



K. K. Wagh Institute of Engineering Education & Research, Nashik
(An Autonomous Institute From A.Y. 2022-23)

SUMMER-2024	
Exam Seat No.:	
Academic Year:2023-2024	Semester:III
Class: SY	Program:MCA
Branch Code:M.C.A.	Pattern:2022
Name of Course:Machine Learning	Course Code:MCA223002
Max. Marks:60	Duration:2.30 Hrs.

Instructions: Candidates should read carefully the instructions printed on the Question Paper and on the cover page of the Answer Book, which is provided for their use.

1. This question paper contains 2 page(s).
2. Answer to each new question is to be started on a new page.
3. Assume suitable data wherever required, but justify it.
4. Draw the neat labelled diagrams, wherever necessary.
5. The last columns indicates the Course Outcome and level of Blooms Taxonomy of the Question/sub-question.

Question No. 1 Attempt following Question

- 1a) Differentiate between Supervised, Unsupervised and Reinforcement Learning (6) CO1

Question No. 2 Attempt following Question

- 2a) Demonstrate principal components analysis as a method of dimensionality reduction (6) CO2

Question No. 3 Attempt following Question

- 3a) Demonstrate K- Nearest Neighbour classification algorithm with suitable examples (8) CO3

OR

- 3b) Underfitting and overfitting are two problems that degrade the performance of machine learning models. Justify your answer with suitable graphs (8) CO3

- 3c) Demonstrate the Naïve Bayes Classifier with an Example (8) CO3

OR

- 3d) Suppose Classifier's prediction is given as follows. Calculate Accuracy, Precision, Recall and F1 score (8) CO3

	Predicted		
Actual		Yes	No
	Yes	60	15
	No	10	15

Question No. 4 Attempt following Question

- 4a) What is hierarchical clustering? Illustrate Agglomerative Nesting and Divisive analysis methods with examples (8) CO4

OR

- 4b) Illustrate that K-means algorithm is based on the use of squared Euclidean distance as the measure of dissimilarity between a data point and a prototype vector (8) CO4

- 4c) What is the drawback of the K-means clustering algorithm? Illustrate K-medoid algorithm (8) CO4

OR

- 4d) Demonstrate Euclidean, Manhattan and Minkowski Distance Metric (8) CO4

Question No. 5 Attempt following Question

- 5a) What is the need of Data mining? Illustrate issues in data mining (8) CO5

OR

- 5b) Illustrate steps involved in the process of knowledge discovery from data(KDD) (8) CO5

- 5c) What is the drawback of Apriori algorithm? Illustrate how it is overcome in the FP-growth algorithm (8) CO5

OR

- 5d) State the Apriori principle. Apply Apriori algorithm to find the frequent itemsets with support threshold=2 for the following data: (8) CO5

Transaction ID	Items
T1	{M, O, N, K, E, Y}
T2	{D, O, N, K, E, Y}
T3	{C, O, O, K, I, E}
T4	{C, A, K, E}
T5	{D, U, C, K}
T6	{R, A, C, K}

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX