Reg. No.																

Question Paper Code	13079

B.E. / B.Tech. - DEGREE EXAMINATIONS, NOV / DEC 2024

Third Semester

Artificial Intelligence and Data Science

(Common to Computer Science and Engineering (AIML))

20AIPC302 - FUNDAMENTALS OF MACHINE LEARNING TECHNIQUES

Regulations - 2020

Duration: 3 Hours | Max. Marks: 100

PART - A (MCQ) (20 × 1 = 20 Marks)

Answer ALL Questions

		Marks	K – Level	CO
1.	What is the main objective of machine learning? (a) To develop software with predefined rules (b) To enable computers to learn from data (c) To create manual algorithms for every task (d) To eliminate the need for data	1	K1	CO1
2.	Which type of learning involves learning to make decisions by receiving rewards or penalties? (a) Supervised learning (b) Unsupervised learning (c) Reinforcement learning (d) Semi-supervised learning	1	K1	CO1
3.	Which of the following algorithms is used for dimensionality reduction? (a) K-Nearest Neighbors (b) K-Means Clustering (c) Principal Component Analysis (d) Support Vector Machines	1	K1	CO1
4.	Which metric is commonly used to evaluate the performance of a classification model? (a) Mean Absolute Error (MAE) (b) Root Mean Squared Error (RMSE) (c) Accuracy (d) R-squared	1	K1	CO2
5.	What is the term for the process of correcting or removing errors and inconsistencies in data? (a) Data augmentation (b) Data normalization (c) Data cleaning (d) Data encoding	1	K1	CO2
6.	Which type of data represents values with meaningful numeric magnitudes, such as age or salary? (a) Categorical data (b) Numerical data (c) Text data (d) Time series data	1	K1	CO2
7.	What is the Hamming distance between binary vectors 10010110 and 01011001? (a) 3 (b) 4 (c) 5 (d) 6	1	K1	CO3
8.	When you find many noises in data, which of the following options would you consider in KNN? (a) Increase the value of k (b) Decrease the value of k (c) Noisedoes not depend on k (d) When K = 0	1	K1	CO3
9.	What is feature construction? (a) The process of selecting existing features. (b) The process of creating new features from existing data. (c) The process of removing irrelevant features. (d) The process of splitting data into training and testing sets.	1	K1	CO3
10.	Which of the following is essentially finding a relationship (or) association between the dependent variable (Y) and the independent variables (X)? (a) Slope (b) Regression (c) Classification (d) Categorization	1	K1	CO4
11.	When we are trying to predict a real-value variable such as '$', 'Weight', the problem falls under the category of_____ (a) Unsupervised learning (b) Supervised regression problem (c) Supervised classification problem (d) Categorical attribute	1	K1	CO4

12.	_____ is the extension of the simple linear model by raising (squaring) each of the original predictors to a power. (a) Multiple Linear Regression (b) Simple Linear Regression (c) Polynomial Regression (d) Ridge Regression	*1*	*K1*	*CO4*
13.	Which of the following is required for K-means clustering? (a) defined distance metric (b) number of clusters (c) initial guess as to cluster centroids (d) all of the mentioned	*1*	*K1*	*CO5*
14.	Which of the following clustering requires merging approach? (a) Partition (b) Hierarchical (c) Naive Baye's (d) None of the mentioned	*1*	*K1*	*CO5*
15.	Which clustering algorithm is based on the concept of centroids? (a) K-Means (b) DBSCAN (c) Agglomerative (d) Mean-Shift	*1*	*K1*	*CO5*
16.	Which unsupervised learning algorithm is used for market basket analysis? (a) Apriori algorithm (b) Decision tree (c) Naive Baye's (d) Linear Regression	*1*	*K1*	*CO5*
17.	What is the primary goal of Active Learning in machine learning? (a) Minimize model complexity (b) Maximize dataset size (c) Minimize annotation cost (d) Maximize overfitting	*1*	*K1*	*CO6*
18.	Instance-based learning is also known as ____ (a) Memory-based learning (b) Lazy-learning (c) Memory-based learning, Lazy-learning (d) None of the mentioned	*1*	*K1*	*CO6*
19.	Which ensemble method combines multiple weak learners to create a strong learner by adjusting the weights of misclassified instances? (a) Bagging (b) Boosting (c) Stacking (d) Random Forest	*1*	*K1*	*CO6*
20.	What type of penalty does Ridge regression add to the loss function? (a) L1 penalty (b)L2 penalty (c) Elastic Net penalty (d) No penalty	*1*	*K1*	*CO6*

PART - B (10 × 2 = 20 Marks)

Answer ALL Questions

21.	What is Machine learning? What is the need of it?	*2*	*K1*	*CO1*
22.	List out any 5 real-time applications of Machine Learning.	*2*	*K1*	*CO1*
23.	How Classification varies from regression?	*2*	*K1*	*CO2*
24.	Why is the Kappa value used in Classification?	*2*	*K1*	*CO2*
25.	What is a Confusion Matrix?	*2*	*K1*	*CO3*
26.	What is Feature Engineering?	*2*	*K1*	*CO3*
27.	State Gauss Markov Theorem.	*2*	*K1*	*CO4*
28.	What are Rise and Run with respect to slope?	*2*	*K1*	*CO4*
29.	Define Association Rule.	*2*	*K1*	*CO5*
30.	List out the Partitioning methods that are involved in the Clustering Process.	*2*	*K1*	*CO6*

PART - C (6 × 10 = 60 Marks)

Answer ALL Questions

31.	a)	Explain in detail about the types of Machine learning with necessary diagrams.	*10*	*K2*	*CO1*
		OR			
	b)	Brief about the applications of Machine learning in the Healthcare domain.	*10*	*K2*	*CO1*
32.	a)	Explain in detail about Box plots and its types.	*10*	*K2*	*CO2*
		OR			
	b)	Discuss about the data pre-processing steps in Machine learning. Explain the steps involved with a neat sketch.	*10*	*K2*	*CO2*

33. a) Explain in detail about Feature engineering in detail. *10 K2 CO3*

OR

b) Discuss Random Forest model in detail. What are the strengths and weaknesses of it? *10 K2 CO3*

34. a) Explain multiple linear regressions with an example. *10 K2 CO4*

OR

b) Explain the assumptions in regression analysis and the BLUE concept. *10 K2 CO4*

35. a) Explain the Apriori algorithm for association rule learning with an example. *10 K2 CO5*

OR

b) Explain the types of Partitioning Clustering Algorithms in detail. *10 K2 CO5*

36. a) What is an Ensemble Learning Algorithm? Discuss various types. *10 K2 CO6*

OR

b) Solve the below mentioned 10 data points using K-Means Clustering problem. *10 K3 CO6*

	Height	Weight
1	185	72
2	170	56
3	168	60
4	179	68
5	182	72
6	188	77
7	180	71
8	180	70
9	183	84
10	180	88