R.V. College of Engineering, Bangalore

Lesson Plan

Semester: **III** Year: **2023-2024**

Course Title: **Fundamentals of Data Structures and Data Analysis (DSDA)**	*Course Code*: **AI233AI**
Total Contact Hours: 45 L + 30 P L:T:P : 3:0:1	*Duration of SEE*: 3 **Hrs.**
SEE Marks: **100**	*CIE Marks*: **100**
Lesson Plan Author: **Dr. B. Sathish Babu**	*Date*: **11/12/2023**
Checked By:	*Date*:

Course Overview:

- Data structures deal with various formats of organizing, storing, retrieving, and processing data.
- Data structures play a critical role in algorithms development and the efficient working of algorithms.
- The creation of appropriate data structures and the development of efficient algorithms are the primary requirements for a programmer.
- Choosing suitable data structures will enable you to do proper logic management.
- Data structures hold the complete internal state of your code.

Course Outcomes: After completing the course, the students will be able to	
CO1	Apply the knowledge of data structures in providing solutions to some software development requirements.
CO2	Perform data analysis of some real-world scientific/business use cases and present the analysis results.
CO3	Investigate appropriate data structures and understand requirements in solving some problems of industry and society.
CO4	Use data analysis tools to illustrate the principles of data interpretation, statistical analysis, and graphical visualizations of the datasets.
CO5	Appraise data structures and analysis knowledge to build a successful career as an AIML engineer, work in teams, and communicate their ideas effectively.

Course Content

Unit-I	
Importance of Data Structures and Data Analysis in AIML engineering with real-world examples. **Introduction:** Data Structure, Classifications of Data Structures, Application of Data Structures, Abstract Data Type, Operations Perform on Data Structure, Overview of Different Data Structures, **(2 Hrs.)**	**09 Hrs**

Pointers: Pointer Declaration, Address of Operator, Indirection Operator, Null Pointer, void Pointer, Generic Functions , Dangling Pointer, Arithmetic Operation with Pointer, Pointer to pointer Pointers and Arrays, Array of Pointers, Pointer to an Array, Pointer to Function, Passing addresses to Function, Function returning Pointer, Dynamic Memory Allocation **(2 Hrs.)** **Linked Lists:** Limitations of Array, Linked List, Singly Linked list, Operations on Singly linked list, Representation of polynomials using linked list, Circular Linked list, Operation on Circular Link List, Josephus Problem, Doubly Linked list, Operation on Doubly Link List, Circular Doubly Linked List, Disadvantages of Linked List **(5 Hrs.)**	
Unit – II	
Stacks and Queues: Stack, Operations on Stack, Stack Representation with Array, Stack Representation with Linked List, Processing of function calls, Evaluation of Arithmetic expressions; Queue, Operations on Queue, Queue Representation with Array, Queue Representation with Linked List, Application of Queue, Drawback of Linear Queue, Circular Queue, Circular Queue Representation with Array, Dequeue, Operation on DeQueue, Priority Queue, Representation of Priority Queue **(04 Hrs.)** **Trees:** Terminology of Tree, Binary Tree, Strictly Binary Tree, Extended Binary Tree, Complete Binary Tree, Full Binary Tree, Skewed Binary Tree, Binary Expression Tree, Balanced Binary Tree, Threaded Binary Tree, Properties of Binary Tree, Representation of Binary Tree, Binary Tree Traversal, Binary Search Tree, Operations on Binary Search Tree, Heaps **(05 Hrs.)**	09 Hrs
Unit –III	
Graphs: Terminology of Graph, Terminology of a Directed Graph, Operations on Graph, Representation of Graph, Graph Traversal, Shortest Paths, Dijkstra's Algorithm, Bellman-Ford Algorithm **(05 Hrs.)** **Hashing:** Hash Table, Hash Function, Division Method, Mid Square method, Folding method, Collision Resolution, Linear Probing, Quadratic Probing, Double Hashing, Separate Chaining, Load Factor **(04 Hrs.)**	09 Hrs
Unit –IV	
Introduction to Data-Analytic Thinking: The ubiquity of data opportunities, Examples: Hurricane Fraces, Predicting Churn, Data Science, Engineering and Data-Driven Decision Making, Data Processing and Big Data, From Big Data 1.0 to Big Data 2.0, Data-Analytic Thinking **(04 Hrs.)** **Business Problems and Data Science Solutions:**	09 Hrs

