Q1-1: K-NN algorithms can be used for:

1. Only classification
2. Only regression
3. Both

Q1-1: K-NN algorithms can be used for:

1. Only classification
2. Only regression
3. Both

Q1-2: Which of the following distance measure do we use in case of categorical variables in k-NN?

1. Hamming Distance
2. Euclidean Distance
3. Manhattan Distance

Q1-2: Which of the following distance measure do we use in case of categorical variables in k-NN?

1. Hamming Distance
2. Euclidean Distance
3. Manhattan Distance

Q1-3: Consider binary classification in 2D where the intended label of a point x = (x1, x2) is positive if x1>x2 and negative otherwise. Let the training set be all points of the form x = [4a, 3b] where a,b are integers. Each training item has the correct label that follows the rule above. With a 1NN classifier (Euclidean distance), which ones of the following points are labeled positive? Multiple answers.

1. [5.52, 2.41]
2. [8.47, 5.84]
3. [7 , 8.17]
4. [6.7 , 8.88]

Q1-3: Consider binary classification in 2D where the intended label of a point x = (x1, x2) is positive if x1>x2 and negative otherwise. Let the training set be all points of the form x = [4a, 3b] where a,b are integers. Each training item has the correct label that follows the rule above. With a 1NN classifier (Euclidean distance), which ones of the following points are labeled positive? Multiple answers.

1. [5.52, 2.41] ⬅
2. [8.47, 5.84] ⬅
3. [7 , 8.17]
4. [6.7 , 8.88]

Nearest neighbors are

[4,3] => positive

[8,6] => positive

[8,9] => negative

[8,9] => negative

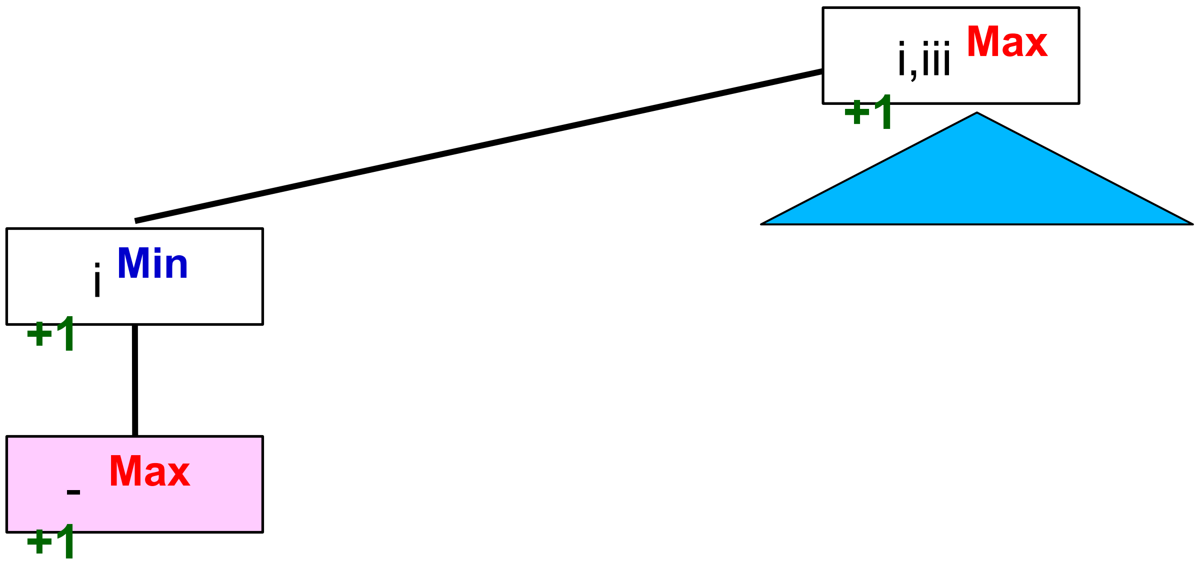
Individually.

Q2-1: Consider a variant of II Nim game where there are 2 piles, with 1 and 3 sticks, respectively. Each time one player takes some stick(s) from only one pile (can take 1 or 2 or 3 sticks). What’s the game theoretic value of the initial state?

A. 1
B. 0
C. -1

Q2-1: Consider a variant of II Nim game where there are 2 piles, with 1 and 3 sticks, respectively. Each time one player takes some stick(s) from only one pile (can take 1 or 2 or 3 sticks). What's the game theoretic value of the initial state?

