CS461 – RECITATION 05

MACHINE LEARNING PRINCIPLES

Daize Dong

2025-10-13

TODAY'S CONTENT

- AdaBoost
- Quiz 02

ADABOOST

Data:

- Iteratively reweight samples
- Misclassified ones get higher weight

Model:

- Chain many weak learners (decision stumps)
- Weight each by its accuracy for final prediction

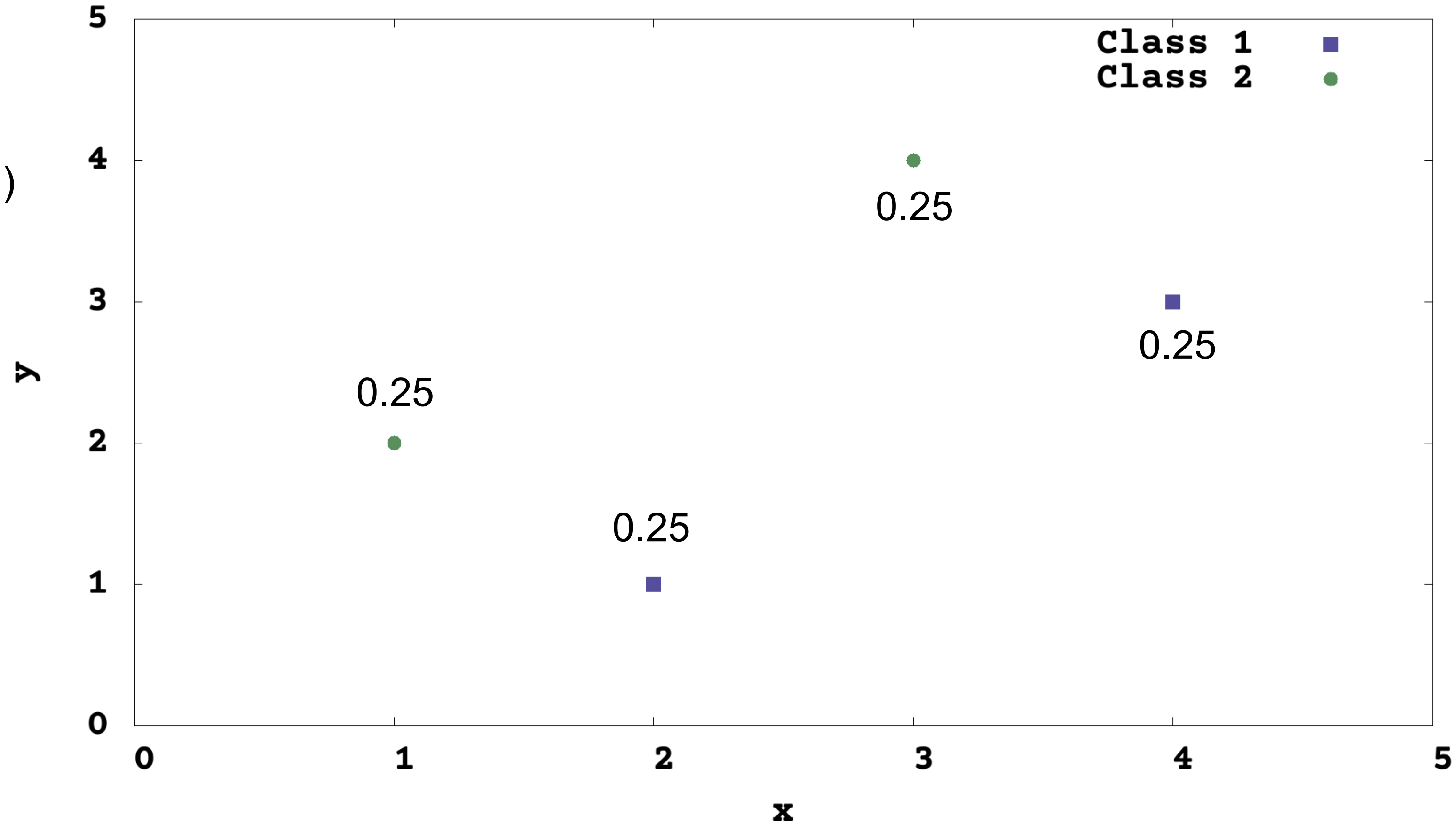
ADABOOST

1. Initialize sample weights uniformly.
2. Repeat for T rounds:
   a. Train a learner h_t on the weighted data.
   b. Compute weighted error ϵ_t.
   c. Set learner weight $\alpha_t = \frac{1}{2}\ln\frac{1-\epsilon_t}{\epsilon_t}$.
   d. Update and renormalize sample weights.
3. Predict X by $\operatorname{sign}(\sum_{t=1}^{T} \alpha_t h_t(X))$.