From Business Problems to Data Mining Tasks, Supervised and Unsupervised Methods, Data mining and its results, The Data mining process: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation and Deployment, Implications of managing the data science team **(05 Hrs.)**	
Unit –V	
Introduction to Predictive Modeling:: Models, Induction and Prediction, Supervised Segmentation: Selecting informative attributes, Attribute selection with information gain, supervised segmentation with Tree-structured models, Visualizing segmentations, Tress as sets of rules, probability estimation, Example of churn problem.	**09 Hrs**

Text Books/Reference Books	
1	**Data Structures and Algorithms with C, Debdutta Pal and Suman Halder, Alpha Science International Ltd, Oxford, UK, 2018. ISBN 978-1-78332-368-5, E-ISBN 978-1-78332-427-9 (Unit 1 to Unit 3)**
2	**Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking 1st Edition, Foster Provost and Tom Fawcett, O'Reilly Media, 2013. ISBN: 978-1449361327 (Unit 4 and Unit 5)**
3	Fundamentals of Data Structures, Ellis Horowitz, Sartaj Sahni, Illustrated Edition, Computer Science Press.
4	R for Beginners, Emmanuel Paradis, 2005

R.V. College of Engineering, Bangalore

Unit and Chapter wise Plan

Unit I

Chapter Number and Title: Chapter-1 Introduction; Chapter-4 Pointers; Chapter-5 Linked List (TEXTBOOK-1)	*Planned Hours: 09*

Learning Objectives:

Sl. No.	Objectives
1	To appraise the need of data structures in computer programming
2	To understand the pointer concept in depth
3	To build dynamic data structures like linked list, circular list, and variants of linked list.

Lesson Schedule

Class No. Portions covered per hour

1. Introduction to Data Structures, its importance, applications, etc.
2. Classification of data structures, ADTs, Operations on data structures and overview of data structures
3. Pointer Declaration, Address of Operator, Indirection Operator, Null pointer, void pointer, generic functions, dangling pointer, arithmetic operation with pointer, pointer to pointer
4. Array of pointers, Pointer to an array, pointer to function, passing addresses to function, function returning pointer, dynamic memory allocation
5. Arrays versus Linked list, and advantageous of linked list, singly linked list insertion, deletion and traversal
6. Polynomial representation using singly linked list
7. Circular linked list creation, insertion, deletion and traversal
8. Doubly linked list creation, insertion, deletion and traversal
9. Doubly linked list creation, insertion, deletion and traversal

Model Questions

1. You are developing a tiny browser, in that you want to include a forward and backward options, suggest which data structure you are going to use for this purpose, why and how?

2. Static Circular Queues uses the memory space efficiently than Static Linear Queues, Justify this statement using an example.

3, By making use of linked list, implement the operation of Circular Queue.

4. Consider a double linked list with a node structure given below, each node stores two links one for predecessor node and one for successor node. A data field is of 30 characters. Assume the list is created with the starting node address as "Start". Write a C program to delete the occurrences of a specific word.

llink	Data (30 characters)	Rlink

5. Write a C program which has functions to create a single linked list with integers, display the list and reverse the linked list.

6. Write recursive function to find the factorial of a number, and trace the working of the function using Stacks.

7. Discuss the logic for merging of two linked lists.

8. Input a polynomial of degree-n, with three variables, and coefficient for every term. Use linked list to store the polynomial and display the same.

9. Use linear queues of size N, to store requests coming for a shared printer. Each request has number of pages to be printed. Demonstrate both insertion and deletion from the queues.

