

USN

1 R V 2 1 A 1 0 3 0

RV COLLEGE OF ENGINEERING®

(An Autonomous Institution Affiliated to VTU)

III Semester B. E. Examinations April-2023

Artificial Intelligence and Machine Learning

DATA STRUCTURES AND DATA ANALYTICS

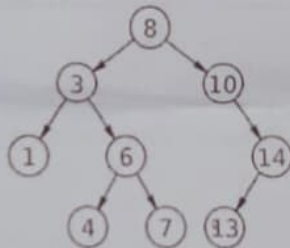
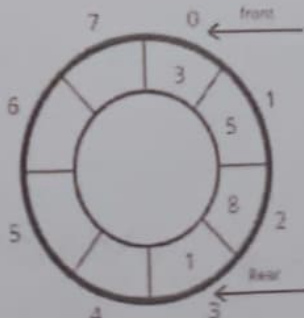
Time: 03 Hours

Maximum Marks: 100

Instructions to candidates:

1. Answer all questions from Part A. Part A questions should be answered in first three pages of the answer book only.
2. Answer FIVE full questions from Part B. In Part B question number 2 is compulsory. Answer any one full question from 3 and 4, 5 and 6, 7 and 8, 9 and 10.

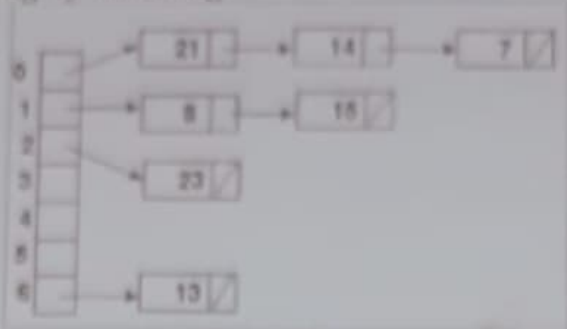
PART-A

1	1.1	Write the expression tree for the given expression: $(A + B)/((C - D) - (E * F)) * G$	02
	1.2	How does graphs are useful in applications like location navigation systems.	02
	1.3	Write a data structure which represents the in order threaded binary tree version of the given binary tree.	
			
	1.4	Differentiate between linear probing and quadratic probing in hashing.	01
	1.5	By considering a medical application, give the scope for using the heap data structure.	02
	1.6	Write a recursive function to do the Postorder traversal on a binary tree.	01
	1.7	Given the following circular queue status, write the front and rear pointers and queue status after the following operations. Insert 10,20,30 Delete 3,5 Insert 40,50,60 and Insert 70.	02
			02

1.8	Write the Max Heap for the given binary tree.	
	<pre> graph TD 8((8)) --> 3((3)) 8 --> 10((10)) 3 --> 1((1)) 3 --> 6((6)) 6 --> 4((4)) 6 --> 7((7)) 10 --> 14((14)) 14 --> 13((13)) </pre>	
1.9	Define exploratory data analysis.	01
1.10	Differentiate between experiential study and observational study.	01
1.11	For the following examples, mention whether they can be solved using regression or classification.	02
	a) Predicting the blood pressure of a patient b) Identifying the type of vehicle on the road.	02
1.12	Write a C function that finds the maximum data in a singly linked list of integers. Assume the list starting address as <i>START</i> and no duplicates.	02

PART-B

2	a	In which scenario do you prefer using linked lists over arrays? Write a C program, which creates a singly linked list of characters, and after creation, delete all the nodes which have vowels. Print the final list.	08
	b	Complete the following C function, which uses a stack 's' to check the given expression 'expr', has balanced pairs of {},[] and (). The function should return 1 if its is balanced ; otherwise 0. Int is balanced (char expr[]) { // write your code here }	08
3	a	Prove that the height of a binary tree with 'n' internal nodes is at least $\log_2(n + 1)$ and at most $n - 1$.	06
	b	How does a Binary Search Tree (BST) different from a regular Binary tree? Mention its advantages? Write a C program that uses a BST to sort the 'n' number of integers and outputs the list in ascending order.	10
		OR	
4	a	Write a C program to create a Min Heap of integers and display the same. (Hint: Use arrays to represent the Min heap).	10
	b	Illustrate with an example of at least five operators, the process used to create expression trees.	06
5	a	How does the DFS traversal differ from the BFS traversal of a Graph, illustrate by considering a graph of 6 nodes? Write a C program to create an adjacency matrix or adjacency list of any directed and weighted graph G with N vertices. Use DFS traversal to find the path length between two vertices, S and D, if it exists. Otherwise, display path not exists.	10
	b	Give the definition and real world applications of the following: <ol style="list-style-type: none"> Simple graphs Connected graphs Spanning trees. 	06
		OR	

6	a	<p>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:</p> 	10
	b	<p>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 scenarios?</p>	06
7	a	<p>By considering a traffic lights operations in a junction, list any three data items and any two types of knowledge? Write a block diagram depicting various phases of the <i>CRISP-DM</i> process and briefly explain the importance of all stages.</p>	08
	b	<p>Assume you have been hired to analyze a type of customer of a mega jewelry showroom to offer some goodies. Propose a sound approach you follow for the same.</p>	08
OR			
8	a	<p>You are asked to create a cognitive map for the frequency of watching a movie on an <i>OTT</i> platform. Assume at least ten attributes and show in your diagram which items are negatively correlated, positively correlated, and relationships depend on the product.</p>	08
	b	<p>Discuss the importance of the following in the context of data analysis:</p> <ul style="list-style-type: none"> i) Model flexibility/adequacy ii) Problems faced in data analysis projects. 	08
9	a	<p>Why visualization of data is important in data analysis? Discuss any four types of visualization we can have on the dataset by mentioning the purpose.</p>	08
	b	<p>Discuss the consequences of missing values in a data set. What do we mean by missing completely at random and observed at random?</p>	08
OR			
10	a	<p>Differentiate between the syntactic and semantic accuracy of the data. Also discuss the meaning of data completeness and data timeliness.</p>	08
	b	<p>Write a technical note on correlation analysis, including Principal Component Analysis (<i>PCA</i>) along with necessary equations and representations.</p>	08