STEP1:
INITIALIZE SAMPLE WEIGHTS

Sample Weights

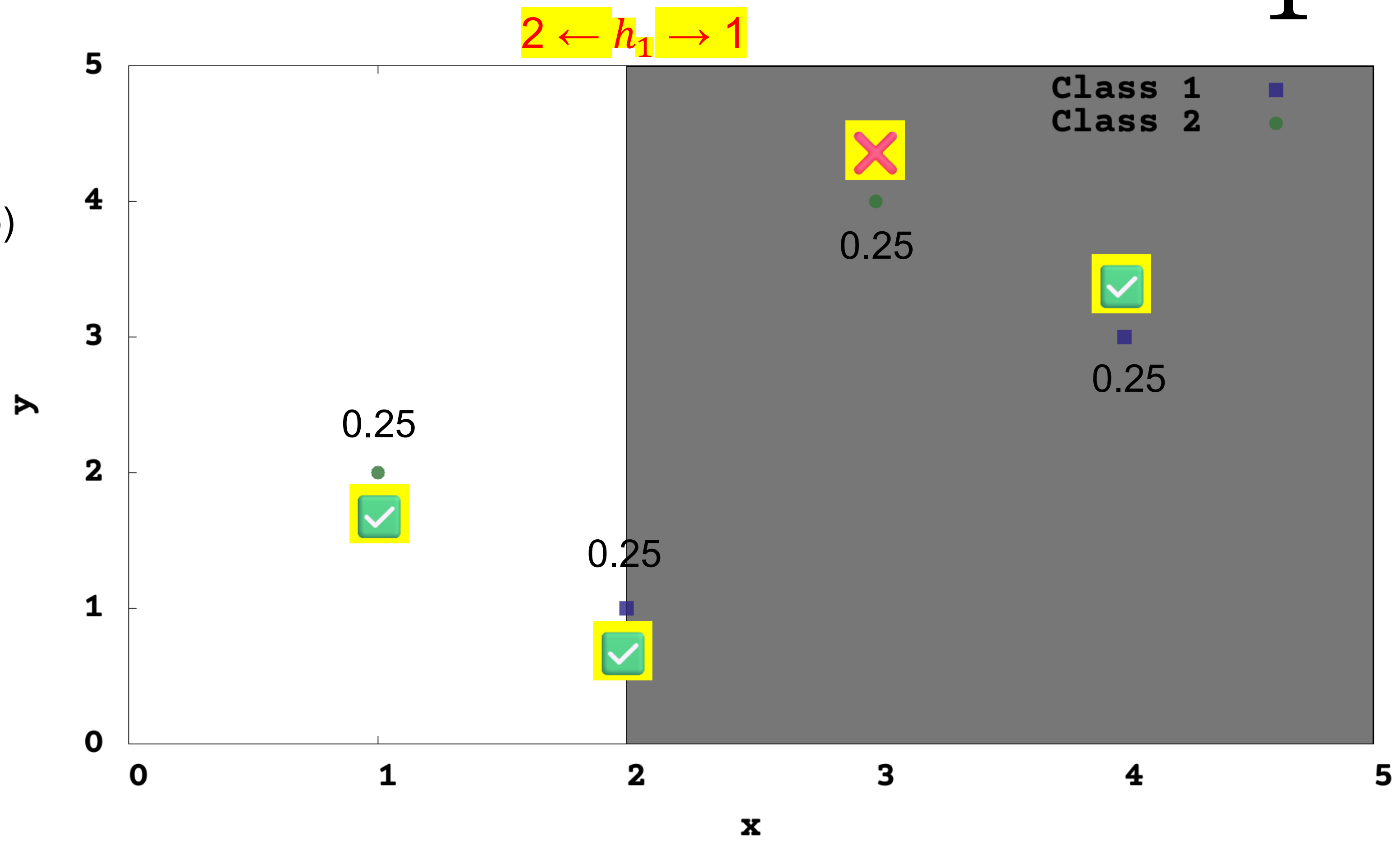
(0.25, 0.25, 0.25, 0.25)



