


<https://swayam.gov.in>

[https://swayam.gov.in/nc\\_details/NPTEL](https://swayam.gov.in/nc_details/NPTEL)

rosy.prabu@gmail.com ▾

NPTEL (<https://swayam.gov.in/explorer?ncCode=NPTEL>) » Introduction to Machine Learning (course)

If already  
registered, click  
to check your  
payment status

Course  
outlineAbout  
NPTEL ()How does an  
NPTEL  
online  
course  
work? ()

Week 0 ()

Week 1 ()

Week 2 ()

Week 3 ()

Week 4 ()

Week 5 ()

Week 6 ()

Week 7 ()

## Week 10 : Assignment 10

The due date for submitting this assignment has passed.

Due on 2025-10-01, 23:59 IST.

### Assignment submitted on 2025-09-27, 15:21 IST

1) In a clustering evaluation, a cluster C contains 50 data points. Of these, 30 belong to class A, 15 to class B, and 5 to class C. What is the purity of this cluster? **1 point**

- ☐ 0.5  
☒ 0.6  
☐ 0.7  
☐ 0.8

Yes, the answer is correct.

Score: 1

Accepted Answers:

0.6

2) Consider the following 2D dataset with 10 points: **1 point**

(1, 1), (1, 2), (2, 1), (2, 2), (3, 3),  
(8, 8), (8, 9), (9, 8), (9, 9), (10, 10)

Using DBSCAN with  $\epsilon = 1.5$  and MinPts = 3, how many core points are there in this dataset?

- ☐ 4  
☐ 5  
☒ 8  
☐ 10

Yes, the answer is correct.

Score: 1

Accepted Answers:

8



**Week 8 ()****Week 9 ()****Week 10 ()**

- ☐ Partitional Clustering (unit? unit=113&lesson=114)
- ☐ Hierarchical Clustering (unit? unit=113&lesson=115)
- ☐ The BIRCH Algorithm (unit? unit=113&lesson=116)
- ☐ The CURE Algorithm (unit? unit=113&lesson=117)
- ☐ Density Based Clustering (unit? unit=113&lesson=118)
- ☐ Week 10 Feedback Form: Introduction to Machine Learning!! (unit? unit=113&lesson=291)
- ☐ Practice: Week 10 : Practice Assignment 10 (assessment? name=337)
- ☒ **Quiz: Week 10 : Assignment 10 (assessment? name=338)**

3) In BIRCH, using number of points **N**, sum of points **SUM** and sum of squared points **1 point SS**, we can determine the centroid and radius of the combination of any two clusters A and B. How do you determine the radius of the combined cluster? (In terms of **N, SUM** and **SS** of both two clusters A and B)

Radius of a cluster is given by:

$$Radius = \sqrt{\frac{SS}{N} - \left(\frac{SUM}{N}\right)^2}$$

Note: We use the following definition of radius from the BIRCH paper: "Radius is the average distance from the member points to the centroid."



$$Radius = \sqrt{\frac{SS_A}{N_A} - \left(\frac{SUM_A}{N_A}\right)^2 + \frac{SS_B}{N_B} - \left(\frac{SUM_B}{N_B}\right)^2}$$



$$Radius = \sqrt{\frac{SS_A}{N_A} - \left(\frac{SUM_A}{N_A}\right)^2} + \sqrt{\frac{SS_B}{N_B} - \left(\frac{SUM_B}{N_B}\right)^2}$$



$$Radius = \sqrt{\frac{SS_A + SS_B}{N_A + N_B} - \left(\frac{SUM_A + SUM_B}{N_A + N_B}\right)^2}$$



$$Radius = \sqrt{\frac{SS_A}{N_A} + \frac{SS_B}{N_B} - \left(\frac{SUM_A + SUM_B}{N_A + N_B}\right)^2}$$

Yes, the answer is correct.

Score: 1

Accepted Answers:

$$Radius = \sqrt{\frac{SS_A + SS_B}{N_A + N_B} - \left(\frac{SUM_A + SUM_B}{N_A + N_B}\right)^2}$$

4) Which of the following properties are TRUE?

1 point



Using the CURE algorithm can lead to non-convex clusters.



K-means scales better than CURE for large datasets.



CURE is a simplification of K-means and hence scales better than k-means for large datasets



K-means being more expensive to run on large datasets, can give non-convex clusters too.

Yes, the answer is correct.

Score: 1

Accepted Answers:

Using the CURE algorithm can lead to non-convex clusters.

5) The pairwise distance between 6 points is given below. Which of the option shows the hierarchy of clusters created by single link clustering algorithm? 1 point