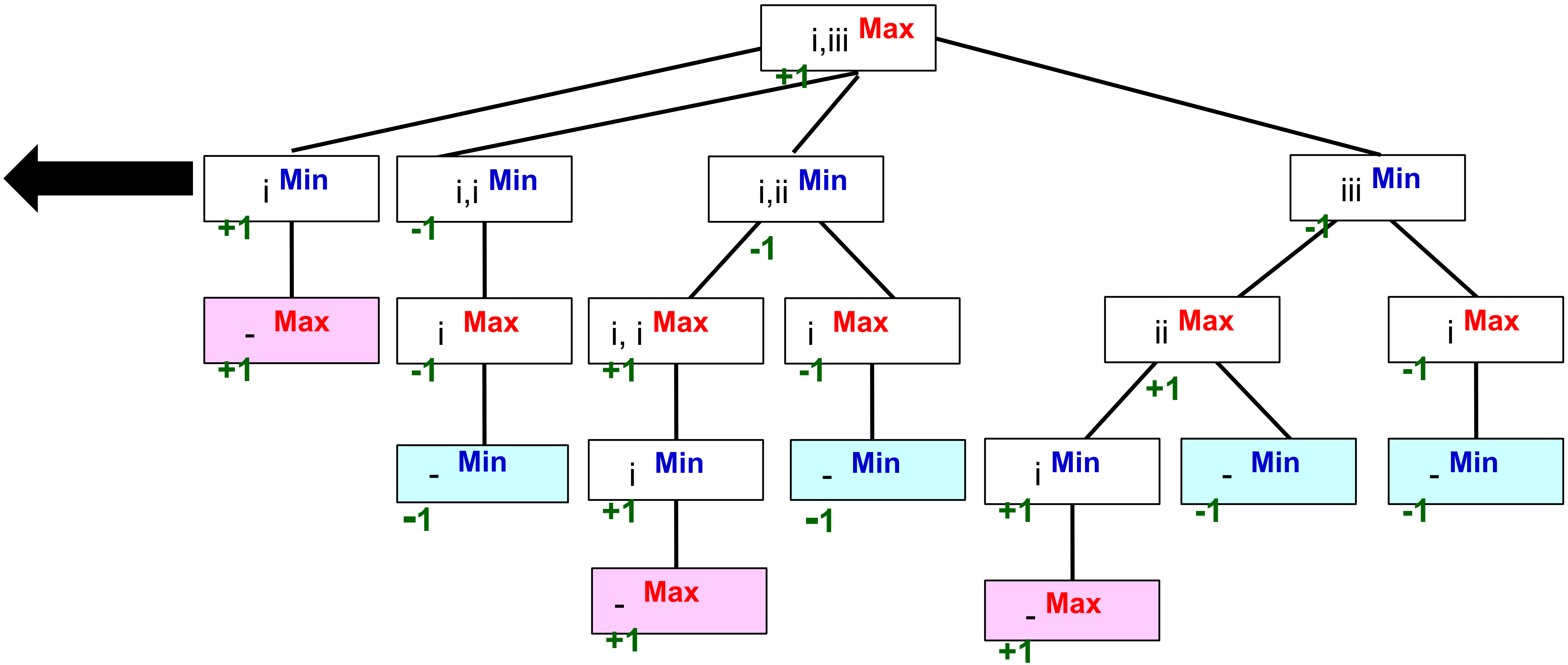
A. 1
B. 0
C. -1



The first player could always take 3 sticks from the pile with 3 sticks, which guarantees the game value +1. No need to check the other branches.

Q2-1: Consider a variant of II Nim game where there are 2 piles, with 1 and 3 sticks, respectively. Each time one player takes some stick(s) from only one pile (can take 1 or 2 or 3 sticks). What's the game theoretic value of the initial state?

