Semester: III						
FUNDAMENTALS OF DATA STRUCTURES AND DATA ANALYSIS **Category :PROFESSIONAL CORE COURSE** **(Theory and Practice)**						
Course Code	**:**	**AI233AI**		**CIE**	**:**	**100+50 Marks**
Credits: L: T: P	**:**	**3:0:1**		**SEE**	**:**	**100+ 50 Marks**
Total Hours	**:**	**45L+30P**		**SEE Duration**	**:**	**3.00 Hours**

Unit-I — **09 Hrs.**

Introduction:
Data Structure, Classifications of Data Structures, Application of Data Structures, Abstract Data Type, Operations Perform on Data Structure, Overview of Different Data Structures

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

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

Unit – II — **09 Hrs.**

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 De Queue, Priority Queue, Representation of Priority Queue

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

Unit –III — **09 Hrs.**

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

Hashing:
Hash Table, Hash Function, Division Method, Mid Square method, Folding method, Collision Resolution, Linear Probing, Quadratic Probing, Double Hashing, Separate Chaining, Load Factor

Unit –IV — **09 Hrs.**

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

Business Problems and Data Science Solutions:
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

Unit –V	09 Hrs.
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.	

List of Laboratory Experiments

Exp. No.	Data Structure Name	Application to be coded using C
1	Stack	• Arithmetic Expression Evaluation o Evaluating the given postfix expression by considering the priority of the operators. o Identify the invalid expression o Operators: +,/,* and - o Nested Parenthesis: (())
2	Queues	• Simulating a shared resource management o Create a simulated version of a shared memory. o Generate N random requests for memory. o Use queues to manage the resource requests.
3	Singly Linked List	• Polynomial Arithmetic o Adding two polynomials of any degrees.
4	Doubly Linked List	• Simple Text Editor o Browsing through the text, line by line in both directions o Insert New lines anywhere in the text o Delete line/s from the text
5	Binary Trees	• Arithmetic Expression Conversion o Building an expression tree o Infix to Prefix conversion o Infix to Postfix conversion
6	Binary Search Trees	• Creating a dictionary of words o Insert a new word into a dictionary o Delete a word from a dictionary o Print Dictionary
7	Graphs	• Implementing Dijkstra's algorithm and finding the shortest route between nodes
8	Hash Table	• Implementing the Rabin-Karp algorithm for pattern matching using Hashing
9	Heaps	• Implement a Max-heap data structure from a binary tree

PART-B OF LAB

A batch of two students develops a prototype using the C/C++ language. The prototype demonstrates the use of data structure in real-time applications. E.g., using trees to index search results, using graphs to navigate places, using graphs for recommendations and match-making, using queues for message passing, developing spell and grammar checkers, using matrices to generate the survey insights, etc. (Ref: https://www.geeksforgeeks.org/real-time-application-of-data-structures/). The innovative applications of data structures attract high marks.

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	Identify appropriate data structures and understand requirements in solving some problems of industry and society.
CO3	Perform data analysis of some real-world scientific/business use cases and present the analysis results.
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.

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
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
3	Fundamentals of Data Structures, Ellis Horowitz, Sartaj Sahni, Illustrated Edition, Computer Science Press.

RUBRICFOR THE CONTINUOUS INTERNAL EVALUATION		
#	**COMPONENTS**	**MARKS**
1.	**QUIZZES:** Quizzes will be conducted in online/offline mode. **TWO QUIZZES** will be conducted & Each Quiz will be evaluated for 10 Marks. Each quiz is evaluated for 10 marks adding up to 20 MARKS	**20**
2.	**TESTS:** Students will be evaluated in test, descriptive questions with different complexity levels (Revised Bloom's Taxonomy Levels: Remembering, Understanding, Applying, Analyzing, Evaluating, and Creating). **TWO tests will be conducted**. Each test will be evaluated for **50 Marks**, adding upto 100 Marks. **FINAL TEST MARKS WILL BE REDUCED TO 40 MARKS.**	**40**
3.	**EXPERIENTIAL LEARNING:** Students will be evaluated for their creativity and practical implementation of the problem. Case study based teaching learning (10), Program specific requirements (10), Video based seminar/presentation/demonstration (10) Designing & Modeling (10) **Phase 2 will be done in the exhibition mode (Demo/Prototype/any outcome). ADDING UPTO 40 MARKS**.	**40**
4.	**LAB:** Conduction of laboratory exercises, lab report, observation, and analysis (30 Marks), lab test (10 Marks) and Innovative Experiment/ Concept Design and Implementation (10 Marks) adding up to 50 Marks. THE FINAL MARKS WILL BE 50 MARKS	**50**
	MAXIMUM MARKS FOR THE CIE(THEORY+LAB)	**150**

RUBRIC FOR SEMESTER END EXAMINATION (LAB)		
Q.NO.	**CONTENTS**	**MARKS**
1	Write Up	10
2	Conduction of the Experiments	20
3	Viva	20
	TOTAL	**50**

RUBRIC FOR SEMESTER END EXAMINATION (THEORY)		
Q.NO.	**CONTENTS**	**MARKS**
	PART A	
1	Objective type of questions covering entire syllabus	20
	PART B (Maximum of THREE Sub-divisions only)	
2	Unit 1 : (Compulsory)	16
3 & 4	Unit 2 : Question 3 or 4	16
5 & 6	Unit 3 : Question 5 or 6	16
7 & 8	Unit 4 : Question 7 or 8	16
9 & 10	Unit 5: Question 9 or 10	16
	TOTAL	**100**