	P1	P2	P3	P4	P5	P6
P1	0	3	8	9	5	4
P2	3	0	9	8	10	9
P3	8	9	0	1	6	7
P4	9	8	1	0	7	8
P5	5	10	6	7	0	2
P6	4	9	7	8	2	0

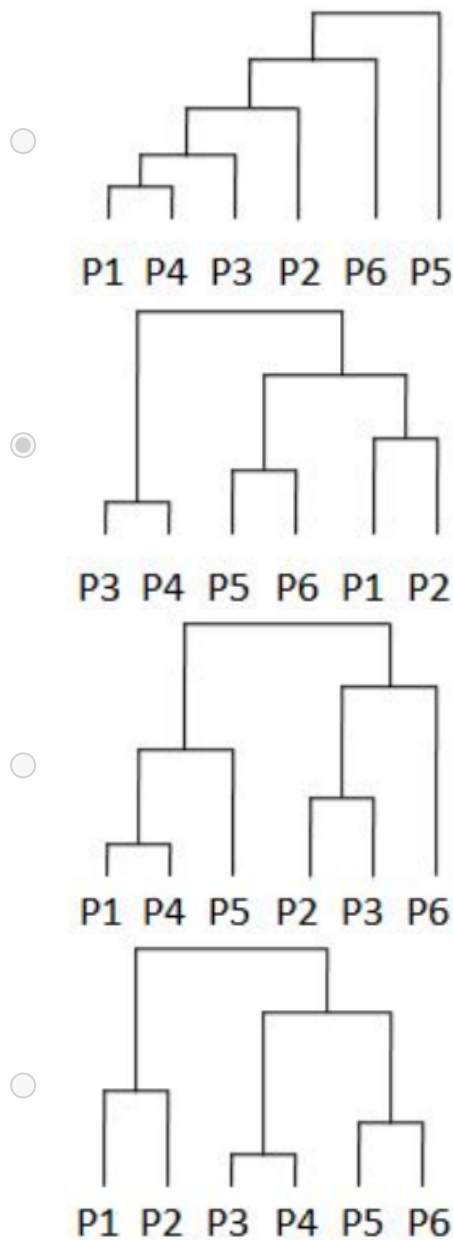
Week 11 ()

Week 12 ()

Solutions ()

Download  
Videos ()

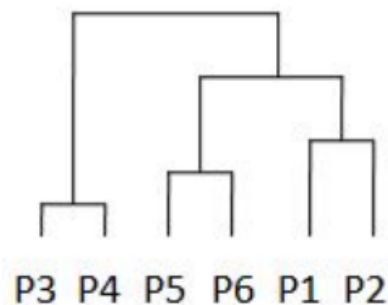
Problem  
Solving  
Session -  
July 2025 ()



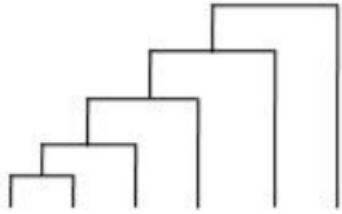
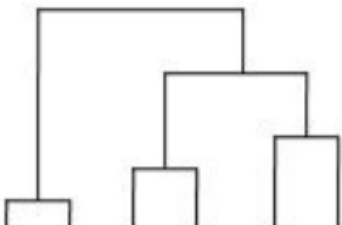
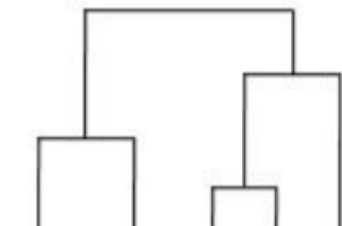
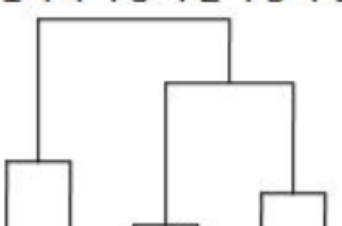
Yes, the answer is correct.

Score: 1

Accepted Answers:



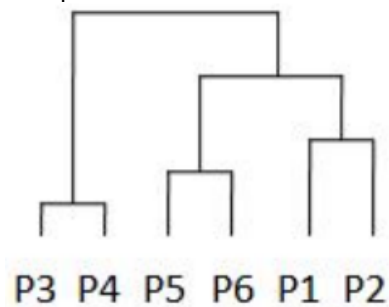
6) For the pairwise distance matrix given in the previous question, which of the following shows the hierarchy of clusters created by the complete link clustering algorithm. **1 point**

- ☐   
P1 P4 P3 P2 P6 P5
- ☒   
P3 P4 P5 P6 P1 P2
- ☐   
P1 P4 P5 P2 P3 P6
- ☐   
P1 P2 P3 P4 P5 P6

Yes, the answer is correct.

Score: 1

Accepted Answers:



For the following questions, we will be using the iris dataset that can be loaded using the following utility from sklearn:

`https://scikit-`

`learn.org/stable/modules/generated/sklearn.datasets.load_iris.html`

([https://scikit-learn.org/stable/modules/generated/sklearn.datasets.load\\_iris.html](https://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_iris.html))

Do not make any changes to the dataset unless directed in the question.



Set seed = 42 for numpy (`np.random.seed(seed)`).

Use `scikit-learn 1.0.2` to run your experiments.

7) Given a bag containing 6 red balls, 4 blue balls and 7 green balls, what is the probability that in 5 trials, at least 3 red balls are drawn from the bag?

**2 points**

- ☒ 0.24  
☐ 0.38  
☐ 0.17  
☐ 0.43

Yes, the answer is correct.

Score: 2

Accepted Answers:

0.24

8) Given a bag containing 6 red balls, 4 blue balls and 7 green balls, what is the probability that in 5 trials, at least 3 red balls are drawn from the bag?

**2 points**

- ☒ 0.24  
☐ 0.38  
☐ 0.17  
☐ 0.43

Yes, the answer is correct.

Score: 2

Accepted Answers:

0.24

