Reg. No.															

B.Tech. DEGREE EXAMINATION, DECEMBER 2023

Fourth & Seventh Semester

18AIO352T – MACHINE LEARNING

(For the candidates admitted from the academic year 2020-2021 & 2021-2022)

Note:

(i) **Part - A** should be answered in OMR sheet within first 40 minutes and OMR sheet should be handed over to hall invigilator at the end of 40th minute.

(ii) **Part - B & Part - C** should be answered in answer booklet.

Time: 3 hours Max. Marks: 100

PART – A (20 × 1 = 20 Marks)

Answer **ALL** Questions

	Marks	BL	CO	PO
1. Impact of high variance on the training set? (A) Under fitting (B) Over fitting (C) Both under fitting and over fitting (D) Depends upon the dataset	1	2	1	1
2. What type of regularization is used to reduce over fitting? (A) L1 (B) L2 (C) L1 and L2 (D) Logistic regression	1	2	1	1
3. Learning curve depicts the relationship between (A) Model complexity and training error (B) Training set size and model performance (C) Model performance and test set size (D) Regularization and test error	1	1	1	1
4. When training error and validation error are high, what can be inter from learning curve? (A) Model is over fitting (B) Model is under fitting (C) Model is good fit (D) Model is neither over fitting nor under fitting	1	1	1	1
5. Which of the following is a disadvantage of K-fold cross validation method? (A) The variance of the resulting estimate is reduced as K increases (B) Usually does not take longer time to compute (C) Reduced bias (D) Training algorithm has to return from scratch K times	1	2	2	5
6. Logistic regression lies between (A) 0 and 1 (B) 1 and 10 (C) 0 and 10 (D) 10 and 100	1	2	2	1

7. In linear regression with multiple variables, input has manly _____ parameters which are mapped to one predicated output. 1 1 2 5
   (A) Dependent (B) Average
   (C) Independent (D) Unpredicted

8. How many types of logistic regression is available? 1 3 2 2
   (A) 1 (B) 2
   (C) 3 (D) 4

9. The size of dataset which is not best suited for SVM. 1 2 3 5
   (A) Large size (B) Small size
   (C) Medium size (D) Size does not matter

10. SVM is very less effective when 1 3 3 2
   (A) Data is linearly separable (B) Data is clean and ready to use
   (C) Data is noisy and contains overlapping points (D) Data is non linear

11. Which of the following K value in KNN would minimize the LOOC validation accuracy? 1 4 3 3
   (A) 3 (B) 5
   (C) 2 (D) 1

12. Which tool is used to reduce the dimension of the datas? 1 1 3 3
   (A) Product component analysis (B) Principal component analysis
   (C) Principal common analysis (D) Pre complex analysis

13. How is initial centroid of K-means clustering algorithm is selected? 1 2 4 1
   (A) Randomly (B) Data distribution
   (C) Data labels (D) Data segregation

14. Which of the following clustering requires merging approach? 1 3 4 2
   (A) Partitional (B) Hierarchical
   (C) Naïve Bayes (D) Heterogonous

15. Which of the following algorithm is most sensitive to outliers? 1 3 4 5
   (A) K-means clustering (B) K-medians clustering
   (C) K-medoids clustering (D) K-modes clustering

16. What is minimum number of variables required to perform clustering? 1 2 4 2
   (A) 0 (B) 1
   (C) 2 (D) 3

17. In random forest, the target attributes indicates the value of? 1 1 5 2
   (A) Decision node (B) Leaf node
   (C) Path (D) Edge

18. A computational model inspired by structure and function of the human brain is? 1 2 5 1
   (A) KNN (B) ANN
   (C) K-means (D) SVM

19.	A basic computation of an ANN is (A) Axon (B) Neuron (C) Dendron (D) Nodes	1	3	5	1
20.	The common activation function used in neural networks is (A) Cosine (B) Sigmoid (C) Exponential (D) Logarithmic	1	2	5	5

PART – B (5 × 4 = 20 Marks)

Answer **ANY FIVE** Questions

		Marks	BL	CO	PO
21.	Describe about supervised learning with an example.	4	2	1	1
22.	Write a note on parametric model and give example for it.	4	3	1	1
23.	What are various python libraries available in machine learning?	4	4	2	2
24.	What do you mean by ridge regression and mention its use?	4	1	3	5
25.	Write a note on divisive clustering.	4	3	4	2
26.	Describe about decision tree with an example.	4	2	5	5
27.	Explain briefly about the types of artificial neural network available.	4	4	5	2

PART – C (5 × 12 = 60 Marks)

Answer **ALL** Questions

		Marks	BL	CO	PO
28. a.	Explain in detail about linear regression with an example.	12	5	1	1
	(OR)				
b.	Describe in detail about the various types of machine learning with an example.	12	4	1	1
29. a.	Explain the steps involved in K-fold cross validations. What do you mean by training data, testing data and validation data?	12	3	2	1
	(OR)				
b.	Explain in detail about performance metrics and its expressions.	12	2	2	2
30. a.	Explain in detail about KNN classification with an example.	12	3	3	3
	(OR)				
b.	What is the technique used for feature extraction, describe it in detail.	12	2	3	4
31. a.	Describe in detail about hierarchical clustering.	12	4	4	1
	(OR)				

b.	Describe about different aspects of clustering validation and explain the types of evaluation of output of clustering methods.	12	3	4	5
32. a.	Describe in detail about random forest algorithm with a neat diagram and list outs its features.	12	4	5	1
	(OR)				
b.	Elaborate about artificial neural network with a neat diagram.	12	3	5	5

* * * * *