STEP2-A:
TRAIN LEARNER h_1

Sample Weights

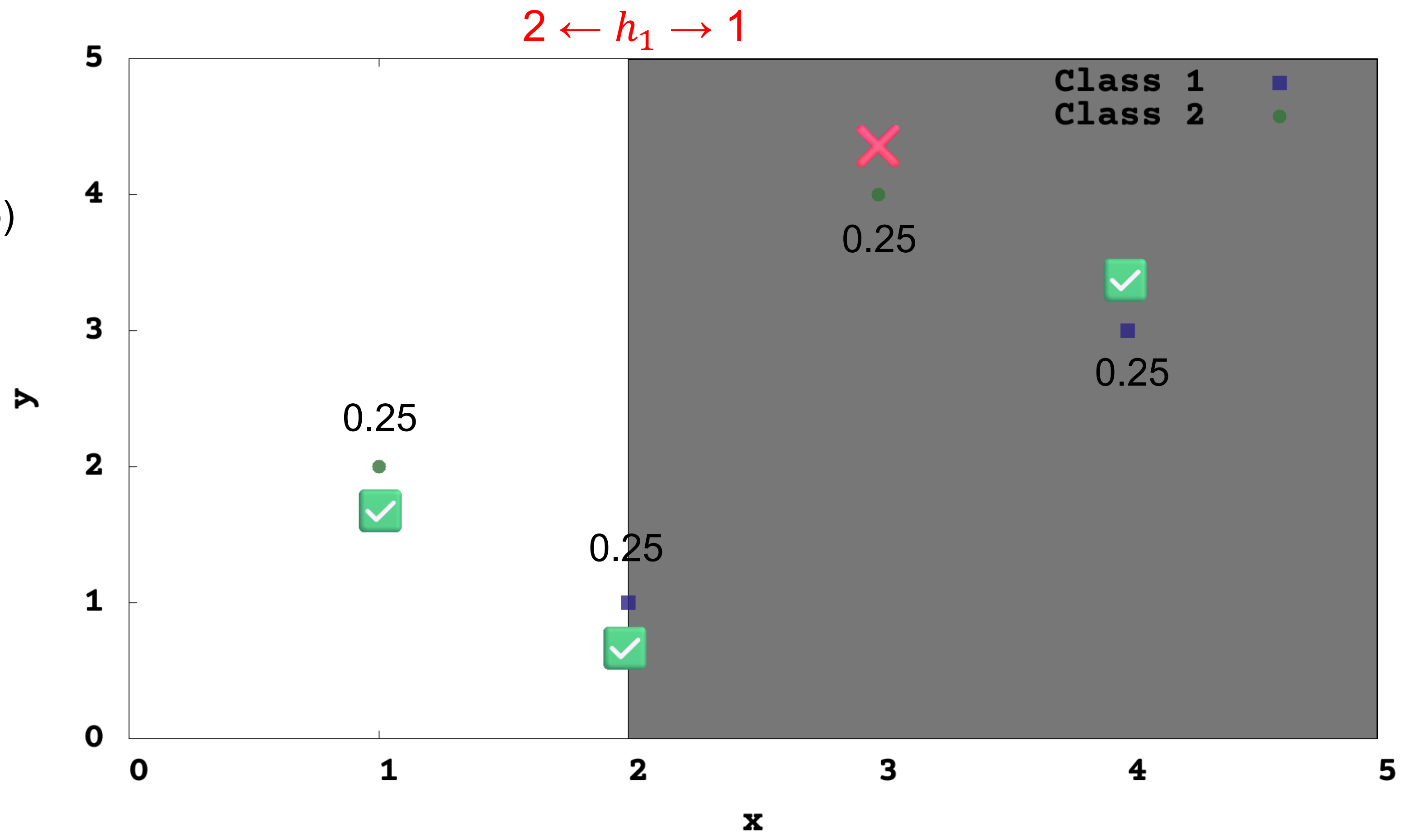
(0.25, 0.25, 0.25, 0.25)



STEP2-B: COMPUTE WEIGHTED ERROR ϵ_1

Sample Weights

(0.25, 0.25, 0.25, 0.25)



