Reg. No.														

B.Tech. DEGREE EXAMINATION, MAY 2023
Fourth Semester

18AIO352T – MACHINE LEARNING
(For the candidates admitted during the academic year 2018-2019 to 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. Output of training process in machine learning is _____. (A) Accuracy (B) Outcome (C) Machine learning model (D) Mean variance	1	2	1	1
2. _____ type of machine learning algorithm is suitable for predicting continuous dependent variance (A) Logistic regression (B) Linear regression (C) Decision tree classifier (D) KNN classifier	1	1	1	1
3. If a ML model output involves target variable then the model is called as _____. (A) Descriptive model (B) Predictive model (C) Reinforcement learning (D) Well defined	1	1	1	1
4. Machine learning model which is built on sample data is known as _____ data. (A) Transfer (B) Training (C) Clean (D) Testing	1	1	1	1
5. What type of regularization is used to reduce over fitting? (A) L2 (B) L1 (C) Both L1 and L2 (D) Linear regression	1	2	2	3
6. Which of the following does not come under logistic regression? (A) Binomial (B) Trinomial (C) Multinomial (D) Ordinal	1	2	2	2
7. Cross validation is used for (A) Estimating expected error (B) Helps in selecting best fit model (C) Avoiding over fit model (D) All the above	1	3	2	2
8. Logistic regression comes under _____ learning. (A) Supervised (B) Unsupervised (C) Semi-supervised (D) Reinforcement	1	1	2	2

9.	KNN is used for _____. (A) Classification (B) Regression (C) Approximation (D) Classification and regression	1	2	3	3
10.	PCA is used for (A) Over fitting case (B) Small data and new features (C) Regularizing (D) Reducing the dimensionality	1	1	3	2
11.	Which of the following is not used as distance function in KNN? (A) Euclidean distance (B) Median distance (C) Manhattan distance (D) Minkowski distance	1	2	3	1
12.	To calculate Eigen value and Eigen vector of the covariance matrix the expressions (A) $\|M * X\| = 0$ (B) $\|M \div X\| = 0$ (C) $\|M + X\| = 0$ (D) $\|M - X\| = 0$	1	3	3	3
13.	Divisive hierarchical clustering is _____ approach. (A) Bottom-up (B) Top-down (C) Middle to top (D) Middle to down	1	2	4	3
14.	Which of the following method determines the number of clusters in hierarchical clustering? (A) Elbow method (B) Silhoulte method (C) Gap statistic (D) All the above	1	3	4	3
15.	The disadvantage of single linkage in hierarchical clustering is? (A) It resulting long and stringy clusters (B) It is sensitive to noise (C) It results in spherical clusters (D) Computationally expensive	1	1	4	2
16.	K-mean clustering is _____ algorithm. (A) Supervised (B) Unsupervised (C) Reinforcement (D) Semi-supervised	1	2	4	5
17.	Which type of neural network is commonly used for recognition of image? (A) RNN (B) CNN (C) Multilayer perceptron (D) Radial basis function network	1	3	5	5
18.	The purpose of back propagation in neural network is? (A) To initialize the weights (B) To propagate input (C) To calculate error between predicted and actual output (D) To regularize	1	1	5	1
19.	Technique used to prevent over fitting in neural network is (A) Lasso (B) Ridge (C) Regularization (D) Dropout	1	2	5	2
20.	Neural network is used to solve (A) Non linear problems (B) Linear problems (C) Classification (D) All the above	1	3	5	3

PART – B (5 × 4 = 20 Marks)

Answer **ANY FIVE** Questions

		Marks	BL	CO	PO
21.	Define under fitting and over fitting.	4	2	1	1
22.	What do you mean by curse of dimensionality?	4	3	1	2
23.	Define precision and recall.	4	2	2	3
24.	Write a note on linear regression with multiple variables.	4	3	2	5
25.	Describe about Bayesian classifier.	4	4	3	3
26.	What do you mean by bi-clustering explain about it?	4	3	4	5
27.	How does a random forest algorithm work?	4	2	5	2

PART – C (5 × 12 = 60 Marks)

Answer **ALL** Questions

		Marks	BL	CO	PO
28. a.i.	Explain in detail about bias and variance.	6	3	1	2
ii.	Write a note on learning curve with a neat diagram.	6	2	1	1
	(OR)				
b.	Describe in detail about L1 and L2 regularization techniques.	12	3	1	2
29. a.	Explain in detail about logistic regression with an example.	12	4	2	2
	(OR)				
b.	What are various performance metrics available? Explain about each metrics.	12	4	2	5
30. a.	Explain in detail about principal component analysis.	12	3	3	2
	(OR)				
b.	Describe in detail about support vector machine and kernels with a neat diagram.	12	4	3	5
31. a.	Explain in detail about K-means clustering with a neat diagram.	12	4	4	5
	(OR)				
b.	Describe in detail about K-medoids clustering with a neat diagram.	12	3	4	2
32. a.	How does a decision tree algorithm works, explain with an example.	12	2	5	5
	(OR)				
b.i.	Write a note on CART.	6	3	5	2
ii.	What is inductive bias in decision tree?	6	2	5	1

* * * * *