A **Reg. No. :** | | | | | | | | | | | | |

Question Paper Code: U5E03

B.E./B.Tech. DEGREE EXAMINATION, NOV 2023

Fifth Semester

Artificial Intelligence & Data Science

21UAD503 - MACHINE LEARNING TECHNIQUES

(Regulations 2021)

Duration: Three hours Maximum: 100 Marks

Answer All Questions

PART A - (10 x 1 = 10 Marks)

1. What is the application of machine learning methods to a large database called? CO1-U

   (a) big data computing (b) internet of things

   (c) data mining (d) artificial intelligence

2. Machine learning is a subset of CO1-U

   (a) Artificial intelligence (b) Deep learning

   (c) Data learning (d) None of the above

3. Among the following options identify the one which is false regarding regression. CO1-U

   (a) It is used for the prediction (b) It is used for interpretation

   (c) It relates inputs to outputs (d) It discovers casual relationships

4. Analysis of ML algorithm needs CO1-U

   (a) Statistical learning theory (b) Computational learning theory

   (c) Both (a) and (b) (d) None of the above

5. The total types of the layer in radial basis function neural networks is _____ CO1-U

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

6. Machine Learning is a field of AI consisting of learning algorithms that CO1-U

(a) At executing some task (b) Over time with experience

(c) Improve their performance (d) All of the above

7. Identify the model which is trained with data in only a single batch. CO1-U

(a) online learning (b) batch learning (c) both (a) and (b) (d) none of the above

8. Among the following identification which one is the dimensionality reduction CO1-U

(a) performance (b) entropy (c) stochastics (d) collinearity

9. What does K stand for in K mean algorithm? CO1-U

(a) Number of clusters (b) Number of data

(c) Number of attributes (d) Number of iterations

10. Among the following option identify the one which is used to create the most common graph types. CO1-U

(a) plot (b) quickplot (c) qplot (d) All of the above

PART – B (5 x 2= 10Marks)

11. What do you understand by Reinforcement Learning technique? CO1-U

12. How Will You Know Which Machine Learning Algorithm to Choose for Your Classification Problem? CO2-App

13. What is the most basic type of neural network? CO1-U

14. What is Principal Component Analysis? CO1-U

15. Define the two components of Bayesian logic program? CO1-U

PART – C (5 x 16= 80Marks)

16. (a) You are working on a machine learning project where you need to pre-process the data before training various models. The dataset includes numerical, categorical, and text data, missing values, and outliers. Construct a complete data pre-processing pipeline, detailing each step and the techniques used to handle various data challenges. Explain how you ensure the pre-processing pipeline is scalable and maintainable as the dataset grows. CO2-App (16)

Or

(b) Working on a fraud detection model for credit card transactions. The dataset contains a small percentage of fraudulent transactions. The business requires a model to detect fraud accurately while minimizing false positives. How would you evaluate the model's performance in this scenario? Which evaluation metrics would you prioritize, and why? How would you balance high recall and low false positive rates? CO2-App (16)

17. (a) Suppose we are building a classifier that says whether a text is about sports or not. Our training data has 5 sentences: CO2-App (16)

Text	Tag
"A great game"	Sports
"The election was over"	Not sports
"Very clean match"	Sports
"A clean but forgettable game"	Sports
"It was a close election"	Not sports

Now, which tag does the sentence "A very close game" belong to? Explain how Naive Bayes can be employed to predict.

Or

(b) A training data set of weather and the corresponding target variable 'Play' (suggesting possibilities of playing). CO2-App (16)

	Outlook	Play		Outlook	Play
0	Rainy	Yes	7	Overcast	Yes
1	Sunny	Yes	8	Rainy	No
2	Overcast	Yes	9	Sunny	No
3	Overcast	Yes	10	Sunny	Yes
4	Sunny	No	11	Rainy	No
5	Rainy	Yes	12	Overcast	Yes
6	Sunny	Yes	13	Overcast	Yes

Solve if the weather is sunny, then the Player should play or not use Naive Bayes to predict.

18.	(a)	Explain briefly about the operation of biological neural network with a simple sketch.	CO1-U	(16)
		Or		
	(b)	Describe Activation Function and explain the types of functions	CO1-U	(16)
19.	(a)	What is supervised learning, and how does it differ from other machine learning paradigms?	CO1-U	(16)
		Or		
	(b)	What is a training dataset, and what is its role in supervised learning?	CO1-U	(16)
20.	(a)	What is a genetic algorithm (GA), and how does it relate to machine learning?	CO1-U	(16)
		Or		
	(b)	What is the difference between local and global optimization in the context of optimization algorithms?	CO1-U	(16)