10. Write a program that performs the following:

    a. Creates a singly linked list of integers
    b. Display the list
    c. Delete all the nodes which have the key and display the list after deletion, or display the key not found in the list

Unit II

Chapter Number and Title: Chapter-6: Stack and Queue; Chapter-8: Tree (TEXTBOOK-1)	*Planned Hours: 09*

Learning Objectives:

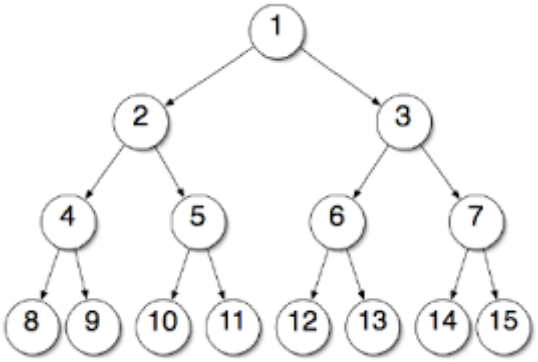
Sl. No.	Objectives
1	To create Stacks using Arrays and Linked list, and perform various operations
2	To create Queue using Arrays and Linked list, and perform various operations
3	To create Binary trees and Binary Search Tree and other variants of Tree

Lesson Schedule

Class No.	*Portions covered per hour*
1.	Stack and Stack Operations using Arrays
2.	Stack operations using linked list and evaluation of arithmetic operations
3.	Queue and Queue Operations using Arrays
4.	Queue operations using linked list, Circular queue, Dequeue, and Priority queue
5.	Tree and Binary Trees, Types
6.	Expression Trees
7.	Balanced Binary Tree & Traversals
8.	Binary Search Tree & Searching
9.	Heap introduction

Model Questions

1. Perform all the three traversals on the given binary tree



2. Prove that "The external path length of any binary tree with n internal nodes is 2n greater than its internal path length."

3, Write a C program to create a BST of integers, and perform all the three traversals.

4. Write and explain the logic for deleting a node from the BST.

5. Create a C function to insert element into Max-Heap or deleting element from the Max-Heap.

6. What advantageous the threaded binary trees over normal binary trees.

7. Discuss briefly an application of BST in developing any mobile applications.

8. Give a real-world application of Max-Heap.

9. Create an expression tree using the input as postfix expression.

10 By using a stack of type characters, write a C program to check the given string has a matching number of curly brackets; EXAMPLE: {{AAABBBBCCC}} Output: It is Balanced

Unit III

Chapter Number and Title: Chapter-9: Graph; Chapter-11: Hashing (TEXTBOOK-1)	*Planned Hours: 09*

Learning Objectives:

Sl. No.	Objectives
1	To understand the working principles of Graphs data structures, Representations, Traversals, and applying the same in solving some real-world problems
2	To understand the working principles of Hashing functions, Hash Tables, and Hash Table probing Techniques

R.V. College of Engineering, Bangalore

Lesson Schedule

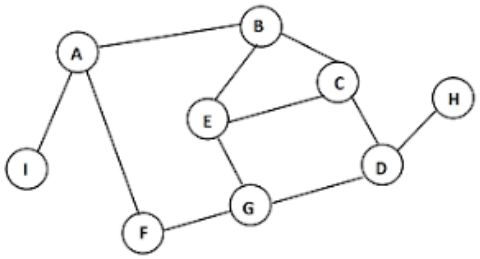
Class No.	*Portions covered per hour*
1.	Introduction to Graph Data Structure, Types and Terminologies
2.	Terminologies of a Directed graph
3.	Representation of Graphs - Adjacency Matrix and Lists
4.	Graph Traversals - DFS and BFS
5.	Dijkstra's Algorithm and Bellman-Ford Algorithm
6.	Hash functions and types
7.	Working of Linear Probing
8.	Working of Quadratic Probing
9.	Double Hashing & Separate Chaining

Model Questions

1. Use the adjacency lists and adjacency matrix to represent the following sparse matrix.

$$\begin{bmatrix} 0 & 9 & 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 5 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

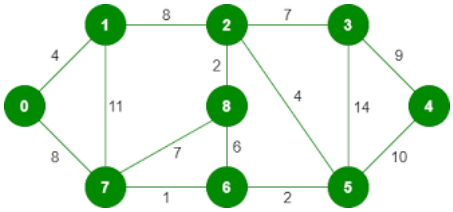
2. Demonstrate the working of DFS and BFS on the following graph.