A. 1
B. 0
C. -1

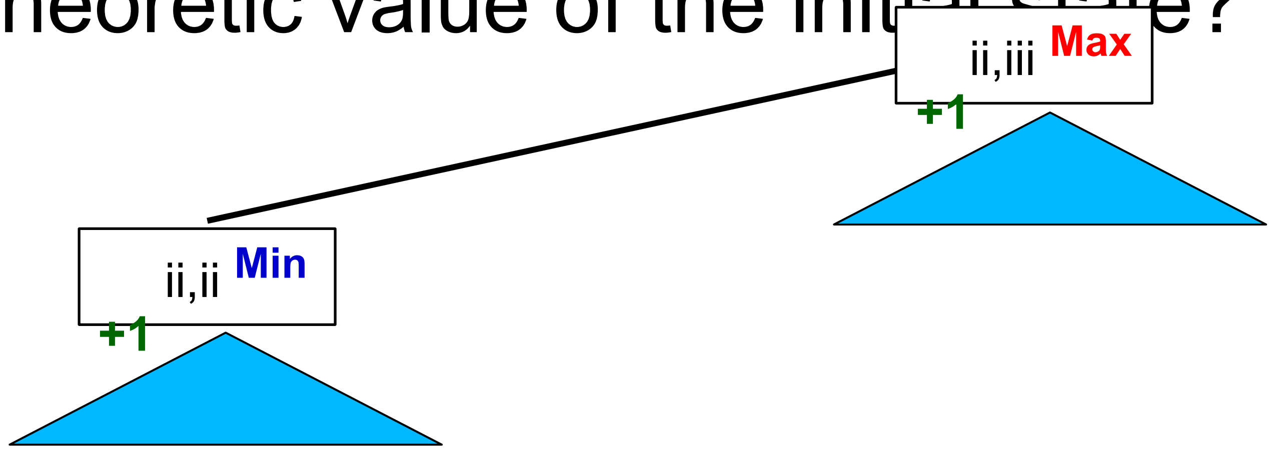


Q2-2: We know that the game theoretic value of the initial state is -1 in II Nim game where there are 2 piles, each with 2 sticks. Now consider a variant of II Nim game where there are 2 piles, with 2 and 3 sticks, respectively. Each time one player takes some stick(s) from only one pile (can take 1 or 2 or 3 sticks). What's the game theoretic value of the initial state?

A. 1
B. 0
C. -1

Q2-2: We know that the game theoretic value of the initial state is -1 in II Nim game where there are 2 piles, each with 2 sticks. Now consider a variant of II Nim game where there are 2 piles, with 2 and 3 sticks, respectively. Each time one player takes some stick(s) from only one pile (can take 1 or 2 or 3 sticks). What's the game theoretic value of the initial state?

A. 1
B. 0
C. -1



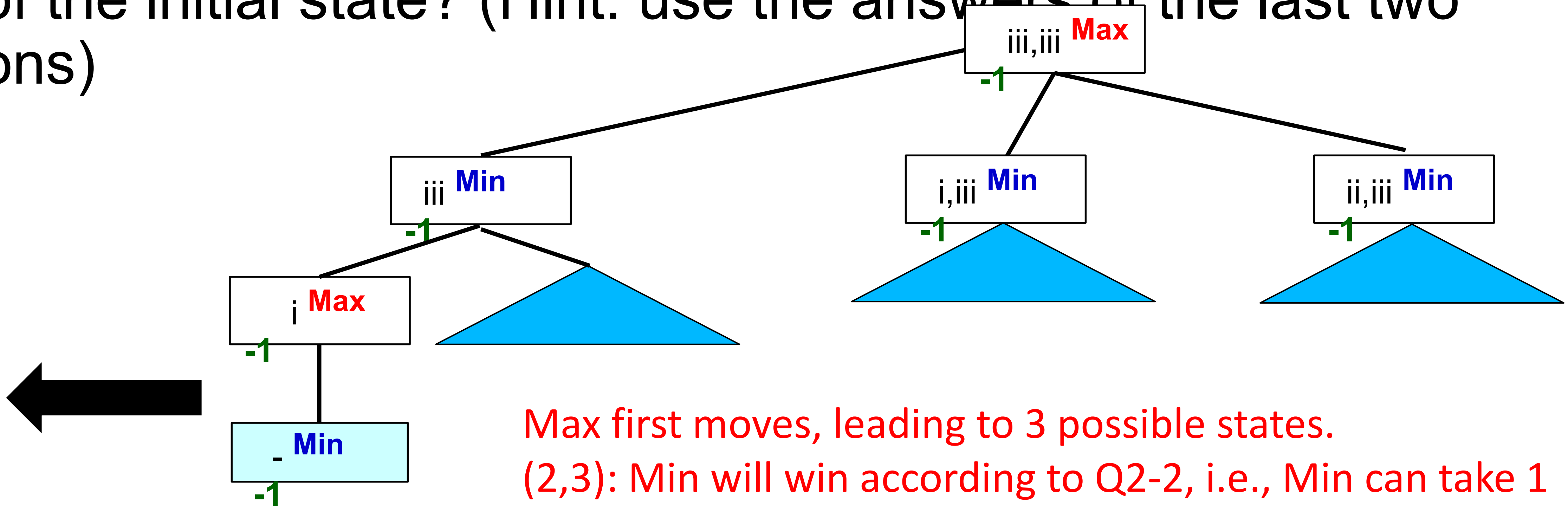
The first player could always take 1 stick from the pile with 3 sticks and make it a (2,2) state, in which case the second player will always lose.

