii. How many bits are required in physical address? (3)
 - (i) Draw the resource allocation graph for the given scenario as follows: (3)
P1 is waiting for R1 and holding R2
P2 is holding R1
P3 is holding R1 and waiting for R2
P4 is holding R2.
State whether a deadlock exists or not with justification.
 - (j) Distinguish between absolute and relative path name (3)

Section B

2.
 - (a) What is demand paging in an operating system. Describe its advantages and disadvantages. (5)
 - (b) A system has a main memory access time of 100 ns, Translation Look aside Buffer (TLB) access time of 20 ns, and a TLB hit ratio of 95%. Calculate effective memory access time with TLB and without TLB (5)
 - (c) What is the layered approach of operating system design. Write its two advantages and two disadvantages. (5)
3.
 - (a) Distinguish between the following: (5)
 - i. Multiprogramming and Multitasking
 - ii. Ready queue and Job queue
 - (b) What is internal fragmentation. Which memory management technique used to address this issue. (5)

- (c) A memory system has free blocks of sizes: 350 KB, 150 KB, 500 KB, and 250 KB. If processes arrive in the order: P1: 200 KB, P2: 120 KB, P3: 450 KB, P4: 180 KB. (5)
- Allocate the memory to the processes using First-Fit and Best-Fit memory management algorithms.
 - Which of the above algorithm results in minimum external fragmentation? Justify your answer briefly.
4. (a) A disk scheduling system has the following queue of pending requests (in order of arrival): [98, 183, 37, 122, 14, 124, 65, 67]. The disk head starts at position 51. The disk tracks range from 0 to 199. Compute the total head movement using the SCAN algorithm (head moving towards higher-numbered tracks first). (5)
- (b) Define time quantum in Round Robin scheduling. How very small and very large values of time quantum affect system performance. (5)
- (c) Explain dual mode of operation and its purpose. (5)
5. (a) Define the following terms: (10)
- Random Access Time in magnetic disk
 - System Program
 - Long-term scheduler
 - Segmentation
- (b) Describe the various operations that can be performed on a file in an operating system. (5)
6. (a) Consider the following set of processes with their arrival times and CPU burst times (in milliseconds): (10)
- | Process ID | Arrival Time | Burst Time |
|------------|--------------|------------|
| P1 | 0 | 7 |
| P2 | 2 | 4 |
| P3 | 4 | 1 |
| P4 | 5 | 4 |
- Using the Shortest Remaining Time-First (SRTF) and First Come First Serve (FCFS) scheduling algorithms:
- Draw the Gantt chart showing the execution order of processes.
 - Calculate the turnaround time and waiting time for each process.
 - Compute the average waiting time and average turnaround time.
- (b) What is deadlock in process management? Write briefly the necessary conditions of deadlock. (5)
7. (a) Consider the following page reference string: 7, 0, 1, 2, 0, 3, 0, 4, 2, 3, 0, 3, 2. (10)
- Assuming demand paging with three frames:
- Compute the page faults using the following page replacement algorithms:
 - First in first out (FIFO)
 - Least Recently Used (LRU)
 - Compare the results and identify which algorithm performs better for the given reference string.
- (b) What is Open-Source Software? Write an advantage and a disadvantage of an Open-Source Operating System? (5)

[This question paper contains 8 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 8645

L

Unique Paper Code : 2343010023

Name of the Paper : Reinforcement Learning

Name of the Course : DSE

Semester : IV/VI/VIII

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Question No. 1 (Section A) is compulsory.
3. Attempt any 4 (four) questions from Section B.
4. Parts of a question should be attempted together.
5. Use of scientific calculator is allowed.

Section A

- Q1 (a) Describe the four main components of a reinforcement learning system. (4)
Explain each briefly.
- (b) Differentiate between continuing tasks and episodic tasks. Explain their key (4)
characteristics with suitable examples.
- (c) What is policy gradient method? Under what circumstances, and for what (4)
reasons, are actor-critic methods preferred over them?

- (d) Describe the concept of bootstrapping in TD learning. How does the TD(0) update make use of it? Also define the TD error δ with its formula and explain the meaning of a positive δ . (4)
- (e) What is an ϵ -greedy policy? Why is it preferred in Monte Carlo Control? Also, specify the probabilities of selecting a greedy action and a non-greedy action under this policy. (4)
- (f) Define the value function $v_\pi(s)$ and the action-value function $q_\pi(s, a)$. Write the Bellman equation for $v_\pi(s)$ and briefly explain the intuition behind it — how does it relate the value of a state to the values of its possible successor states? (5)
- (g) Consider the following scenarios: (5)
- i. Episodic Task:
An agent receives rewards +2, +3, +5 over three time steps and then reaches a terminal state. The discount factor $\gamma = 1$.
Compute the total return G .
 - ii. Continuing Task:
An agent receives a constant reward of +2 at each time step indefinitely. The discount factor $\gamma = 0.9$.
Compute the total return G .

Section B

- Q2 (a) Consider a Markov Decision Process (MDP) with two states S_1 and S_2 (8)
and a discount factor $\gamma=0.9$. The transition and reward structure
are provided in the table below:

State	Action	Next State	Reward
S_1	A1	S_1	5
S_1	A2	S_2	10
S_2	B1	S_1	-1
S_2	B2	S_2	2

Using the Value Iteration algorithm, perform one iteration starting with

initial values $V_0(S_1) = 0$ and $V_0(S_2) = 0$.

Compute:

- $V_1(S_1)$
 - $V_2(S_2)$
- (b) Explain the difference between ordinary importance sampling and weighted (7)
importance sampling for off-policy MC estimation. Which one is unbiased?
Which typically has lower variance? State one advantage of each.

- Q3 (a) Consider an agent following a fixed policy π in an episodic task with (8)
discount factor $\gamma=1$ and $V(A)=0$. The observed episode is given
below:

Time (t)	State	Reward
0	A	0
1	B	2
2	A	3
3	Terminal	--

Using both the First-Visit Monte Carlo and Every-Visit Monte Carlo methods, compute the value of state $V(A)$. Show all steps involved in the calculation.

- (b) In Monte Carlo control, why must we estimate action-value functions $Q(s, a)$ rather than state-value functions $V(s)$ for model-free settings? Also, explain the difference between first-visit and every-visit Monte Carlo prediction — how does each method compute $V(s)$? (7)

- Q4 (a) Consider a short corridor problem with switched actions consisting of 3 non-terminal states $S1, S2, S3$ and a terminal goal state G . The agent always starts from $S1$. The reward is -1 per step, and reaching G ends the episode. (8)

Given Action dynamics:

In $S1$ and $S3$:

Action Right moves right

Action Left moves left (no movement if already at boundary)

In $S2$: actions are switched

Right moves left

Left moves right

A policy selects Right with probability p and Left with probability $1-p$ at every state.

Assume:

Discount factor $\gamma=1$

From experiments, the expected number of steps to reach G is:

$$T(p=0.4) = 25$$

$$T(p=0.6) = 18$$

$$T(p=0.8) = 30$$

Perform the following tasks: -

- i. Create a suitable diagram for the above problem.
 - ii. Compute the expected return for each value of p .
 - iii. Identify which policy (value of p) is best.
- (b) Explain stochastic-gradient descent (SGD) and semi-gradient methods in the context of function approximation highlighting how they are used to update value function parameters. (7)

Q5 (a) Compare Q-learning and SARSA in terms of how they handle exploration and the nature of the policies they learn. Also write their update rules. (8)

(b) Consider an episode: (7)

- States: $S_1 \rightarrow S_2 \rightarrow \text{Terminal}$
- Rewards: $R_1=4, R_2=6$
- Initial values: $V(S_1)=1, V(S_2)=2$
- $\alpha=0.5, \gamma=1$

- i. Update $V(S_1)$ using **TD(λ) with $\lambda = 0$** (i.e., TD(0)).