3, Write a C program to demonstrate the working of hashing by Division.

4. Write a C program to demonstrate the working of hashing by Multiplication.

5. Create a spanning tree for the given graph



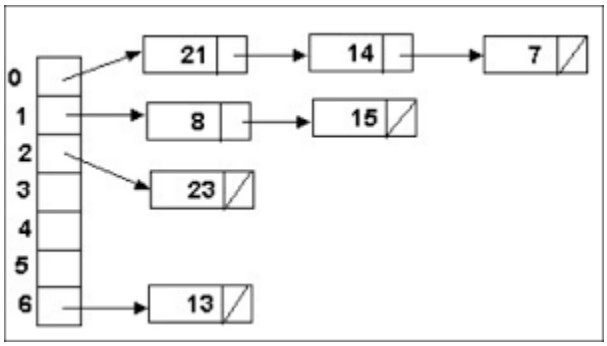
6. Write a C program to demonstrate the working of DFS.

7. Write a C program to demonstrate the working of BFS.

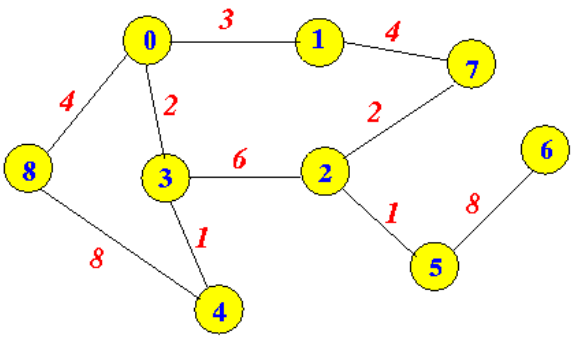
8. Compare Quadratic Probing with Linear Probing.

9. Load balancing refers to efficiently distributing incoming service requests across a group of backend servers, also known as a server farm or server pool (Assume all the servers are executing same task). Discuss how hashing can be used to implement such scenario?

10. Write a C program to demonstrate the working of hashing by chaining. Assume the maximum length of the hash table as m. Sample of hashing by chaining:



11. Write the DFS traversal for the given weighted graph. Assume starting vertex as 3.



12. Given the following C function, which is written for BFS traversal of a graph, with source vertex as v. Correct the logic if there are any errors, and write the corrected version. Assume adj[][], visited[], front, and rear have been properly initiallised before the function was called.

```
void bfs(int v)
{
   for (i = 1; i <= n; i++)
      if (adj[v][i] && visited[i])
         queue[rear++] = i;
   if (front <= rear)
   {
      visited[queue[front]] = 1;
      printf("Node visited = %d\n", queue[front]);
      bfs(adj[v][i]);
   }
}
```

13. Given the input {4371, 1323, 6173, 4199, 4344, 9679, 1989}, a fixed table size of 10, and a hash function h(x) = x mod 10, show the working of; a. Hashing with Linear Probing and b. Hashing with Quadratic probing

Unit IV

Chapter Number and Title: Chapter-1: Introduction : Data Analytic Thinking; Chapter-2: Business Problems and Data Science Solutions (TEXTBOOK-2)	*Planned Hours: 09*

Learning Objectives:

Sl. No.	Objectives
1	Introducing the concept of Data science, and data-driven decision making
2	To learn various concepts in data mining
3	To understand various phases of data mining process

Lesson Schedule

Class No.	*Portions covered per hour*
1.	The ubiquity of data opportunities
2.	Examples : Hurricane Frances, Predicting Customer Churn
3.	Data Science, Engineering and Data-driven decision making
4.	Data Processing and "Big Data", Data-Analytic Thinking
5.	From Business Problems to Data Mining Tasks
6.	Supervised and Unsupervised Methods
7.	Data mining and its results
8.	Data mining process
9.	Data mining process

Model Questions

1 Illustrate by considering an example, how patterns are extracted by doing data analysis.

2. By considering an example of Hurricane France, illustrate the need for data-driven decision making.

3. By considering an example of Customer churn, illustrate the need for data-driven decision making.

4. List various advantageous of data-driven decision making.

5. Explain the role and impact of Big data in data-driven decision making.

6. Give suitable examples for Profiling, Clustering, Similarity matching and Regression.

7. Explain the CRISP data mining process.

Unit V

R.V. College of Engineering, Bangalore

Chapter Number and Title: Chapter-3: Introduction to Predictive Modeling (TEXTBOOK-2)	*Planned Hours: 09*

