Reg No.:_____ Name:_____

**APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY**

B.Tech Degree S6 (S,FE) Examination December 2025 (2019 Scheme)

**Course Code: ADT308**

**Course name: COMPREHENSIVE COURSE WORK**

Max. Marks: 50 Duration: 1Hour

***Instructions:***
***(1) Each question carries one mark. No negative marks for wrong answers***
***(2) Total number of questions: 50***
***(3) All questions are to be answered. Each question will be followed by 4 possible answers of which only ONE is correct.***
***(4) If more than one option is chosen, it will not be considered for valuation.***

1 Let the page fault service time be 10ms in a computer with average memory access time being 20ns. If one page fault is generated for every 10^6 memory accesses, what is the effective access time for the memory?

a) 21ns b) 30ns c) 23ns d) 35ns

2 A system uses FIFO policy for page replacement. It has 4 page frames with no pages loaded to begin with. The system first accesses 100 distinct pages in some order and then accesses the same 100 pages but now in the reverse order. How many page faults will occur?

a) 196 b) 192 c) 197 d) 195

3 Consider three processes (process id 0, 1, 2 respectively) with compute time bursts 2, 4 and 8 time units. All processes arrive at time zero. Consider the longest remaining time first (LRTF) scheduling algorithm. In LRTF ties are broken by giving priority to the process with the lowest process id. The average turn around time is:

a) 13 units b) 14 units c) 15 units d) 16 units

4 Increasing the RAM of a computer typically improves performance because:

a) Virtual memory increases b) Larger RAMs are faster c) Fewer page faults occur d) Fewer segmentation faults occur

5 A computer system supports 32-bit virtual addresses as well as 32-bit physical addresses. Since the virtual address space is of the same size as the physical address space, the operating system designers decide to get rid of the virtual memory entirely. Which one of the following is true?

a) Efficient implementation of multi-user support is no longer possible
b) The processor cache organization can be made more efficient now
c) Hardware support for memory management is no longer needed
d) CPU scheduling can be made more efficient now

6 Consider a disk pack with 16 surfaces, 128 tracks per surface and 256 sectors per track. 512 bytes of data are stored in a bit serial manner in a sector. The capacity of the disk pack and the number of bits required to specify a particular sector in the disk are respectively:

a) 256 Mbytes, 19 bits
b) 256 Mbytes, 28 bits
c) 512 Mbytes, 20 bits
d) 64 Gbytes, 28 bits

7 Consider three CPU-intensive processes, which require 10, 20 and 30 time units and arrive at times 0, 2, and 6 respectively. How many context switches are needed if the operating system implements a shortest remaining time first scheduling algorithm ? Do not count the context switches at time zero and at the end.

a) 1
b) 2
c) 3
d) 4

8 Which of the following scheduling reduces process flow time?

a) FCFS
b) LIFO
c) SJF
d) None of these

9 How many types of buffer overflow in the operating system?

a) 1
b) 2
c) 3
d) 4

10 Consider a machine with 64 MB physical memory and a 32-bit virtual address space. If the page size is 4KB, what is the approximate size of the page table?

a) 16 MB
b) 8 MB
c) 2 MB
d) 24 MB

11 How many queues are required to implement a stack? Where no other data structure like arrays and the linked list is available to you.

a) 1
b) 2
c) 3
d) 4

12 A Priority-Queue is implemented as a Max-Heap. The level-order traversal is 10, 8, 5, 3, and 2. Elements 1 and 7 are inserted in a heap in, respectively. The level-order traversal of the heap will be

a) 10, 8, 7, 5, 3, 2, 1
b) 10, 8, 7, 2, 3, 1, 5
c) 10, 8, 7, 1, 2, 3, 5
d) 10, 8, 7, 3, 2, 1, 5

13 The height of a tree is the length of the longest root-to-leaf path in it. The maximum and minimum number of nodes in a binary tree of height 5 are

a) 63 and 6
b) 64 and 5
c) 32 and 6
d) 31 and 5

14 The maximum number of binary trees that can be formed with three unlabeled nodes is:

a) 1 b) 5 c) 4 d) 3

15 A program P reads in 500 integers in the range [0, 100] representing the scores of 500 students. It then print the frequency of each score above 50. What would be the best way for P to store the frequencies?

a) An array of 50 numbers b) An array of 100 numbers c) An array of 500 numbers d) None of these

16 The result evaluating the postfix expression is: 10 5 + 60 6 / * 8 -

a) 284 b) 213 c) 142 d) 71

17 A hash table contains 10 buckets and uses linear probing to resolve collisions. The key values are integers and the hash function used is key % 10. If the values 43, 165, 62, 123, 142 are inserted in the table, in what location would the key value 142 be inserted?

a) 2 b) 3 c) 4 d) 6

18 The best data structure to check whether an arithmetic expression has balanced parentheses is a

a) Stack b) Queue c) Tree d) List

19 Consider the tree arcs of a BFS traversal from a source node W in an unweighted, connected, undirected graph. The tree T formed by the tree arcs is a data structure for computing

a) The shortest path between every pair of vertices b) The shortest path from W to every vertex in the graph c) The longest path in the graph d) None of these

20 What is the maximum number of swaps that can be performed in the Selection Sort algorithm?

a) n-1 b) n c) 1 d) n-2

21 In a relational data model, which one of the following statements is TRUE?

a) A relation with only two attributes is always in BCNF b) If all attributes of a relation are prime attributes, then the relation is in BCNF c) Every relation has at least one non-prime attribute d) BCNF decompositions preserve functional dependencies