- ii. Update $V(S_1)$ using TD(λ) with $\lambda = 1$ (i.e., TD(1) / Monte Carlo).
- iii. Compare the results

Q6 (a) Differentiate between Monte Carlo (MC), Dynamic Programming (DP), and Temporal Difference (TD) methods highlighting their key features. Also compare their approaches to learn value functions. (8)

(b) An agent is navigating an environment. Using the given experience sequence and parameters, compute the updated action-value for the state-action pair (S_0, A_0) using 2-step SARSA. (7)

The experience sequence:

$t = 0: (S_0, A_0) \rightarrow R_1 = 2 \rightarrow (S_1, A_1)$

$t = 1: (S_1, A_1) \rightarrow R_2 = 4 \rightarrow (S_2, A_2)$

$t = 2$: The agent reaches (S_2, A_2)

Parameters: $n = 2$, learning rate $(\alpha) = 0.5$, discount factor $(\gamma) = 0.9$, current estimates $Q(S_0, A_0) = 5$, $Q(S_2, A_2) = 10$.

Q7 (a) A mobile robot collects soda cans in an office. It has two battery states: $\{S\} = \{\text{high}, \text{low}\}$. The robot must choose actions to maximize its expected reward (cans collected). (9)

System Dynamics:

State $s = \{\text{high}\}$

- Search: Battery stays high with probability α and drops to low with probability $1 - \alpha$.
- Wait: Battery always stays high.

State $s = \{\text{low}\}$

- Search: Battery stays low with probability β .
- With probability $1 - \beta$, the battery is depleted, the robot must be rescued, and it is reset to high.
- Wait: Battery always stays low.
- Recharge: Battery is restored to high with 100% certainty.

Reward Structure:

- Expected reward for searching is r_{search}
 - Expected reward for waiting is r_{wait} where $r_{\text{search}} > r_{\text{wait}}$
 - If the robot's battery is depleted during a search (requiring rescue), it receives a penalty of -3.
 - Recharging yields 0 reward.
- i. Construct the transition probability table $p(s' | s, a)$ for all valid state-action pairs. Compute the probabilities for:
 - $(s=\text{high}, a=\text{search})$
 - $(s=\text{low}, a=\text{search})$
 - $(s=\text{low}, a=\text{recharge})$
 - ii. Compute the expected reward $r(s, a, s')$ for the following specific transitions:
 - Searching from low and staying low.
 - Searching from low and being rescued (transitioning to high).

- (b) In asynchronous dynamic programming, states can be updated in any order (6) rather than in systematic sweeps. Give two advantages of asynchronous DP over synchronous DP. Explain why asynchronous DP is still guaranteed to converge.

[This question paper contains 4 printed pages]

Your Roll No. :

Sl. No. of Q. Paper : **275** **I**

Unique Paper Code : 2343010020

Name of the Paper : Research Methodology

Name of the Course : **B.A(Prog)/B.Sc(Prog)**

Semester : VI

Time : 3 Hours

Maximum Marks : 90

Instructions for Candidates :

- (a) Write your Roll No. on the top immediately on receipt of this question paper.
- (b) All parts of Question 1 (**Section-A**) are compulsory
- (c) Attempt any **four** questions from **Section-B**.
- (d) Parts of a question must be answered together
- (e) Use of scientific calculator is allowed

Section - A

1. (a) Explain any three tools used for research briefly. 3
- (b) What are academic search engines? List any three academic search engines for scientific research. 3
- (c) What is the purpose of including endnotes and footnotes in a research report? 3
- (d) What is a DOI? How are they better than URLs? 3

P.T.O.

कालिन्दी महाविद्यालय पुस्तकालय
KALINDI COLLEGE LIBRARY

- (e) Explain the difference between plagiarism and piracy. 3
- (f) Explain copyrights with respect to IPR. What happens when a copyright expires? 3
- (g) Differentiate between Type I and Type II error in Hypothesis testing 3
- (h) Explain the concept of 'voluntary participation' with respect to ethics in research 3
- (i) Categorise the following research as basic or applied research and explain your answer. 3
 - (i) Development of a new drug for diabetes
 - (ii) Study on how stress affects memory in brain
 - (iii) Designing a stress management program for students before exams
- (j) Given the z-score value of 3.4, calculate IQ and stanine score. 3

Section - B

- 2. (a) What is the peer review process? What are the type of questions that a referee is expected to answer while reviewing a paper? 7.5
- (b) Explain GI tag with the help of an example. Who governs GI Tags in India? Can it be owned by an individual? What is the validity of GI Tags? 7.5
- 3. (a) What are the key practical considerations that must be evaluated for analysing whether a research idea is feasible? Briefly explain any three. 7.5

- (b) Journal of Advanced Data Analytics has 200 articles published in 2022, 180 articles in 2023 and 220 articles in 2024. In 2025, 300 citations were obtained for articles published in 2022, 320 citations for articles in 2023 and 310 citations in 2024. Calculate the impact factor and cite score for 2025

7.5

4. (a) Classify the following three cases as either Direct, accidental, mosaic or self plagiarism. Justify your answer.

7.5

Case A: Ravi is writing his thesis on sustainable agriculture. He comes across a paragraph in a journal that perfectly summarizes his point. He copies it word for word into his dissertation without citing it.

Case B: Priya submitted a paper on consumer buying behaviour last semester. This semester, she submits a new assignment with large sections of the same paper including the literature review and methodology.

Case C: Arjun is writing a report on food security. He picks specific phrases from three different sources, weaves them together with minor word changes, and presents it as a new paragraph.

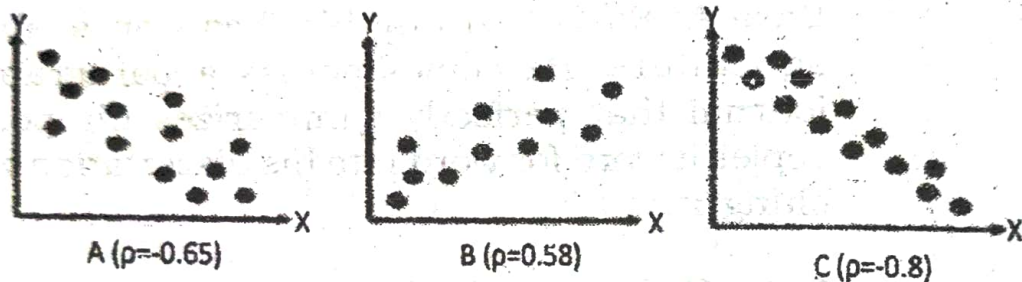
- (b) A company claims that the average weight of its product is 500 grams. A sample of 36 items has a mean weight of 490 grams with a standard deviation of 30 grams.

7.5

- (i) State the null and alternative hypotheses.
(ii) Calculate the test statistic.

(iii) Based on your result, explain whether the sample provides enough evidence to reject the company's claim at a 5% significance level.