Q2-3: Consider a variant of II Nim game with 2 piles, both with 3 sticks. Each time one player takes some stick(s) from only one pile (can take 1 or 2 or 3 sticks). What’s the game theoretic value of the initial state? (Hint: use the answers of the last two questions)

A. 1
B. 0
C. -1

Q2-3: Consider a variant of II Nim game with 2 piles, both with 3 sticks. Each time one player takes some stick(s) from only one pile (can take 1 or 2 or 3 sticks). What's the game theoretic value of the initial state? (Hint: use the answers of the last two questions)

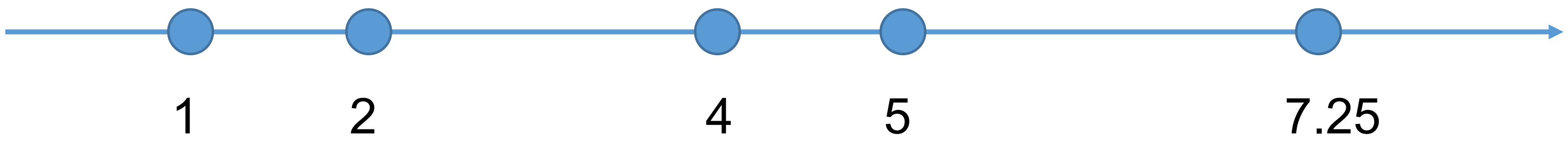
A. 1
B. 0
C. -1



Max first moves, leading to 3 possible states.
(2,3): Min will win according to Q2-2, i.e., Min can take 1 stick from the pile with 3 sticks and make it (2,2).
(1,3): Min will win according to Q2-1, i.e., Min can take all the second pile, and make it (1,-).
(0,3): Min can take 2 sticks and will win.

Q3-1: Suppose we run average linkage on the following data set to get two clusters. What is the result?

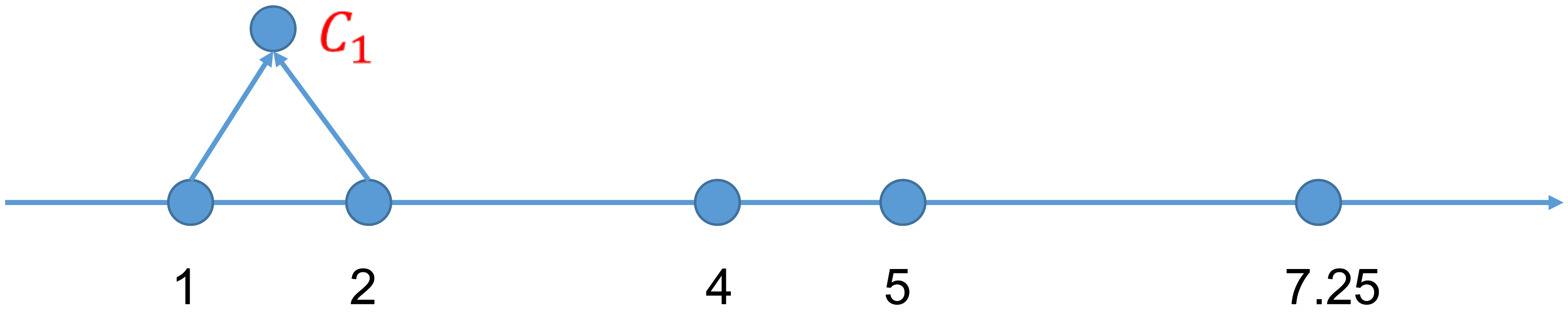
A. {1}, {2,4,5,7.25}
B. {1,2}, {4, 5, 7.25}
C. {1,2,4}, {5, 7.25}
D. {1,2,4,5}, {7.25}



Q3-1: Suppose we run average linkage on the following data set to get two clusters. What is the result?

