

Reg. No.:

Name :

**VIT**

Vellore Institute of Technology

(Approved by the Tamil Nadu State Council of Higher Education, Chennai, in the year 1984)

Continuous Assessment Test-II – October 2022

Programme :	B.Tech CSE (Artificial Intelligence and Robotics)	Semester :	Fall 2022-23
Course :	Machine Learning and its Applications	Code :	CSE3105
Faculty :	Dr.J. Priyadarshini	Slot(s) :	B2
		Class Nbr(s) :	CH2022231000531
Time :	1½ Hours	Max. Marks :	50

ANSWER ALL THE QUESTIONS

1

a)

Consider the training and testing data given in Table 1 and Table 2, respectively.

10

Table 1: Training Data

Sl. No	A1	A2	A3	C
1	{A1}	{A2}	{A3}	C1
2	{A1}	{A2}	{A3}	C1
3	{A1}	{A2}	{A3}	C1
4	{A1}	{A2}	{A3}	C1
5	{A1}	{A2}	{A3}	C2
6	{A1}	{A2}	{A3}	C2
7	{A1}	{A2}	{A3}	C3
8	{A1}	{A2}	{A3}	C3
9	{A1}	{A2}	{A3}	C3
10	{A1}	{A2}	{A3}	C3

Table 2: Test data

Sl. No	A1	A2	A3	Predicted Class
1	{A1}	{A2}	{A3}	?

{A1}, {A2}, {A3} – You can replace it between the given range in each transaction of training and test data

{A1} – {6/7/8}, {A2} – {4/5}, {A3} – {1/2/3} // Represents: categorical data.

For eg: Training Data

Sl. No	A1	A2	A3	C
1	7	5	1	C1
2	6	4	2	C1

Find the class label for the given test instance in table 2 using naïve bayes approach

b)

Consider all the points of any two attributes and all the points from any two classes from Table 1 and discover the support vectors that accurately discriminates the two classes.

10

2

a)

Table 3: Dataset

#	Attribute			
	A1	A2	A3	A4
1	{A1}	{A2}	{A3}	{A4}
2	{A1}	{A2}	{A3}	{A4}
3	{A1}	{A2}	{A3}	{A4}
4	{A1}	{A2}	{A3}	{A4}
5	{A1}	{A2}	{A3}	{A4}
6	{A1}	{A2}	{A3}	{A4}

Consider Table 3, replace {A1} to {A4} values ranging from 5.1 to 9.9, choose the initial centroids optimally and form suitable Clusters using K-Means clustering algorithm.

Eg:

	A1	A2	A3	A4
1	5.1	5.5	8.0	7.2
2	7.5	7.0	8.5	8

Note:

- Assume the K value
- If final clusters couldn't be formed till three iterations, stop the iteration and specify the clusters after Iteration 3.

b) Apply the KNN algorithm to form Clusters using the data specified in Table 3. Assume the threshold optimally

10

c) Consider the data given in Table 3 and perform the following

- Apply complete link agglomerative hierarchical clustering (7 Marks)
- Illustrate the clustering with the dendrogram by specifying the distance (3 Marks)

10



Reg. No.:

Name :

**VIT**

Vellore Institute of Technology

(Approved by the Tamil Nadu State Council of Higher Education, Chennai, India)

Continuous Assessment Test-II – October 2022

Programme :	B.Tech CSE (Artificial Intelligence and Robotics)	Semester :	Fall 2022-23
Course :	Machine Learning and its Applications	Code :	CSE3105
Faculty :	Dr.J. Priyadarshini	Slot(s) :	B2
		Class Nbr(s) :	CH2022231000531
Time :	1½ Hours	Max. Marks :	50

ANSWER ALL THE QUESTIONS

1

a)

Consider the training and testing data given in Table 1 and Table 2, respectively.

10

Table 1: Training Data

Sl. No	A1	A2	A3	C
1	{A1}	{A2}	{A3}	C1
2	{A1}	{A2}	{A3}	C1
3	{A1}	{A2}	{A3}	C1
4	{A1}	{A2}	{A3}	C1
5	{A1}	{A2}	{A3}	C2
6	{A1}	{A2}	{A3}	C2
7	{A1}	{A2}	{A3}	C3
8	{A1}	{A2}	{A3}	C3
9	{A1}	{A2}	{A3}	C3
10	{A1}	{A2}	{A3}	C3

Table 2: Test data

Sl. No	A1	A2	A3	Predicted Class
1	{A1}	{A2}	{A3}	?

{A1}, {A2}, {A3} – You can replace it between the given range in each transaction of training and test data

{A1} – {6/7/8}, {A2} – {4/5}, {A3} – {1/2/3} // Represents: categorical data.

For eg: Training Data

Sl. No	A1	A2	A3	C
1	7	5	1	C1
2	6	4	2	C1

Find the class label for the given test instance in table 2 using naïve bayes approach

b)

Consider all the points of any two attributes and all the points from any two classes from Table 1 and discover the support vectors that accurately discriminates the two classes.

10

2

a)

Table 3: Dataset

#	Attribute			
	A1	A2	A3	A4
1	{A1}	{A2}	{A3}	{A4}
2	{A1}	{A2}	{A3}	{A4}
3	{A1}	{A2}	{A3}	{A4}
4	{A1}	{A2}	{A3}	{A4}
5	{A1}	{A2}	{A3}	{A4}
6	{A1}	{A2}	{A3}	{A4}

Consider Table 3, replace {A1} to {A4} values ranging from 5.1 to 9.9, choose the initial centroids optimally and form suitable Clusters using K-Means clustering algorithm.

Eg:

	A1	A2	A3	A4
1	5.1	5.5	8.0	7.2
2	7.5	7.0	8.5	8

Note:

- Assume the K value
- If final clusters couldn't be formed till three iterations, stop the iteration and specify the clusters after Iteration 3.

b) Apply the KNN algorithm to form Clusters using the data specified in Table 3. Assume the threshold optimally

10

c) Consider the data given in Table 3 and perform the following

- Apply complete link agglomerative hierarchical clustering (7 Marks)
- Illustrate the clustering with the dendrogram by specifying the distance (3 Marks)

10