5. (a) Explain placebo effect with respect to medication. Explain any two ethical issues while administering a placebo? 7.5
- (b) Categorise the following relationship of X and Y as strong/weak positive/negative correlation. How does a change in X influence the change in Y for each case? 7.5



6. (a) Height of 5 students are as follows- 130, 129, 127, 103, 110. Calculate the normalised heights using z-score normalization. 7.5
- (b) Explain the seven steps of the cyclic nature of research. 7.5
7. (a) Write the six essential elements of a research report. 7.5
- (b) Differentiate between the following with suitable examples 7.5
- (i) Descriptive vs Analytical research
 - (ii) Quantitative vs Qualitative research

[This question paper contains 2 printed pages.]

Your Roll No.....

L

Sr. No. of Question Paper : 2234

Unique Paper Code : 2342573601

Name of the Paper : Computer Networks

Name of the Course : **B.A./B.Sc. (Prog)**

Semester : VI

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** is compulsory.
3. Attempt **any 4** questions from **Section B**.
4. Parts of a question must be answered together.

Section A

1. a) Distinguish between IPv4 and IPv6 protocol. 4
b) Assume that a voice channel occupies a bandwidth of 4 kHz. We need to multiplex 10 voice channels with guard bands of 500 Hz using FDM. Calculate the required bandwidth. 5
c) What is the remainder obtained by dividing $x^7 + x^5 + 1$ by the generator polynomial $x^3 + 1$? Show the binary steps and the transmitted data. 5
d) Frames of 1000 bits are sent over a 1 Mbps channel using a geostationary satellite whose propagation time from the earth is 270 msec. Acknowledgements are always piggybacked into data frames. The headers are very short. Three-bit sequence numbers are used. What is the maximum achievable channel utilization for: Stop-and-wait, Go-back-N and Selective repeat algorithms. 6
e) Describe different parts of a URL. What is the role of HTTP protocol and at which layer does it work? What are the steps to be followed to fetch a web page from a server? 5
f) Describe the responsibilities of physical layer in TCP/IP model? Which of the model among TCP/IP and OSI models is used on Internet these days? 5

Section B

2. a) Compare and contrast TCP and UDP (User Datagram Protocol) in terms of their features, advantages and disadvantages. Discuss scenarios where each protocol would be suitable for data transmission. 7
b) Write short notes on: 8
(i) File Transfer Protocol
(ii) Simple Mail Transfer Protocol

3. a) What do you understand by Multiplexing? Two channels, one with a bit rate of 190 kbps and another with bit rate of 180 kbps are to be multiplexed using pulse stuffing TDM with no synchronization bits. What is the size of the frame in bits? What is the frame size and data rate? 5
- b) (i) What is transmission impairment? Explain its types. 4+6
(ii) Differentiate between wired and wireless media. Should we use optical fibre for data transmission?
4. a) Write a short note on the Internet Protocol (IP) that works at the network layer of the TCP/IP model. Discuss how packets are routed to the destination network. 5
- b) Suppose that a message 1100 1001 0011 1010 is transmitted using Internet Checksum (4-bit word). What is the value of the checksum? What kind of errors will not be detected by Checksum? 5
- c) Describe the job of a router and how it is different from switches? How is point to point connection different from a multipoint connection? Give an example of each. 5
5. a) Explain stop and wait protocol for noise free channels. What are the issues of stop and wait protocol? 6
- b) Differentiate the following: 9
(i) Flow control and Error control
(ii) HUB and Gateway
(iii) Bus Topology and Mesh Topology
6. a) (i) Differentiate between LAN, MAN and WAN. 3+3
(ii) Explain simplex data flow. How is it different from half duplex communication?
- b) A 12-bit odd-parity Hamming code whose binary value is 111001001111 arrives at the receiver. Find the original value of the message. Assume that not more than 1-bit is in the error. 4+5
7. a) Write any two network commands: 5
(i) Netstat (ii) Ping (iii) Tracert
- b) Explain the layers in which following tasks are performed in the OSI reference model. 10
(i) Physical addressing (iv) Session Management
(ii) Logical addressing (v) Encryption/Decryption
(iii) Routing

[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 2273

L

Unique Paper Code : 2342573601

Name of the Paper : COMPUTER NETWORKS

Name of the Course : B.A/B.SC. (PROG)

Semester : VI

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Section A is compulsory.
3. Attempt any 4 questions from Section B.
4. Parts of a question must be answered together.

Section A

1. (a) Define Computer Networks. explain any two objectives of computer network. 3
- (b) Find out the following: 3
 - (i) Class of IP address: 213.10.45.21
 - (ii) Network ID of IP address: 168.168.168.168
 - (iii) Default Subnet Mask of IP address: 131.132.133.134
- (c) Calculate the transmission delay for a 1000-byte packet over a 1 mbps link. 3
- (d) If a subnet mask is 255.255.255.192, how many subnets and hosts per subnet are possible? 3
- (e) Discuss the difference between Layer-II and Layer-III Switch. 3
- (f) Differentiate between connection less and connection-oriented protocols with the help of examples. 3

- (g) Name the Protocols to be used in the following scenarios: 3
- i. To map an IP address to a MAC address.
 - ii. To assign the IP address to the host automatically.
 - iii. To map a MAC address to an IP address.
- (h) Explain the modes of communication with appropriate examples. 3
- (i) What are the purposes of the FIN, RST, and URG flags in the TCP header. 3
- (j) Discuss how both switches and bridges handle frame forwarding and MAC address learning. 3

Section B

2. (a) Given the IP address 10.0.0.0/24, an organization needs to create two subnets with the following requirements: 8
- Subnet 1: 100
 - Subnet 2: 50
- For each of these subnets, give the
- i. First IP address and last IP address assigned
 - ii. Subnet mask in the w.x.y.z/s notation
 - iii. Identify the range of IP addresses available
 - iv. Number of host and network bits
- (b) Write any four differences between Star Topology and Mesh Topology. Also, find the number of links required to establish a mesh topology with n devices. (4+3)
3. (a) Differentiate among Stop-and-Wait, Go-Back-N and Selective Repeat protocols. 8
- (b) The received hamming code is 1011010: 7
- i. Detect if any error exists.
 - ii. Find the error position.
 - iii. Correct the code.

4. (a) Explain the functions of each layer of OSI model. Also, compare it with TCP/IP model. 8
- (b) (i) Differentiate between IPv4 and IPv6. (4+3)
- (ii) Explain CIDR notation and its advantages.
5. (a) Consider a network with 4 nodes: A, B, C, and D. the link costs are as follows: 8
 $A \rightarrow B: 2$, $A \rightarrow C: 5$, $B \rightarrow C: 2$, $B \rightarrow D: 4$, $C \rightarrow D: 1$
- i. Show the final converged routing table for Node A.
- ii. Show the final converged routing table for Node C.
- (b) A 1500-byte packet is sent over a 10 Mbps link with a distance of 1000 km and 7
propagation speed 2.5×10^8 m/s. Find:
- i. Transmission delay
- ii. Propagation delay
6. (a) Consider a network comprising five routers labelled as A, B, C, D, and E. The 8
interconnections and their associated costs are given as follows:
- Link between A and B: Cost 2
 - Link between A and C: Cost 4
 - Link between B and C: Cost 1
 - Link between B and D: Cost 5
 - Link between C and D: Cost 3
 - Link between C and E: Cost 7
 - Link between D and E: Cost 2

Using Dijkstra's algorithm, illustrate the step-by-step process of determining the shortest path from router A to router E.