A. {1}, {2,4,5,7.25}
B. {1,2}, {4, 5, 7.25}
C. {1,2,4}, {5, 7.25}
D. {1,2,4,5}, {7.25}

$$d(C_1, \{4\}) = \frac{3+2}{2} = 2.5,$$
$$d(\{4\},\{5\}) = 1$$

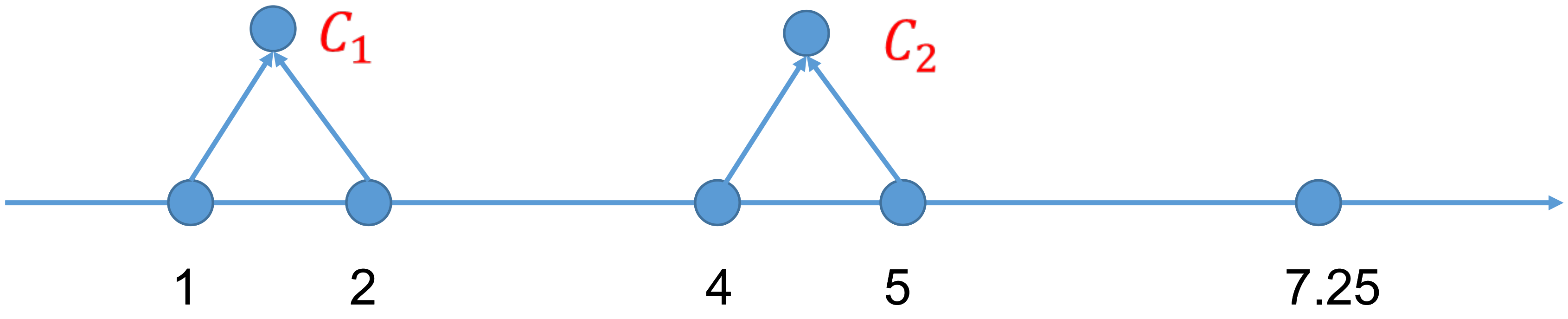


Q3-1: Suppose we run average linkage on the following data set to get two clusters. What is the result?

A. {1}, {2,4,5,7.25}
B. {1,2}, {4, 5, 7.25}
C. {1,2,4}, {5, 7.25}
D. {1,2,4,5}, {7.25}

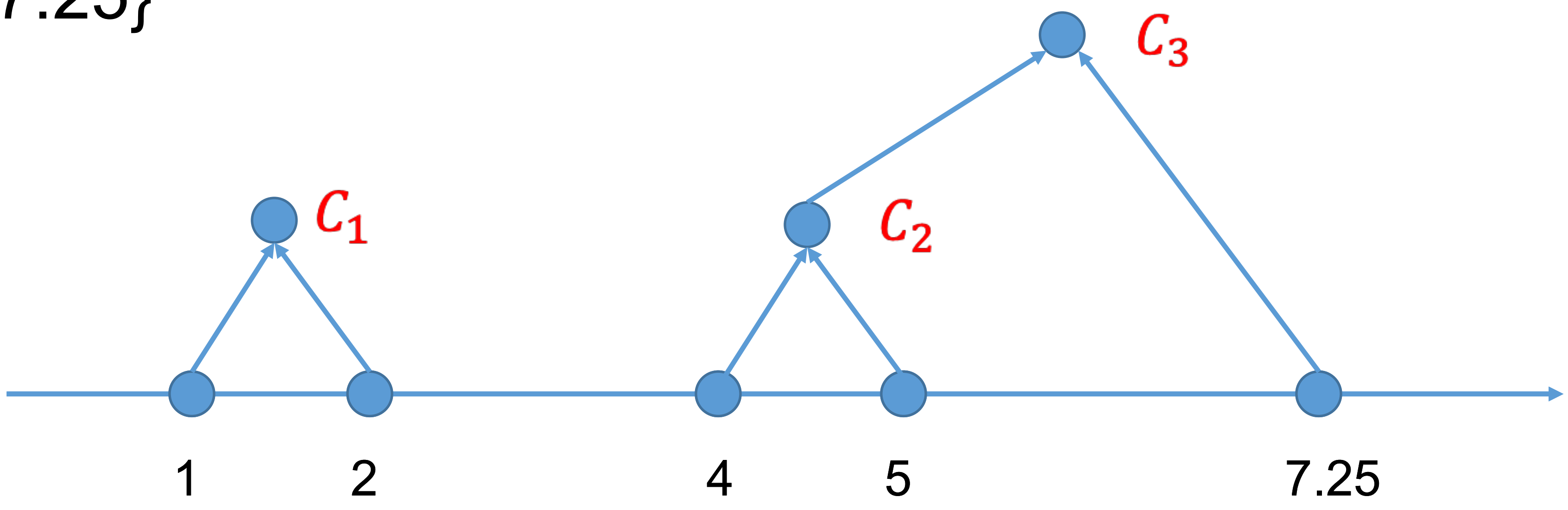
$$d(C_1, C_2) = \frac{3+2+4+3}{4} = 3,$$