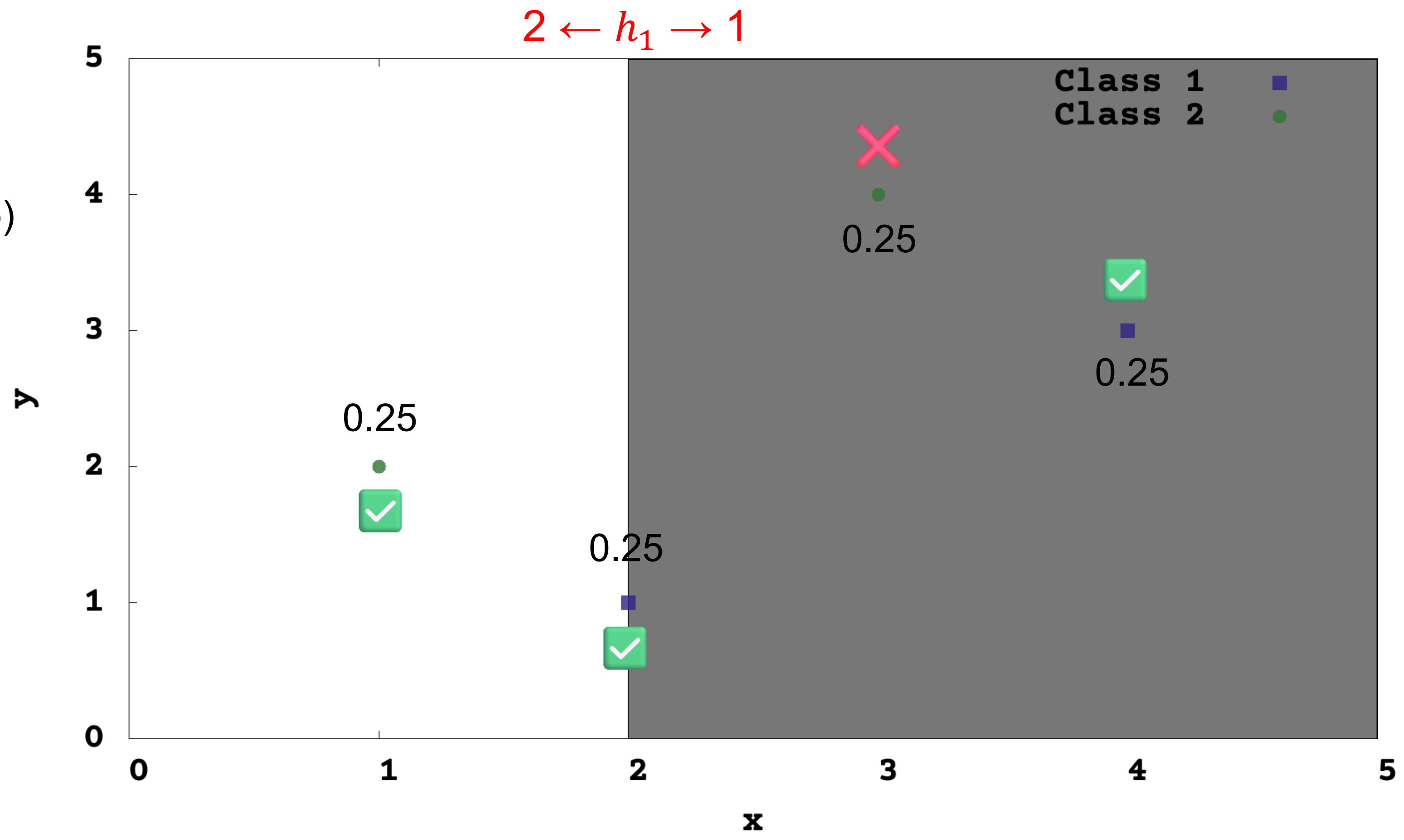
Weighted Error

$\epsilon_1 = 0.25 * 1 = 0.25$

STEP2-C: COMPUTE LEARNER WEIGHT α_t

Sample Weights

(0.25, 0.25, 0.25, 0.25)