- (b) What is the CSMA/CD Protocol? What is the minimum frame size required to ensure collision detection in the CSMA/CD protocol? 7
7. (a) What are the key characteristics of infrared waves and microwaves? 3
- (b) Write short notes on the following terms: 2*6=12
- (i) WWW
 - (ii) Jitter
 - (iii) SMTP
 - (iv) Subnetting
 - (v) NIC
 - (vi) DNS

[This question paper contains 2 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 2306
Unique Paper Code : 2342203602
Name of the Paper : Deep Learning
Name of the Course : B.A. (P) DSC / SOL
Semester : VI
Duration : 3 Hours

L

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Question No. 1 (Section A) is compulsory.
3. Attempt any 4 (four) questions from Section B.
4. Parts of a question should be attempted together.
5. Use of scientific calculator is allowed.

Section A

- Q1 (a) What are the activation functions? Compare Sigmoid and Tanh function. (3)
- (b) Explain the parameter norms penalties in regularization? Why are the weights penalized but biases usually left unregularized. (3)
- (c) What do you mean by L1 regularization? How does it differ from L2 regularization in terms of sparsity. (3)
- (d) Consider a 1-D input signal $x = [2, 1, 0]$ and a kernel $w = [1, -1]$. Compute the discrete convolution $s(t) = (x * w)(t)$. (3)
- (e) What do you mean by parameter sharing in convolutional neural networks? How does it improve efficiency compared to fully connected layers. (3)
- (f) Consider the recurrent relations (3)
 $s(t) = s(t - 1) + 2$.
With initial state $s(0) = 1$.
Compute $s(1)$, $s(2)$ and $s(3)$.
- (g) Describe Bidirectional RNN. How does it differ from standard RNN. (3)
Give one application of Bidirectional-RNN.
- (h) What do you mean by undercomplete autoencoder? What loss function does it optimize? (3)
- (i) Explain computational graphs and chain rule used in backpropagation. (3)
- (j) What is the objective of speech recognition? Give one example of traditional model used for speech recognition and one deep learning model that has improved its performances. (3)

Section B

- Q2 (a) What are Rectified Linear Units (ReLU)? Explain why ReLU is easy to optimize and discuss its variants. (8)

- (b) Explain how the recursive application of the chain rule is used to compute gradients in deep feedforward networks. Why naive implementation becomes infeasible, and how back-propagation overcomes this challenge to achieve efficient gradient computation? (7)
- Q3 (a) Explain how the AdaGrad optimizer adapts learning rates for different parameters during training. Discuss its key limitations in the deep learning models. How does RMSProp address these limitations when dealing with non-convex optimization problems? (8)
- (b) What is Batch Normalization? Explain how it works during training and testing? and discuss its role in optimizing deep neural networks. (7)
- Q4 (a) Explain the role of convolution in feature extraction. (3)
- (b) An input grayscale image of dimension 3×3 is given as follows: (6)
- $$I = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 2 \\ 0 & 1 & 1 \end{bmatrix}$$
- Perform the convolution operation using the following 2×2 kernel with stride = 1 and no padding:
- $$K = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$
- (i) Compute the output feature map.
- (ii) Find the size of the output feature map.
- (c) Explain the concept of pooling in convolutional neural networks. Explain how pooling provides translation invariance? (6)
- Q5 (a) Explain the architecture of Long Short-Term Memory (LSTM) networks. Describe the role of the self-loop and the gating units in controlling information flow. Why are LSTMs more effective than simple recurrent networks for learning long-term dependencies? (8)
- (b) What is a sparse autoencoder? Give one advantage of sparse autoencoder over standard autoencoder. Explain how sparse autoencoders learn useful features and how the sparsity penalty is applied? (7)
- Q6 (a) Explain the architecture and learning process of Generative Adversarial Networks (GANs). Describe the roles of the generator and discriminator. What is a major challenge in a GAN training? (8)
- (b) Explain the architecture of a recurrent neural network (RNN) with suitable diagram. Justify the statement- "RNNs are powerful but computationally expensive to train". (7)
- Q.7 (a) What is the role of preprocessing in deep learning for computer vision? Explain with examples of contrast normalization and dataset augmentation (7)
- (b) What are neural language models (NLMs)? How do distributed word representations help overcome the curse of dimensionality in language modelling? (4)
- (c) Explain DropOut and DropConnect as regularization techniques. (4)

[This question paper contains 4 printed pages]

Your Roll No. :

Sl. No. of Q. Paper : 2490 I

Unique Paper Code : 2343010020

Name of the Paper : Research Methodology

Name of the Course : B.A(Prog)/B.Sc(Prog)

Semester : VI

Time : 3 Hours

Maximum Marks : 90

Instructions for Candidates :

- (a) Write your Roll No. on the top immediately on receipt of this question paper.
- (b) **All** parts of Question 1 (**Section-A**) are compulsory.
- (c) Attempt any **four** questions from **Section-B**.
- (d) Parts of a question must be answered together.
- (e) Use of scientific calculator is allowed.

Section - A

1. (a) Why are objectivity and systematic approach important characteristics of research? 3
- (b) What do you mean by problem identification in research? 3

P.T.O.

कालिन्दी महाविद्यालय पुस्तकालय
KALINDI COLLEGE LIBRARY

- (c) What is the purpose of conducting a literature review? 3
- (d) Explain the three measures of central tendency. 3
- (e) Why is interpretation of data important in research? 3
- (f) How are conclusions drawn from research findings? 3
- (g) Explain the role of descriptive statistics in data analysis. 3
- (h) What do you mean by peer review in academic publishing? Describe the role of reviewers in the peer review process? 3
- (i) Define citation index and its importance. 3
- (j) What is a conflict of interest in research? How can a personal conflict of interest affect ethical publication practices? 3

Section - B

2. (a) Differentiate between the following types of research with suitable examples 7.5
- (i) Experimental vs Non-Experimental Research
 - (ii) Prospective vs Retrospective Research

- (b) A Student has scored the following marks in different tests 10 subjects. 7.5

78 89 67 76 57 88 92 83 85 72

Calculate the mean, median, mode and inter quartile range of the student's marks.

3. (a) A journal had published 120 articles in 2019, 130 articles in 2020 and 150 articles in 2021. What will be the cite score and impact factor if journal receive 2200 citations in 2022? 7.5
- (b) Discuss the complete research process in detail, explaining each step and its importance. 7.5
4. (a) Researcher has published 8 papers and citations count for each paper are: 7.5
20, 16, 15, 14, 12, 19, 17, 18
Calculate the h-index, g-index, and i10-index for this researcher.
- (b) Explain standard error of mean and estimated standard error of mean. A factory produces metal rods. A quality control team is analysing the length of rods, which is known to have a population standard deviation of 0.05 cm. They take a random sample of 36 rods to estimate the mean length. Calculate the standard error of mean for this sample. 7.5

5. (a) Describe the steps involved in hypothesis testing with the help of an example. 7.5
(b) What is plagiarism? Mention any two plagiarism checker tools and explain their purpose. 7.5
6. (a) Explain fraud and misconduct in research. Explain its types, effects on society, and ways to prevent it. 7.5
(b) Explain copyright and royalty. What type of works are protected under copyright? What happens when a copyright expires? 7.5
7. Explain the placebo effect with suitable examples. How is it used in experimental research? 7.5
(b) Explain any three intellectual property rights (IPR) in research. How do they protect the work of researchers? 7.5

[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 2916

L

Unique Paper Code : 2342573601

Name of the Paper : Computer Networks

Name of the Course : B.A. / B.Sc. (Prog)

Semester : VI

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Section A is compulsory.
3. Attempt any 4 questions from Section B.
4. Parts of a question must be answered together.