Learning Objectives:

Sl. No.	Objectives
1	To understand the notions of models, induction, prediction and segmentation
2	To work with Trees, Segmentation visualization, and Estimations
3	To study the customer churn problem from data science perspective

Lesson Schedule

Class No.	*Portions covered per hour*
1.	Models, Induction and Prediction
2.	Supervised segmentation
3.	Selecting informative attributes
4.	Attribute selection with information gain
5.	Supervised segmentation with tree-structured models
6.	Visualizing segmentation
7.	Trees as sets of rules
8.	Probability estimation
9.	Churn Problem

Model Questions

1. By considering an example discuss the concept of prediction.
2. By considering an example discuss the notion of descriptive modeling.
3, Give suitable examples for syntactic accuracy and semantic accuracy.
4. Discuss various challenges while selecting informative attributes.
5. By considering any three types of data visualization discuss the need.
6. Numerical on Entropy and Information Gain.
7. How does the tree-structured models are employed in supervised segmentation?
8. Which techniques are used to visualize segmentations?

CIE Evaluation Scheme for Theory

Continuous Internal Evaluation (CIE) (Theory – 100 Marks)	
Evaluation method	Course with EL
Quiz -1	10
Test -1	50
Quiz -2	10
Test-2	50
MAKEUP-TEST/QUIZ	50
EL	40
Total – theory	**100 (Best Two Quizzes (20) +Best of Two Tests Reduced to 40 +EL (40))**

Experiential Learning Components

Component #1 (30 Marks)

- Demonstration of CRISP Data Mining Process for the selected business problem from a domain, e.g., supplier selection in a supply chain management, inventory management in a selected business-like jewelry, etc. (15 Marks)
- Using Open source tools like Orange, RapidMiner, WEKA, etc., demonstrate the data mining process (15 Marks)

Component #2 (10 Marks)

Expected Outcome (any of the following):

- Video-based Presentation on a real-world application of Data Structures
- The Outcome of participating in Hackathons (Inter-collegiate or https://www.hackerearth.com/challenges/hackathon/)
- Producing international conference/journal paper/book chapter
- Part of organizing an event
- Resource person for a topic

- Part of department-level industrial consultancy or research activity
- Any other activity with the permission of the subject faculty

Semester End Evaluation Theory (100)	
Part - A Objective type questions	20
Part - B Five questions from five units. Each question is for a maximum of 16 Marks. The UNIT-1 is compulsory. UNIT-2 and UNIT-5 have an internal choice. Both questions should be of the same complexity in terms of COs and Bloom's taxonomy level.	80
Total	100

CIE and SEE Evaluation Scheme for Lab

CIE	SEE
Conduction: 25 Lab Test: 05 Experiential Learning: 20 Total: 50	Conduction: 40 Viva-Voce: 10 Total: 50

Course Articulation Matrix (3: High 2: Medium 1: Low)

CO-PO Mapping												
CO/PO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	2	-	-	-	-	-	-	-	-	-	-	-
CO2	-	2	-	2	-	-	-	-	-	-	-	-
CO3	_	-	2	_	_	1	-	-	-	-	-	-
CO4	_	-	-	_	3	-	-	-	-	-	-	-
CO5	_	-	-	_	_	-	-	-	2	2	-	2

R.V. College of Engineering, Bangalore

CO –PSO Mapping		
CO/PSO	PSO1	PSO2
CO1	2	
CO2		2
CO3	1	
CO4		1
CO5	2	

Innovative teaching method/s used

Sl. No.	Type	Purpose
1	Flipped Classroom	To learn data structures, coding and operations.
2	Think-Pair-Share	Used in the classroom to develop logic and coding skills.
3	Google Classroom	Used to upload supporting study materials and to give assignments.

Name of the faculty with Sign. HOD

Dr. B. Sathish Babu

Professor & HoD, AI & ML