Weighted Error

$\epsilon_1 = 0.25 * 1 = 0.25$

Learner Weight

$$\alpha_1 = \frac{1}{2} \ln \frac{1-\epsilon_1}{\epsilon_1} \approx 0.55$$

STEP2-D: UPDATE SAMPLE WEIGHTS

Sample Weights

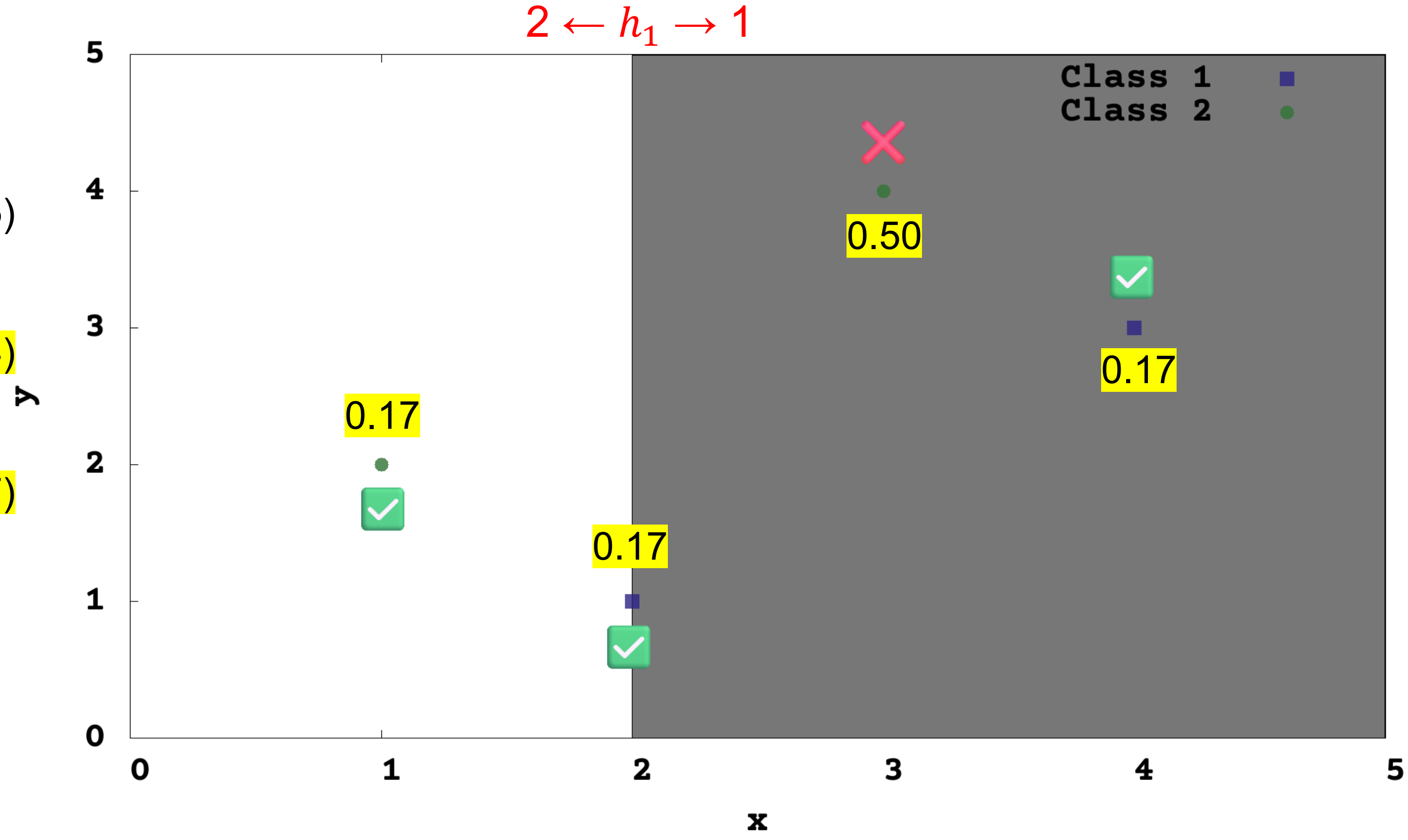
(0.25, 0.25, 0.25, 0.25)

↓ multiply factor

(0.14, 0.14, 0.43, 0.14)

↓ renormalize

(0.17, 0.17, 0.50, 0.17)



Weighted Error

$\epsilon_1 = 0.25 * 1 = 0.25$

Learner Weight

$$\alpha_1 = \frac{1}{2}\ln\frac{1-\epsilon_1}{\epsilon_1} \approx 0.55$$

Update Factor

✅: $e^{-\alpha_1}$

❌: e^{α_1}

(0.58, 0.58, 1.73, 0.58)