Section A

- 1 (a) Write the notion of transmission impairment. Briefly discuss the difference between distortion and attenuation. (1+2=3)
- (b) Consider a noiseless channel with a bandwidth of 4000 Hz that transmits a signal with 8 signal levels. Calculate the maximum bit rate of the channel. (3)
- (c) Identify the OSI layer responsible for each of the following operations: (3)
 - (i) Ensures reliable end-to-end message delivery between processes.
 - (ii) Responsible for moving packets from the source to the destination and enabling internetworking between networks.
 - (iii) Responsible for transmitting raw bits over a physical medium and defining the mechanical and electrical specifications of the interface.
- (d) What are the purposes of the URG, SYN, and PSH flags in a TCP header? (3)
- (e) Why is HTTP a stateless protocol? (3)
- (f) What are the periodic signals, and why are they commonly used in analog transmission? (3)
- (g) What is the full form of DNS? What is its use? (3)

- 4 (a) Give three differences between connectionless and connection-oriented services in networking. (3)
- (b) What is the HDLC protocol? Explain I-Frame, S-Frame and U-Frame. (2+3=5)
- (c) Consider a network comprising five routers labelled as A, B, C, D, and E. The interconnection and its associated costs are as follows:
- Link between A and B: Cost 2
 - Link between A and C: Cost 4
 - Link between B and C: Cost 1
 - Link between B and D: Cost 5
 - Link between C and D: Cost 3
 - Link between C and E: Cost 7
 - Link between D and E: Cost 2
- Using the Shortest Path Algorithm (Dijkstra's algorithm), illustrate the step-by-step process of determining the shortest path from router A to router E. (7)
- 5 (a) What does the following address mean, and when are they used? (3)
- (i) 0.0.0.0
 - (ii) 127.xx.yy.zz
 - (iii) 156.76.255.255
- (b) What do you mean by CSMA/CD protocol? Discuss the binary exponential back-off algorithm. (2+4=6)
- (c) For Class A, Class B, and Class C IP addresses, specify their address ranges and default subnet masks. (6)
- 6 (a) What is the role of the HTTP protocol in data communication? (3)
- (b) Write the purpose of using Guard Bands in a multiplexed channel, and explain with a suitable example. (2+2=4)
- (c) Describe the structure of an IPv4 address. Why is the IPv6 address structure better than IPv4? (5+3=8)

- 7 (a) What is wavelength division multiplexing? How does it differ from the time division multiplexing technique? (2+3=5)
- (b) Write short notes on the following terms: (2*5=10)
- (i) Switch
 - (ii) Router
 - (iii) HDLC
 - (iv) SMTP
 - (v) ICMP

[his question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 7290

L

Unique Paper Code : 2343010018

Name of the Paper : Ethical Hacking

Name of the Course : DSE

Semester : VI / VIII

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Question No. 1 (Section A) is compulsory.
3. Attempt any 4 (four) questions from Section B.
4. Parts of a question should be attempted together.

Section A

1. (a) How is penetration testing different from vulnerability assessment? (3)
(b) What are Google directives? List any two Google search operators useful in reconnaissance. (3)
(c) Differentiate between advanced persistent threat attacks and ransomware attacks. (3)

[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 7290

L

Unique Paper Code : 2343010018

Name of the Paper : Ethical Hacking

Name of the Course : DSE

Semester : VI / VIII

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Question No. 1 (Section A) is compulsory.
3. Attempt any 4 (four) questions from Section B.
4. Parts of a question should be attempted together.

Section A

1. (a) How is penetration testing different from vulnerability assessment? (3)
(b) What are Google directives? List any two Google search operators useful in reconnaissance. (3)
(c) Differentiate between advanced persistent threat attacks and ransomware attacks. (3)

- (d) How does the CIA triad help in securing information systems? (3)
- (e) Discuss the role of log files in system security. Why might they be deleted or cleared? (3)
- (f) Define MetaGooFil. What kind of data does it collect? (3)
- (g) What is Cross-Site Scripting (XSS)? How does it exploit web applications? (3)
- (h) Write any three differences between a virus and a worm? (3)
- (i) Define Alternate Data Streams (ADS). How are they used in file systems? (3)
- (j) Write any three differences between active scanning and passive scanning? (3)

Section B

- 2. (a) Explain any five stages of the Cyber Kill Chain. (5)
- (b) Discuss the key implementation guidelines for an effective threat and vulnerability management program in an organization. (5)

- (c) What are post-testing findings? Explain their significance in security assessment. (5)
3. (a) Differentiate between Active reconnaissance and Passive reconnaissance. (5)
- (b) Explain the role of HTTrack in website footprinting and analysis. (5)
- (c) Write short note on the following :
- (i) Nslookup
- (ii) Netstat (5)
4. (a) What is phishing? Write any three methods to prevent phishing attacks. (5)
- (b) What is the role of Nmap tool in network security assessment? Explain its any three features. (5)
- (c) What is insecure design in web applications? Explain its consequences. (5)
5. (a) Explain session hijacking with examples. Discuss the causes and impacts of session hijacking attacks. (5)
- (b) A login page accepts user input without validations. Explain how an attacker can exploit it using SQL Injection. (5)

- (c) What is Metasploit? Describe any three payloads available in Metasploit. (5)
6. (a) Explain any five preventive measures to protect a system from malware attacks. (5)
- (b) Explain any two password attack techniques with suitable examples. Suggest two preventive measures to protect systems against these attacks. (5)
- (c) How does port scanning and vulnerability scanning assist an attacker during the reconnaissance phase of penetration testing? (5)
7. (a) Define web server hacking. Identify any two common vulnerabilities associated with web servers. How can these vulnerabilities be mitigated? (5)
- (b) Discuss any three major causes of identity theft. Also, recommend appropriate preventive measures to avoid the risk of such attacks. (5)
- (c) Write short notes on the following : (5)
- (i) The Harvester Tool
- (ii) DDoS attack

[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 7967

L

Unique Paper Code : 2343010018

Name of the Paper : Ethical Hacking

Name of the Course : DSE

Semester : VI/VIII

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

छात्रों के लिए निर्देश.

1. Write your Roll No- on the top immediately on receipt of this question paper.
2. **Section A** is compulsory (Question 1).
3. Attempt any 4 questions from **Section B**.
4. Parts of a question should be attempted together.

