Q1-1: Consider we are working on an image classification problem. Which of the following could be considered as unlabeled data?

A. Vehicle images with the type of the vehicle
B. Fruit images with the height and width
C. Digit images with the class of the digit (0-9)
D. Furniture images with the name of the Furniture

Q1-1: Consider we are working on an image classification problem. Which of the following could be considered as unlabeled data?

A. Vehicle images with the type of the vehicle
B. Fruit images with the height and width 
C. Digit images with the class of the digit (0-9)
D. Furniture images with the name of the Furniture

The height and width of the fruit images are the features, not labels.

Q1-2: Which is true about machine learning?

A. The process doesn’t involve human inputs
B. The machine is given the training and test data for learning
C. In clustering, the training data also have labels for learning
D. Supervised learning involves labeled data

Q1-2: Which is true about machine learning?

A. The process doesn't involve human inputs
B. The machine is given the training and test data for learning
C. In clustering, the training data also have labels for learning
D. Supervised learning involves labeled data 

A. The labels are human inputs
B. The machine should not have test data for learning
C. No labels available for clustering

Q1-3: Which is true about feature vectors?

A. Feature vectors can have at most 10 dimensions
B. Feature vectors have only numeric values
C. The raw image can also be used as the feature vector
D. Text data don’t have feature vectors

Q1-3: Which is true about feature vectors?

A. Feature vectors can have at most 10 dimensions
B. Feature vectors have only numeric values
C. The raw image can also be used as the feature vector 
D. Text data don't have feature vectors

A. Feature vectors can be in high dimen.
B. Some feature vectors can have other types of values like strings
D. Bag-of-words is a type of feature vector for text

Q2-1: Which of the following is not a common task of unsupervised learning?

A. Clustering
B. Anomaly detection
C. Dimensionality reduction
D. Classification

Q2-1: Which of the following is not a common task of unsupervised learning?

A. Clustering
B. Anomaly detection
C. Dimensionality reduction
D. Classification

Q2-1: Which is true about the unsupervised learning tasks?

A. There are only 3 types of unsupervised learning tasks
B. Anomaly detection doesn’t have test data
C. PCA is a type of dimensionality reduction
D. Kmeans clustering is a type of hierarchical clustering

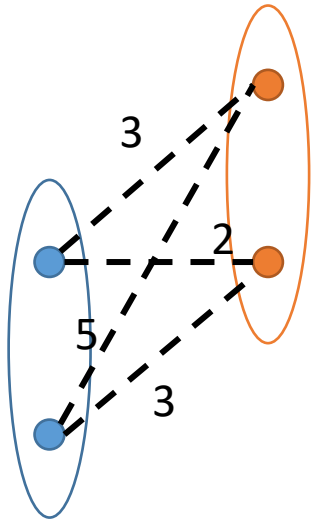
Q2-1: Which is true about the unsupervised learning tasks?

A. There are only 3 types of unsupervised learning tasks
B. Anomaly detection doesn't have test data
C. PCA is a type of dimensionality reduction 
D. Kmeans clustering is a type of hierarchical clustering

Notice that Anomaly detection also has test data

Q3-1: If we use single linkage to measure the distance from two clusters, what is the distance of these two clusters in the following example?

A. 2
B. 3
C. 5
D. 2.5



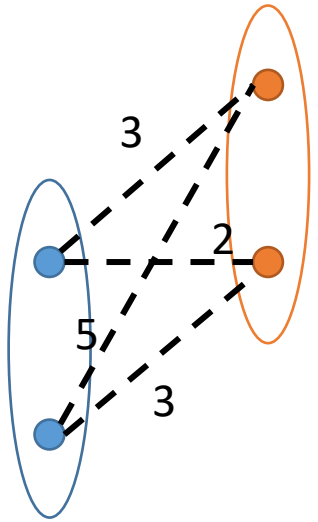
Q3-1: If we use single linkage to measure the distance from two clusters, what is the distance of these two clusters in the following example?

A. 2

B. 3

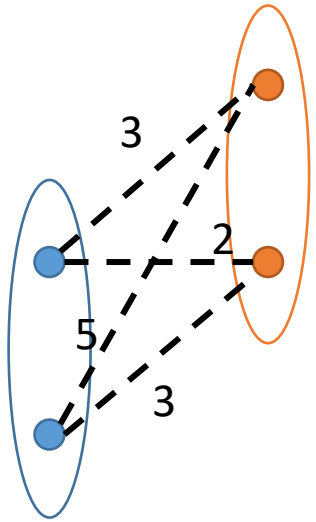
C. 5

D. 2.5



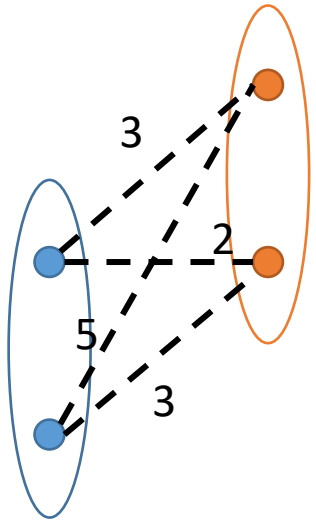
Q3-2: If we use complete linkage to measure the distance from two clusters, what is the distance of these two clusters in the following example?