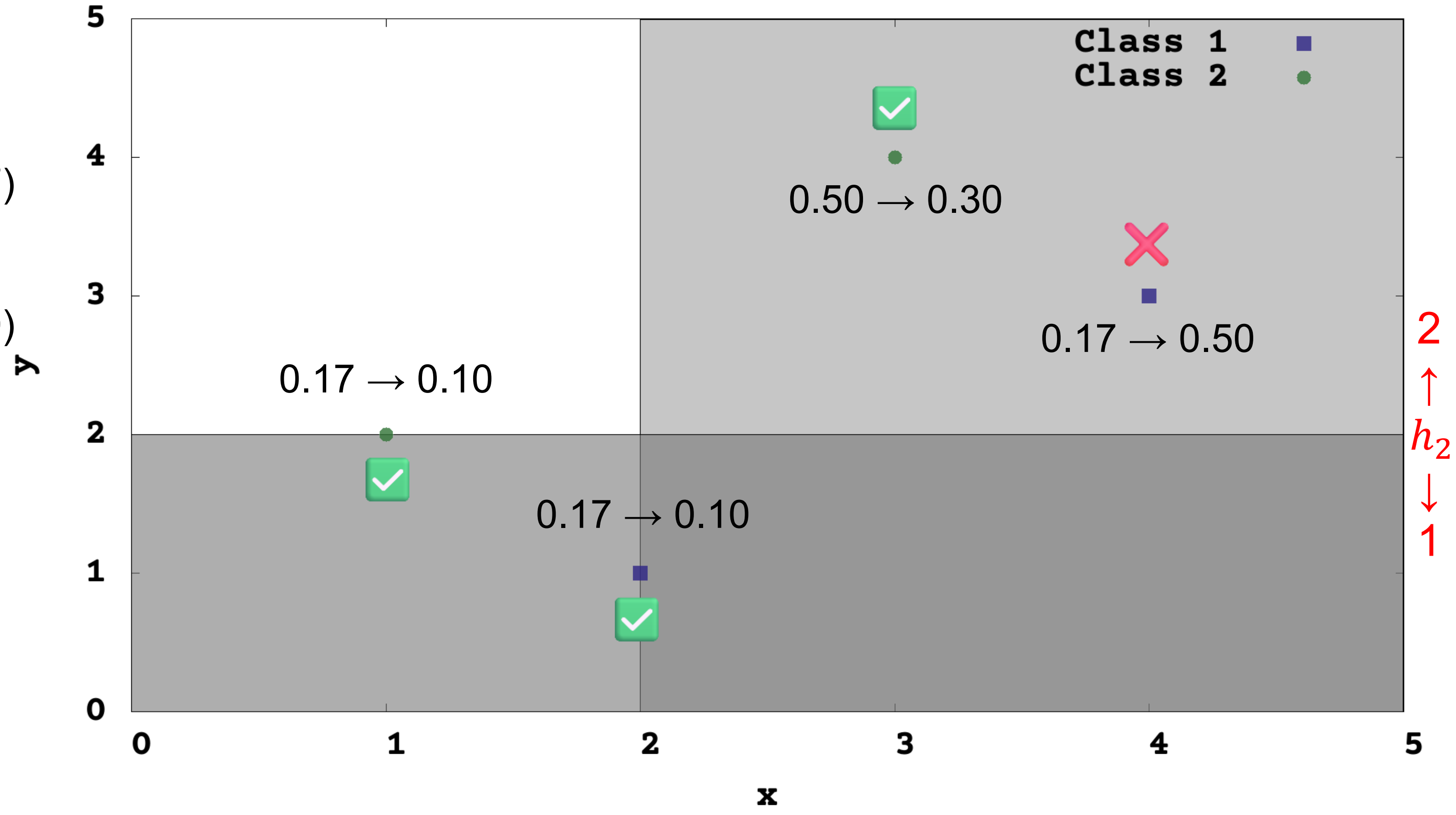
STEP2:
REPEAT UNTIL CONVERGENCE

Sample Weights

(0.17, 0.17, 0.50, 0.17)

↓ mul & norm

(0.10, 0.10, 0.30, 0.50)



Weighted Error

$\epsilon_2 = 0.17 * 1 = 0.17$

Learner Weight

$$\alpha_2 = \frac{1}{2}\ln\frac{1-\epsilon_2}{\epsilon_2} \approx 0.80$$

Update Factor

✅: $e^{-\alpha_2}$

❌: e^{α_2}

(0.45, 0.45, 0.45, 2.24)