Section A

- Q1 (a) What are ISO standards in cyber security? 3
- (b) Define Broken Access Control. List any two common causes of Broken Access Control. 3
- (c) What is the importance of reporting in cyber security? 3
- (d) What is the purpose of the fping command? What kind of information is obtained from its output? 3
- (e) A user downloads free software from an untrusted website. After installation of the software, the system became slow, and multiple unwanted pop-ups appeared. Identify the type of malware likely involved in this scenario and explain its impact on system performance and user security. 3

- (f) What is DNS enumeration? What kind of information can be extracted from DNS records? 3
- (g) Explain the role of Wireshark in troubleshooting networks. 3
- (h) Give any three disadvantages of clearing logs. 3
- (i) Write any three differences between web server hacking and web application hacking. 3
- (j) What is a code injection attack? Give one example of a code injection attack. 3

Section B

- Q2 (a) List and describe the five primary stages of penetration testing. 5
- (b) Explain any two vulnerabilities from the OWASP Top 10 with suitable examples. 5
- (c) A system administrator observes multiple consecutive failed login attempts followed by a successful login. Identify the password cracking technique that may have been used. Explain how this technique works and discuss possible measures to prevent such attacks. 5
- Q3 (a) Explain the concept of foot printing and reconnaissance in ethical hacking. 5
- (b) An attacker tries to find valid email addresses and mail server details of a company. Explain the technique that can be used for extracting information from email servers? 5
- (c) A user receives an email appearing to be from a bank, asking to verify account details through a link. After entering the credentials, the user's account is misused. Identify the attack technique and explain how it is carried out. How can such attacks be prevented? 5

- Q4 (a) What is the NIST Cybersecurity Framework? What are the three core functions of the NIST Cybersecurity Framework. 5
- (b) Differentiate between white box and black box penetration testing. 5
- (c) Explain the role of social engineering in ethical hacking. How does it help to identify security weaknesses? Also, suggest two preventive measures to reduce such attacks. 5
- Q5 (a) Differentiate between threat and risk? Describe the role of System and Services Acquisition (SA) in maintaining information security. 5
- (b) A security analyst needs to scan a network to identify open ports and services. Explain how Nmap can be used to perform this task, including TCP connect scan. 5
- (c) List any five common techniques of unauthorized system access? 5
- Q6 (a) Explain John the Ripper password cracking tool. How is the SAM file used to crack the password? Mention two ethical considerations associated with its usage. 5
- (b) Explain how DoS/DDoS attacks affect network availability, performance, and security. 5
- (c) List three types of non-adversarial threat sources and explain how they impact the CIA triad with suitable examples. 5
- Q7 (a) What is SQL Injection? A website stores user data in a database and does not sanitize inputs. Explain how SQL Injection can compromise the system. 5
- (b) Define Advanced Persistent Threat (APT). Explain its three characteristics and describe how it differs from other types of cyber-attacks. 5

7967

4

5

- (c) Write short notes on the following :
- (i) Ipconfig Command
 - (ii) HTTRACK Tool

[This question paper contains 8 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 111

L

Unique Paper Code : 2342574801

Name of the Paper : Information Security

Name of the Course : B.A. (P.)

Semester : VIII

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** (Question 1) is compulsory.
3. Attempt **any 4** questions from **Section B** (Questions 2 to 7).

Section A

1. (a) Is every threat necessarily an attack? Justify your answer. (3)

- (b) Consider the following code fragment: (3)

```
legitimate code
```

```
if (date == Friday the 13th);
```

```
    crash_computer();
```

```
legitimate code
```

What type of malware is this? Give justification for your answer.

- (c) What is the difference between a false positive and a false negative in the context of an intrusion detection system? (3)
- (d) What do you mean by a digital signature? Provide two examples of digital signature algorithms. (3)
- (e) What is meant by 'patching vulnerability' in the context of IoT devices? (3)
- (f) Define 'backscatter traffic' and explain its significance in detecting network attacks. (3)

- (g) Briefly explain the difference between in-band SQL injection and inferential SQL injection attacks. Illustrate each with an example. (3)
- (h) What are mutually exclusive role constraints and cardinality constraints in the RBAC₂ model? Explain with examples. (3)
- (i) What types of programming languages are vulnerable to buffer overflows?
- (j) Consider an Automated Teller Machine (ATM) system in which users provide a Personal Identification Number (PIN) and a card to access their accounts. Give one example each of confidentiality, integrity, and availability requirements associated with this system. (3)

Section B

2. (a) A university is launching an online portal where students can log in to access grades, submit assignments, and pay fees. To ensure that students are connecting to the genuine university server

and that their data remains secure, the IT team deploys digital certificates issued by a trusted authority.

With reference to the above scenario, explain what a Public Key Certificate is. Describe its main components and explain how the trusted authority helps in issuing and validating these certificates to ensure secure communication. (7.5)

(b) Briefly explain the different phases in the lifecycle of a computer virus. (7.5)

3. (a) Is symmetric encryption alone suitable for data authentication? If yes, please provide relevant justification. If no, explain the possible security threats associated with using symmetric encryption alone for authentication. Also, suggest a more secure cryptographic mechanism that can be used to ensure authentication. (7.5)

(b) A web application allows users to log in using a username and password. The following code is used to validate user credentials : (7.5)

```
query = "SELECT *
```

```
FROM users
```

```
WHERE username = ' " + input_username + "'
```

```
AND password = ' " + input_password + "' ; ;
```

A user enters the following input in the login form:

Username: admin' – –

Password: (left blank)

- (i) What will be the final SQL query executed by the system?
 - (ii) Identify the type of attack being performed.
 - (iii) Explain how the attack works in this case.
 - (iv) Suggest any one method to prevent such attacks.
4. (a) A financial web application allows users to input transaction details. During testing, it was observed that unexpected inputs (such as special characters and extremely large values) caused system crashes and inconsistent outputs. How can defensive

programming practices be applied to address these issues? Illustrate with suitable examples. (7.5)

(b) State and explain any five security threats to wireless networks. (7.5)

5. (a) A small business installs a basic firewall that only filters traffic based on IP addresses and ports but still faces security issues. Identify the type of firewall being used. What are its limitations? Suggest a more advanced firewall type and explain why it would be more effective. (7.5)

(b) Define access control policy. Discuss the four major types of access control policies used in computer security. Are these policies mutually exclusive? Justify your answer. (7.5)

6. (a) An organization is evaluating different cloud options based on cost, control, and security requirements. What are the four cloud deployment models? Explain how each model differs in terms of ownership and usage. (7.5)

(b) Differentiate between : (2.5×3)

- (i) Reactive and Proactive Password Checking
- (ii) Access Control List and Capability List
- (iii) Active Attacks and Passive Attacks

7. (a) An application stores user passwords in a database using a simple hashing technique : (7.5)

Password $\rightarrow$ Hash(password)

After a security breach, the developer updates the system to use salting:

Password $\rightarrow$ Hash(password + salt)

Each user is assigned a unique random salt value, which is stored along with the hash.

- (i) Why was storing passwords using only hashing considered insecure in this case?
- (ii) How does adding a salt improve password security?

- (iii) Suppose two users have the same password "abc@123". Explain how salting affects their stored hash values.
- (iv) Is it safe to store the salt along with the password hash? Justify your answer.
- (b) What are the differences among a backdoor, a trojan, a keylogger, spyware, and a rootkit? Can they all be present in the same malware? (7.5)

[This question paper contains 4 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 200

L

Unique Paper Code : 2343010023

Name of the Paper : Reinforcement Learning

Name of the Course : DSE

Semester : VIII

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** is compulsory.
3. Attempt **any 4** questions from **Section B**.
4. Parts of a question must be answered together.
5. Use of Scientific Calculator is allowed.

