CBCS SCHEME

USN | | | | | | | | | | |

BCS602

Sixth Semester B.E./B.Tech. Degree Examination, June/July 2025
Machine Learning

Time: 3 hrs. Max. Marks: 100

Note: 1. Answer any FIVE full questions, choosing ONE full question from each module.
2. M : Marks , L: Bloom's level , C: Course outcomes.

PESIT LIBRARY BANGALORE

		Module – 1	M	L	C
Q.1	a.	State Tom Mitchell's definition of machine learning. List and explain the challenges of machine learning.	7	L1	CO1
	b.	List and explain the visualization aids available for univariate data analysis with example for each.	7	L2	CO1
	c.	For the patients age list {12, 14, 19, 22, 24, 26, 28, 31, 34}. Find the IQR.	6	L3	CO1
		OR			
Q.2	a.	Explain in detail the machine learning process with a neat diagram.	7	L2	CO1
	b.	Explain data preprocessing with measures to solve the problem of missing data.	7	L2	CO1
	c.	Find the 5-point summary of the list {13, 11, 2, 3, 4, 8, 9} and plot the box plot for the same.	6	L3	CO1
		Module – 2			
Q.3	a.	Let the data points be $\begin{pmatrix}2\\6\end{pmatrix}$ and $\begin{pmatrix}1\\7\end{pmatrix}$. Apply Principal Component Analysis (PCA) and find the transformed data.	10	L3	CO1
	b.	Apply candidate elimination algorithm on the dataset given in Table Q.3(b) to obtain the complete version space.	10	L3	CO2

Table Q.3(b):

CGPA	Interactiveness	Practical knowledge	Communication skills	Logical thinking	Job offer
≥ 9	Yes	Excellent	Good	Fast	YES
≥ 9	Yes	Good	Good	Fast	YES
≥ 8	No	Good	Good	Fast	NO
≥ 9	Yes	Good	Good	Slow	YES

		OR			
Q.4	a.	Find Singular Value Decomposition (SVD) of the matrix $A = \begin{pmatrix}1 & 2\\4 & 9\end{pmatrix}$.	10	L3	CO2

b. Write Find-S algorithm. Apply the algorithm to obtain the hypothesis for the dataset given in the Table Q.4(b). — 10 | L3 | CO2

Table Q.4(b)

Sky	Air temp	Humidity	Wind	Water	Forecast	Enjoy sport
Sunny	Warm	Normal	Strong	Warm	Same	YES
Sunny	Warm	High	Strong	Warm	Same	YES
Rainy	Cold	High	Strong	Warm	Change	NO
Sunny	Warm	High	Strong	Cool	Change	YES

Module – 3

Q.5 a. Apply K-nearest neighbor algorithm, for the dataset given in Table Q.5(a). Given a test instance (6.1, 40, 5), use the training set to classify the test instance. Choose K = 3. — 6 | L3 | CO3

Table Q.5(a)

CGPA	Assessment	Project submitted	Result
9.2	85	8	PASS
8	80	7	PASS
8.5	81	8	PASS
6	45	5	FAIL
6.5	50	4	FAIL
5.8	38	5	FAIL

b. Explain types of regression methods and limitations of regression methods. — 7 | L2 | CO3

c. Explain the structure of a decision tree and write the procedure to construct a decision the using ID3 algorithm. — 7 | L2 | CO3

OR

Q.6 a. Write the nearest-centroid classifier algorithm. Apply the same to predict the class for the given test instance (6, 5) using the training dataset given in Table Q.6(a). — 7 | L3 | CO3

X	Y	Class
3	1	A
5	2	A
4	3	A
7	6	B
6	7	B
8	5	B

Table Q.6(a)

b. Distinguish between — 6 | L2 | CO3

i) Regression and correlation

ii) Regression and causation

iii) Linearity and non-linearity relationships.

c. Explain the advantages and disadvantages of decision tree. Write the general algorithm for decision tree. — 7 | L2 | CO3

Module – 4

Q.7	**a.**	Using Naïve bayes classifier classify the new data (Red, SUV, Domestic) using the training dataset given in Table Q.7(a). Table Q.7(a) Color | Type | Origin | Stolen Red | Sports | Domestic | YES Red | Sports | Domestic | NO Red | Sports | Domestic | YES Yellow | Sports | Domestic | NO Yellow | Sports | Imported | YES Yellow | SUV | Imported | NO Yellow | SUV | Imported | YES Yellow | SUV | Domestic | NO Red | SUV | Imported | NO Red | Sports | Imported | YES	**10**	**L3**	**CO4**
	b.	Explain the simple model of an artificial neuron along with the artificial neural network structure.	**10**	**L2**	**CO4**

OR

Q.8	**a.**	Explain Bayes theorem, Maximum A Posteriori (MAP) hypothesis and Maximum Likelihood (ML) hypothesis in detail.	**10**	**L2**	**CO4**
	b.	Explain different activation functions used in artificial neural network.	**10**	**L2**	**CO4**

Module – 5

Q.9	**a.**	Consider the following set of data given in Table Q.9(a). Cluster it using K-means algorithm with initial value of objects 2 and 5 with the coordinate values (4, 6) and (12, 4) as initial seeds. Table Q.9(a) Objects | X-coordinate | Y-coordinate 1 | 2 | 4 2 | 4 | 6 3 | 6 | 8 4 | 10 | 4 5 | 12 | 4	**10**	**L3**	**CO5**
	b.	Explain the various components of reinforcement learning.	**10**	**L2**	**CO5**

OR

Q.10	**a.**	Find the Manhattan and Chebyshev distance if the coordinates of the objects are (0, 3) and (5, 8).	**4**	**L3**	**CO5**
	b.	Explain the mean shift clustering algorithm.	**6**	**L2**	**CO5**
	c.	List and explain the i) Characteristics of reinforcement learning ii) Challenges of reinforcement learning iii) Applications of reinforcement learning	**10**	**L3**	**CO5**

* * * * *