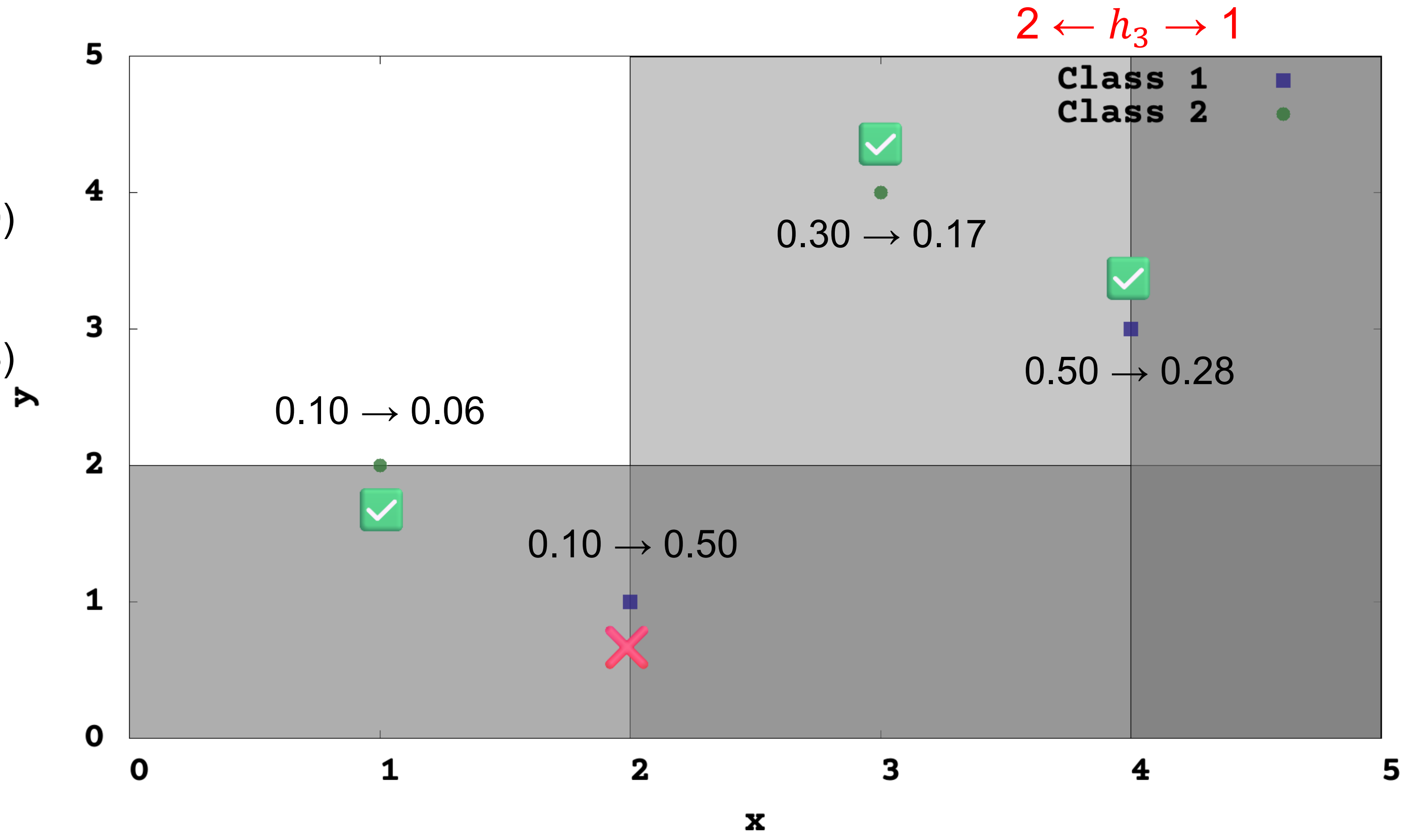
STEP2:
REPEAT UNTIL CONVERGENCE

Sample Weights

(0.10, 0.10, 0.30, 0.50)

↓ mul & norm

(0.06, 0.50, 0.17, 0.28)