Section A

1. (a) Define a Markov Decision Process (MDP). What are its five components? (3)
(b) What is a Policy (π)? Distinguish between deterministic and stochastic policies. (3)
(c) Differentiate between the "Planning" Approach and the "Learning" Approach. If an agent has a perfect model of the environment, does it still need to interact with the environment to learn? (3)
(d) An agent receives a constant reward of +2 at each time step indefinitely. The discount factor $\gamma = 0.9$. Compute the total return G . (3)
(e) A state S is visited in three episodes with returns Episode 1: $G=8$; Episode 2: $G=4$ and Episode 3 : $G=10$. Using incremental MC updates with $\alpha = 0.1$ and $V_0(S) = 0$: (i) Compute $V_1(S)$ after episode 1. (ii) Compute $V_2(S)$ after episode 2. (iii) Compute $V_3(S)$ after episode 3. (3)
(f) Describe the Agent-Environment interface in a Reinforcement Learning Task. (3)
(g) A state S is visited in 2 episodes with returns: Episode 1: $G = 6$ and Episode 2: $G = 10$. Compute the Monte Carlo estimates of $V(S)$. (3)

- (h) Assume the followings are given: $Q(S_t, A_t)=3$, $R=1$, $Q(S_{t+1}, A_{t+1})=4$, $\gamma=0.8$, $\alpha=0.1$.
Using SARSA, compute updated Q . (3)
- (i) In the TD (λ) algorithm, what is the role of the parameter λ ? How does the algorithm behave when $\lambda = 0$ and when $\lambda = 1$? (3)
- (j) Define Q-Learning. Why is it called an off-policy method? (3)

Section B

- 2 (a) Differentiate between a Reward Signal and a Value Function in Reinforcement Learning. (3)
- (b) An agent uses ϵ -greedy policy with $\epsilon = 0.2$ and 4 actions. (4)
- (i) Calculate P (greedy action).
 - (ii) Calculate P (each non-greedy action).
 - (iii) If ϵ changes to 0.0, what is the agent's behavior?
 - (iv) Verify your probabilities sum to 1.
- (c) In Reinforcement Learning, an agent must balance Exploration and Exploitation to maximize its cumulative reward. Define these two concepts and explain the fundamental trade-off between them using a real-world example. Why is it impossible for an agent to achieve an optimal policy by focusing solely on one of these strategies? (8)
- 3 (a) Consider a Markov Decision Process (MDP) with two states S_1 and S_2 and a discount factor γ (gamma) = 0.8. The transition and reward structure is given in the table below: (8)

State	Action	Next State	Reward
S_1	A_1	S_1	5
S_1	A_2	S_2	10
S_2	B_1	S_1	-1

S2	B2	S2	2
----	----	----	---

Using the Value Iteration algorithm, perform two iterations starting with initial values $V_0(S_1) = 0$ and $V_0(S_2) = 0$.

Compute:

- (i) $V_1(S_1)$ and $V_1(S_2)$ — values after the first Iteration.
- (ii) $V_2(S_1)$ and $V_2(S_2)$ — values after the second Iteration.

- (b) An agent at state s takes action a (stochastic): with probability 0.75 it transitions to state s_e and receives reward $r = +20$; with probability 0.25 it transitions to state s_j and receives reward $r = -10$. (7)

Given: $\gamma = 0.7$, $V(s_e) = 50$, $V(s_j) = -20$. Calculate the action-value $q(s, a)$.

4. (a) The value function $V(s)$ is estimated as 100 based on 100 samples and the 101st sample is 201. Which formula is used to update the estimate incrementally? Apply the formula and compute the new value of $V(s)$. Show all steps. (5)

- (b) An agent receives: $R_1 = 2$, $R_2 = 5$, $R_3 = 3$, $R_4 = 8$, with $\gamma = 0.9$. (4)
(i) Compute G_3 . (ii) Compute G_2 . (iii) Compute G_1 . (iv) Compute G_0 .

- (c) An agent generates two episodes. Compute the value estimate $V(S_1)$ using First-visit Monte Carlo (MC) assuming a discount factor $\gamma = 1$. (6)
Episode 1: $(S_1, A_1, 2) \rightarrow (S_2, A_2, 3) \rightarrow (S_1, A_3, 5) \rightarrow \text{Terminated}$.
Episode 2: $(S_1, A_1, 1) \rightarrow (S_3, A_2, 9) \rightarrow \text{Terminated}$.

Would the value of $V(S_1)$ change if Every-visit Monte Carlo is used?

Show your calculation.

- 5 (a) If a robot's ability to move depends on its current position and how much battery it has left, but *not* on the path it took to get there, does this satisfy the Markov Property? Why is this assumption critical for the Bellman Equations to work? (5)
- (b) Explain $TD(\lambda)$. What is the λ -return? How does it combine all the n -step returns? (5)
- (c) An agent is in state S_t with $V(S_t) = 25$. It takes an action, receives a reward of -5, and transitions to S_{t+1} where $V(S_{t+1}) = 30$. If the learning rate is 0.2 and discount factor is 0.8, calculate the updated value $V(S_t)$. (5)

- 6 (a) Explain the difference between ordinary importance sampling and weighted importance sampling for off-policy MC estimation. Which one of them is unbiased? Which one typically has a lower variance? State one advantage of each. (10)
- (b) Define Policy Gradient Methods. Why is REINFORCE called a "Monte Carlo policy gradient"? (5)
- 7 (a) Compare SARSA and Q-Learning: write the update rule for each. Why is SARSA considered on-policy and Q-Learning considered off-policy? Explain the key difference in next-action selection. (7)
- (b) In asynchronous dynamic programming, states can be updated in any order rather than in systematic sweeps. Give two advantages of asynchronous DP over synchronous DP. Explain why asynchronous DP is still guaranteed to converge. (8)

[This question paper contains 8 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 7968

L

Unique Paper Code : 2343010023

Name of the Paper : Reinforcement Learning

Name of the Course : Bsc. Computer Science

Semester : VIII

Duration : 3 Hours

Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Section A is compulsory.
3. Attempt any 4 questions from Section B.
4. Parts of a question must be answered together.
5. Use of Scientific Calculator is allowed.

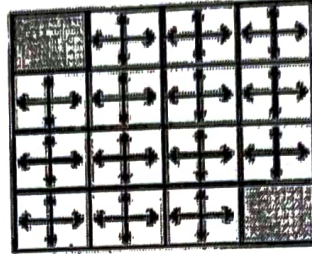
Section A

1. (a) What is Reinforcement Learning? How is it different from Supervised learning? (3)
- (b) Differentiate between a Reward Signal and Value Function. Which one of them represents the long-term potential of a state? (3)
- (c) A robot's next position depends only on its current coordinates and its current velocity. Does this system satisfy the Markov Property? Justify your answer. (3)
- (d) An agent receives a sequence of rewards $R_1 = 2$, $R_2 = 4$, $R_3 = 10$. If the discount factor $\gamma = 0.5$, calculate the total discounted return G_0 . (3)

7968

- (e) Consider a 4x4 grid world. The top-left and bottom-right are the terminal states, while the rest are non-terminal states. Initial state-values v are all 0s and corresponding greedy policy is shown using arrows for all directions achieving the maximum. (3)

0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0



After a series of policy iteration steps, the new state-values are shown as below. Show the greedy policy (using arrows) with respect to these given state values:

0.0	-1.7	-2.0	-2.0
-1.7	-2.0	-2.0	-2.0
-2.0	-2.0	-2.0	-1.7
-2.0	-2.0	-1.7	0.0

- (f) An agent is using ϵ -greedy policy with $\epsilon = 0.1$. If there are 5 possible actions, compute the probabilities of choosing the greedy (best-known) action and the non-greedy (random) actions. (3)
- (g) In the Actor-Critic algorithm, which component is responsible for the policy, and which is responsible for value function computation? (3)
- (h) Using a suitable example, describe the trade-off between exploration and exploitation. What are the consequences of favoring one over the other? (3)
- (i) In a TD(λ) algorithm, the eligibility trace is updated as: (3)

$$z_t = \gamma \lambda z_{t-1} + 1$$

Given $\gamma = 0.9$, initial trace $z_0 = 0$, compute z_1 and z_2 when $\lambda = 1$.