22 Which of the following statements is correct?
1. A self-join may be implemented in SQL using table aliases.
2. The outer-join operation is a fundamental relational algebra procedure.
3. The procedures of natural join and outside join are the same.

a) 1 is correct b) 1 and 2 is correct c) 2 and 3 is correct d) None of these

23 Location transparency in distributed databases allows database users, programmers, and administrators to handle data as if it were stored in a single location. A SQL query with location transparency must provide the following information:

a) Fragments b) Locations c) Local formats d) Inheritances

24 A SQL command changes one or more fields of a table:

a) Look Up b) Insert c) Change d) Modify

25 What do you mean by one to many relationships?

a) One class may have many teachers b) One teacher can have many classes c) Many classes may have many teachers d) Many teachers may have many classes

26 The term "FAT" is stands for_____

a) File Allocation Table b) File Allocation Tree c) File Allocation Template d) None of these

27 Which one of the following refers to the "data about data"?

a) Directory b) Meta Data c) Sub Data d) None of these

28 Which one of the following commands is used to modify a column inside a table?

a) Drop b) Update c) Alter d) None of these

29 Which of the following is generally used for performing tasks like creating the structure of the relations, deleting relation?

a) DML b) Query c) Relational Schema d) DDL

30 Which one of the following given statements possibly contains the error?

a) select * from emp where empid = 10003; b) select empid from emp where empid = 10006; c) select empid from emp; d) select empid where empid = 1009 and Lastname = 'GELLER';

31 K-Nearest Neighbors (KNN) is classified as what type of machine learning algorithm?

a) Instance-based learning b) Parametric learning c) Non-parametric learning d) Model-based learning

32 Which of the following is not a supervised machine learning algorithm?

a) K-means b) Naive Bayes c) SVM d) Decision Tree

33 Which of the following statements is false about Ensemble learning?

a) It is a supervised learning algorithm
b) It is an unsupervised learning algorithm
c) More random algorithms can be used to produce a stronger ensemble
d) Ensembles can be shown to have more flexibility in the functions they can represent

34 Which of the following statements is not true about boosting?

a) It mainly increases the bias and the variance
b) It mainly decreases the bias and the variance
c) It is a technique for solving two-class classification problems
d) None of these

35 Which machine learning algorithm is suitable for solving regression problems?

a) K-Means clustering
b) Random Forest
c) KNN
d) Apriori Algorithm

36 What is the purpose of the bias term in a linear regression model?

a) To handle outliers
b) To avoid overfitting
c) To model the noise in the data
d) To shift the regression line up or down

37 Which evaluation metric is commonly used for classification tasks when class imbalance is present?

a) MSE
b) Accuracy
c) F1 Score
d) R squared

38 In the context of machine learning, what is the term "bias-variance trade-off" referring to?

a) The trade-off between the quality of training data and testing data
b) The trade-off between the complexity of a model and its ability to generalize
c) The trade-off between the number of features and the size of the dataset
d) The trade-off between precision and recall in classification

39 Which algorithm is used for hierarchical clustering?

a) K-Means clustering
b) Agglomerative clustering
c) DBSCAN
d) PCA

40 Which method can be used to handle missing data in a dataset?

a) Removing the entire column with missing data
b) Replacing missing data with the median value of the column
c) Ignoring the rows with missing data during analysis
d) All of the above

41 Gradient Boosting is an ensemble technique that:

a) Combines predictions using a weighted average
b) Trains multiple models on different subsets of the data
c) Constructs an ensemble by iteratively updating weights
d) Uses a committee of experts to make predictions

42 Which ensemble learning algorithm can handle both numerical and categorical data without requiring one-hot encoding?

a) AdaBoost
b) Random Forest
c) Gradient Boosting
d) Stacking

43 Which one of the following terms is not used in the Bayes' Theorem?

a) Prior
b) Unlikelihood
c) Posterior
d) Evidence

44 For what purpose, the analysis tools pre-compute the summaries of the huge amount of data?

a) In order to maintain consistency
b) For authentication
c) For data access
d) To obtain the queries response

45 Which one of the following can be considered as the final output of the hierarchal type of clustering?

a) A tree which displays how the close thing are to each other
b) Assignment of each point to clusters
c) Finalize estimation of cluster centroids
d) None of these

46 Which of the following also used as the first step in the knowledge discovery process?

a) Data selection
b) Data cleaning
c) Data transformation
d) Data integration

47 In certain cases, it is not clear what kind of pattern need to find, data mining should_____

a) Try to perform all possible tasks
b) Perform both predictive and descriptive task
c) It may allow interaction with the user so that he can guide the mining process
d) All of the above

48 Which of the following clustering requires merging approach?

a) Partitional b) Hierarchical c) Naive Bayes d) None of these

49 Which clustering algorithm is capable of handling noise and outliers effectively?

a) K-Means b) DBSCAN c) Agglomerative d) Mean-Shift

50 Suppose we are using a random forest algorithm to solve regression problem and there are 4 data points. The value returned by the model and the actual value for the data points 1, 2, 3, and 4 are 11, 14, 9, 10 and 8, 10, 12, 14 respectively. What is the mean squared error?

a) 11.5 b) 14 c) 12.5 d) 10

***