A. 2
B. 3
C. 5
D. 2.5



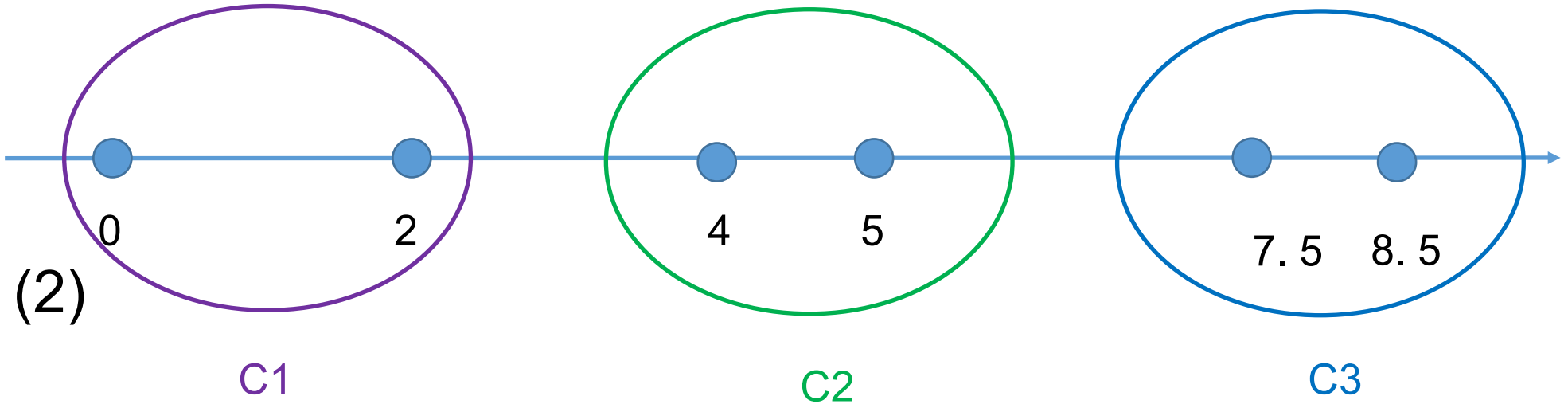
Q3-2: If we use complete linkage to measure the distance from two clusters, what is the distance of these two clusters in the following example?

A. 2
B. 3
C. 5
D. 2.5



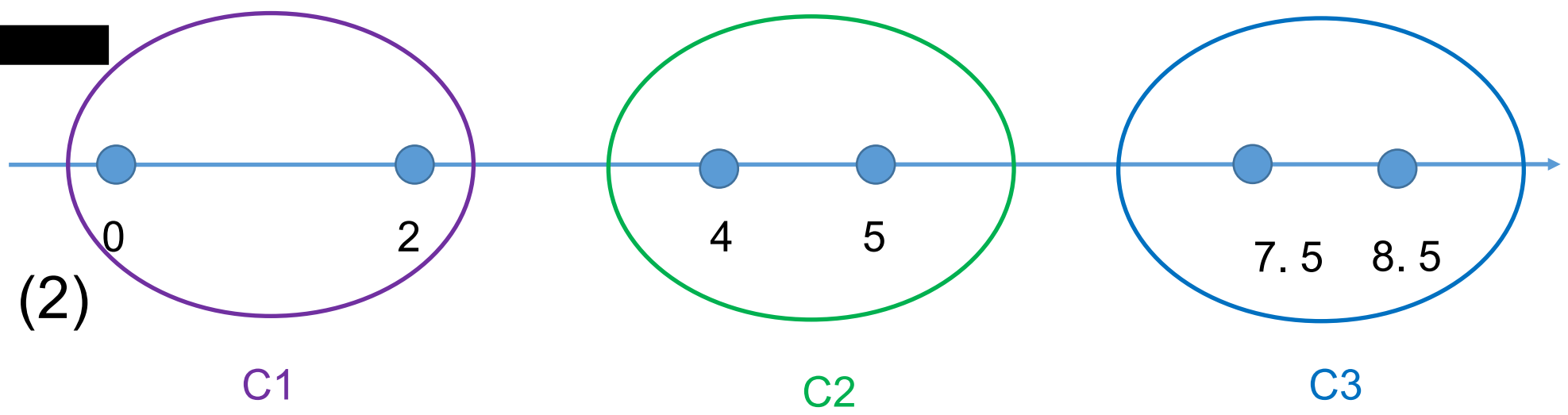
Q3-3: Consider the dataset in 1-dimension below. Now we have 3 clusters C1={0,2}, C2={4,5}, C3={7.5,8.5}.
(1) Single-linkage will merge C1 and C2.
(2) Complete-linkage will merge C1 and C2.
Which statement is true?

A. Only (1)
B. Only (2)
C. None
D. Both (1) and (2)



Q3-3: Consider the dataset in 1-dimension below. Now we have 3 clusters C1={0,2}, C2={4,5}, C3={7.5,8.5}.
(1) Single-linkage will merge C1 and C2.
(2) Complete-linkage will merge C1 and C2.
Which statement is true?

A. Only (1)
B. Only (2)
C. None
D. Both (1) and (2)



Single linkage: $d(C1, C2) = d(2,4) = 2, d(C2, C3) = d(5,7.5) = 2.5$
Complete linkage: $d(C1, C2) = d(0,5) = 5, d(C2, C3) = d(4,8.5) = 4.5$