- (j) What is the difference between n-step TD and TD(λ) in terms of how they use future rewards? (3)

Section B

- 2 (a) Describe the Agent-Environment interaction in a Reinforcement Learning Task. Draw a diagram showing the interaction loop of S_t , A_t , and R_t and explain how the time step progresses from one interaction to the next. (5)

- (b) Consider a state s_0 with two possible actions a_1 and a_2 : (5)

- Action a_1 leads to state s_1 with reward = 4
- Action a_2 leads to state s_2 with reward = 2

The agent follows a policy:

$$\pi(a_1|s_0) = 0.6, \pi(a_2|s_0) = 0.4$$

Given $\gamma = 0.9$, $V(s_1) = 5$, $V(s_2) = 3$, compute $V(s)$ using the Bellman Expectation Equation.

- (c) Consider a Reinforcement learning agent being trained to play the game of Chess. Identify and briefly describe the following components of this RL system: Agent, Environment, State, Action, Reward. (5)

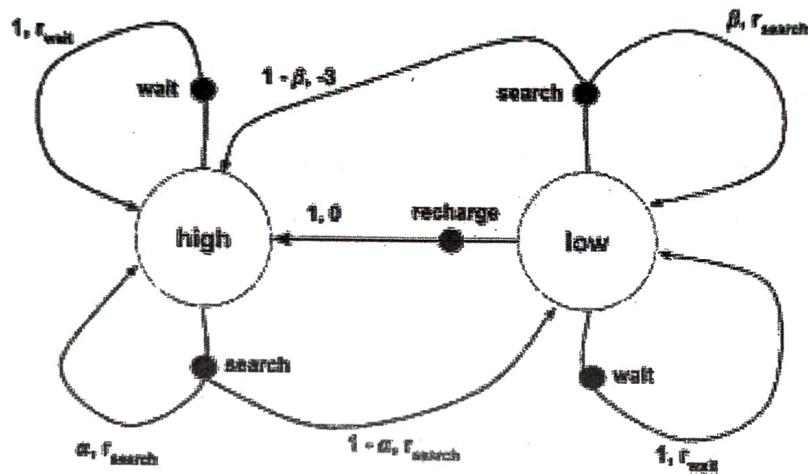
- 3 (a) Define Value Functions in Reinforcement Learning. Write the Bellman Equation for the state-value function $v_\pi(s)$. Identify and explain the meaning of each mathematical term used in the equation such as π , s , a , r , and γ . (5)

- (b) Explain the difference between Policy Iteration and Value Iteration algorithms. Describe the primary steps involved in each. Which one of the two is generally faster in terms of the number of iterations to converge? (5)

- (c) Consider a state s where taking action ' a ' leads to state s_1 with probability 0.7 (reward +5) and state s_2 with probability 0.3 (reward -1). If discount factor $\gamma = 0.9$, $v(s_1) = 10$, and $v(s_2) = 2$, calculate $q(s, a)$. (5)

4 (a)

(5)



Above diagram shows the transition graph for a finite MDP with following details:

States, $S = \{high, low\}$

Actions, $A = \{search, wait, recharge\}$

Starting in a state s and taking action a moves you along the line from state node s to action node (s, a) . Each arrow corresponds to a triple (s, s', a) , where s' is the next state and the arrow is labelled with transition probability, $p(s'|s, a)$, where s' is the next state and the expected reward for that transition $r(s, a, s')$.

For this finite MDP system, fill in the transition probabilities and the expected rewards in the table below:

s	a	s'	$p(s' s, a)$	$r(s, a, s')$
high	search	high		
high	search	low		
low	search	high		
low	search	low		
high	wait	high		
high	wait	low		
low	wait	high		
low	wait	low		
low	recharge	high		
low	recharge	low		

- (b) Explain the assumption of *exploring starts* in Monte Carlo estimation of Action values? Why is it difficult to satisfy in practice, and how does an ϵ -greedy policy address this issue? (5)
- (c) An agent generates two episodes. Compute the value estimate $V(S_1)$ using First-visit Monte Carlo (MC) assuming a discount factor $\gamma = 1$. (5)
- Episode 1: $(S_1, A_1, 2) \rightarrow (S_2, A_2, 3) \rightarrow (S_1, A_3, 5) \rightarrow \text{Terminated}$.
- Episode 2: $(S_1, A_1, 1) \rightarrow (S_3, A_2, 9) \rightarrow \text{Terminated}$.

- 5 (a) Compare SARSA and Q-learning in the context of temporal-difference learning. Write their update rules and justify why SARSA is classified as an on-policy method while Q-learning is considered off-policy. (5)
- (b) Discuss how n-step TD methods are a generalization of Reinforcement Learning algorithms. Show how by taking specific values of parameter 'n' allows these methods converge to either TD(0) or Monte-Carlo. (5)
- (c) An agent is in state S_t with $V(S_t) = 10$. It takes an action, receives a reward of 2, and transitions to S_{t+1} , where $V(S_{t+1}) = 12$. If the learning rate $\alpha = 0.5$ and discount factor $\gamma = 0.9$, calculate the updated value $V(S_t)$. (5)
- 6 (a) Why is it impractical to use a simple lookup table (tabular method) for complex games like Chess? Compare Function Approximation with Tabular methods in terms of memory constraints, and the ability to generalize to new situations. (5)
- (b) State the Policy Gradient Theorem. Why do policy gradient methods work well when actions are continuous or when policy needs to be stochastic? (5)
- (c) Consider a short corridor problem with switched actions consisting of 3 non-terminal states S_1, S_2, S_3 and a terminal goal state G . The agent always starts from S_1 . The reward is -1 per step, and reaching G ends the episode. (5)

Given Action dynamics:

In S_1 and S_3 :

Action Right moves right

Action Left moves left (no movement if already at boundary)

In S_2 : actions are switched

Right moves left

Left moves right

A policy selects Right with probability p and Left with probability $1-p$ at every state.

Assume:

Discount factor $\gamma=1$

From experiments, the expected number of steps to reach terminal Goal state, i.e. $T(p)$ is:

$$T(p = 0.4) = 25$$

$$T(p = 0.6) = 18$$

$$T(p = 0.8) = 30$$

Perform the following tasks: -

- i. Create a suitable diagram for the above problem.
- ii. Compute the expected return for each value of p .
- iii. Identify which policy (value of p) is best.

Comment briefly on why performance decreases when p is too small or too large.

$$7 \text{ (a) Given: } Q(S_t, A_t) = 2, R_{t+1} = 3, Q(S_{t-1}, a_1) = 7, Q(S_{t-1}, a_2) = 8, \gamma = 0.9, \alpha = 0.1 \quad (5)$$

Assume that under the current behavior policy, the next action selected at state S_{t+1} is a_1 .

- Compute the updated value of $Q(S_t, A_t)$ using the SARSA update rule.
- Compute the updated value of $Q(S_t, A_t)$ using the Q-learning update rule (assume greedy action selection).

- (b) Compare Dynamic Programming (DP), Monte Carlo (MC) , and Temporal Difference (TD) methods. Evaluate the methods based on following basis: (5)
- Model Requirement
 - Bootstrapping
 - Suitability for Episodic/Continuous tasks.
- (c) In the Monte Carlo (MC) Prediction task, an agent has visited state S five times, resulting in an average return of $V(S) = 2.0$. On the sixth visit to state S , the agent observes a new return (G_t) of 4.0. (5)

Calculate the new value estimate $V(S)$ using the incremental update rule.