$$d(C_2, \{7.25\}) = \frac{3.25+2.25}{2} = 2.75$$



Q3-1: Suppose we run average linkage on the following data set to get two clusters. What is the result?

A. {1}, {2,4,5,7.25}
B. {1,2}, {4, 5, 7.25}
C. {1,2,4}, {5, 7.25}
D. {1,2,4,5}, {7.25}



Q3-2: Assume, you want to cluster 7 points in 2-dim into 3 clusters using K-Means clustering algorithm. After the first iteration clusters, C1, C2, C3 has the following points:

C1: {(1,1), (3,3), (5,5)}

C2: {(0,4), (-2,4)}

C3: {(7,7), (9,9)}

What will be the cluster centroids if you want to proceed for the second iteration?

A. C1: (3,3), C2: (-1,4), C3: (8,8)
B. C1: (3,3), C2: (0,4), C3: (8,8)
C. C1: (3,3), C2: (-1,4), C3: (7,8)
D. C1: (2,2), C2: (-1,4), C3: (7,8)

Q3-2: Assume, you want to cluster 7 points in 2-dim into 3 clusters using K-Means clustering algorithm. After first iteration clusters, C1, C2, C3 has following points:

C1: {(1,1), (3,3), (5,5)}

C2: {(0,4), (-2,4)}

C3: {(7,7), (9,9)}

What will be the cluster centroids if you want to proceed for second iteration?

A. C1: (3,3), C2: (-1,4), C3: (8,8) ⟵
B. C1: (3,3), C2: (0,4), C3: (8,8)
C. C1: (3,3), C2: (-1,4), C3: (7,8)
D. C1: (2,2), C2: (-1,4), C3: (7,8)

Compute the average of the data points in each cluster

Q3-3: Which are true about linear regression?
1. When $\lambda \rightarrow +\infty$, ridge regression reduces to OLS.
2. The regression function must be linear in the original input features.
3. Gradient descent can be used to solve OLS.

A. 1,2
B. 1,3
C. 2,3
D. None of the above

Q3-3: Which are true about linear regression?
1. When $\lambda \to +\infty$, ridge regression reduces to OLS.
2. The regression function must be linear in the original input features.
3. Gradient descent can be used to solve OLS.

A. 1,2
B. 1,3
C. 2,3
D. None of the above

1: No. when $\lambda \to 0$, ridge regression reduces to OLS
2: No. It only needs to be linear in the parameter
3: Yes. Gradient descent is a general method for optimization.

Q3-4: Which are true about logistic regression?

1. When $\theta^T x_i = 0$, the model will predict label $+1$ with probability close to 1
2. There is ground-truth θ^* that can achieve 0 classification error on the training set
3. Gradient descent can be used to solve the regularized logistic regression problem

A. 1,2
B. 1,3
C. 2,3
D. None of the above

Q3-4: Which are true about logistic regression?

1. When $\theta^T x_i = 0$, the model will predict label $+1$ with probability close to 1
2. There is ground-truth θ^* that can achieve 0 classification error on the training set
3. Gradient descent can be used to solve the regularized logistic regression problem

A. 1,2

B. 1,3

C. 2,3

D. None of the above ←

1: No. predicts ½ for +1, ½ for -1.

2: No. Not 0 classification error. There is randomness in the labels.

3: Yes. Gradient descent is a general method for optimization.

Q3-5: Suppose we have $\theta = [0.5, 0.6, 1]$ for logistic regression. What label will the model predict for $x = [1, -5, 2]$?

A. -1
B. +1
C. Equal probabilities for -1 and +1
D. None of the above

Q3-5: Suppose we have $\theta = [0.5, 0.6, 1]$ for logistic regression. What label will the model predict for $x = [1, -5, 2]$?

A. -1
B. +1
C. Equal probabilities for -1 and +1
D. None of the above

The $\theta^T x = -0.5$, so p[y=+1|x]<0.5, so we will predict -1