Weighted Error

$$\epsilon_3 = 0.10 * 1 = 0.10$$

Learner Weight

$$\alpha_3 = \frac{1}{2}\ln\frac{1-\epsilon_3}{\epsilon_3} \approx 1.10$$

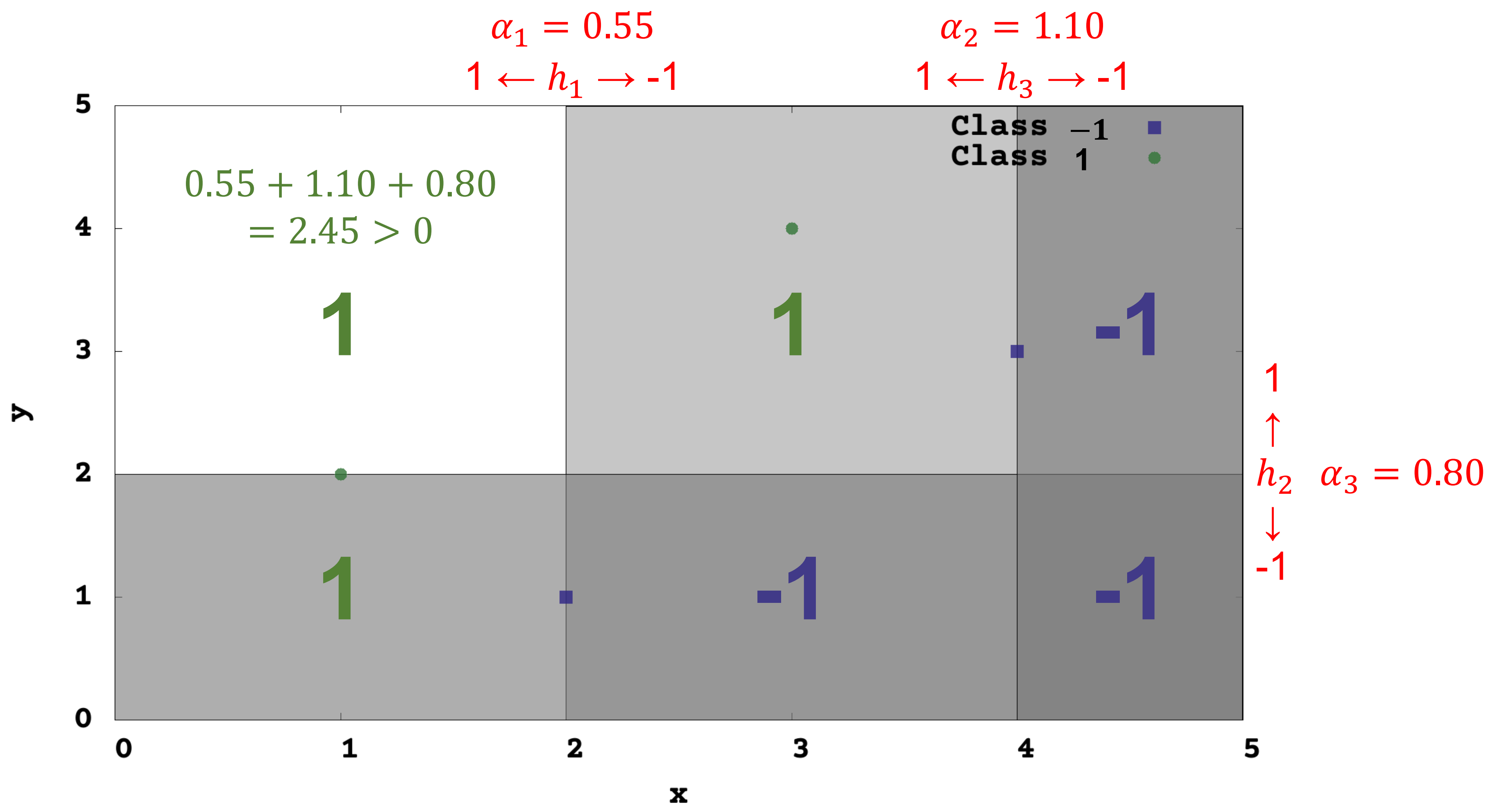
Update Factor

✅: $e^{-\alpha_3}$

❌: e^{α_3}

(0.33, 3.00, 0.33, 0.33)

STEP3: PREDICT X BY $\mathrm{sign}(\sum_{t=1}^{T} \alpha_t h_t(X))$



QUIZ 02

A	B	C	D
3	1	14	3

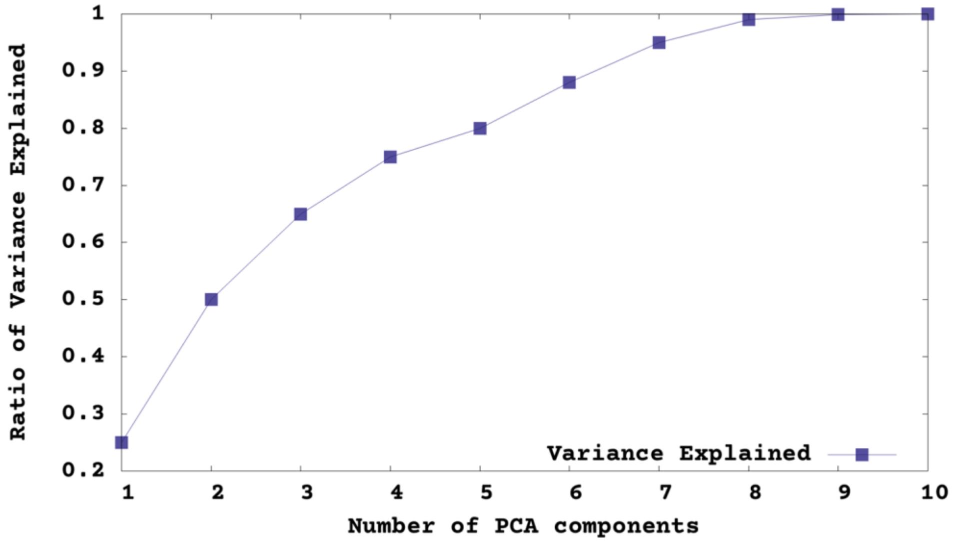
1) When preprocessing data using principal component analysis with eigenvalues, which of the following statements is false?

(a) The eigenvalue from the first component is the largest.

(b) The first two components explain about half of the dataset variance.

(c) The first components explain the most variance. They should be dropped.

(d) The last components explain the least variance. They should be dropped.



Accumulative values

QUIZ 02

A	B	C	D
1	2	0	18

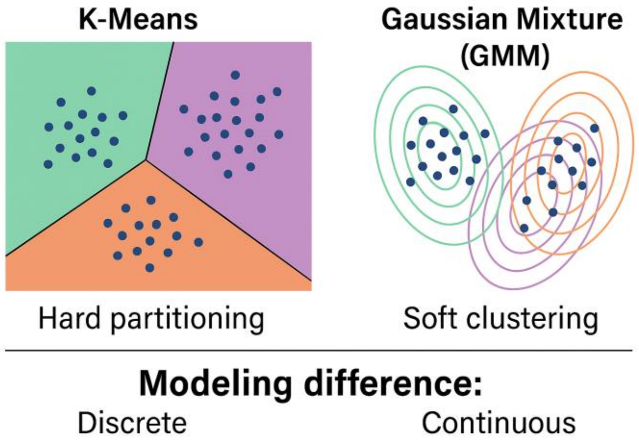
2) Which is true about soft clustering?

(a) Soft clustering is used in K-Means.

(b) Soft clustering is used in K Nearest Neighbors.

(c) Soft clustering is used when making final cluster assignments after training, but not during training.

(d) Soft clustering allows a point's likelihood to be explained by multiple clusters.



K-Means: hard
GMM: soft

QUIZ 02

A	B	C	D
2	15	4	0

3) Which of the following is a supervised technique?

(a) Guassian Mixture Models

(b) Logistic Regression

(c) Hierarchical Agglomerative Clustering

(d) Principal Component Analysis

All clusterings are un-supervised.

QUIZ 02

A	B	C	D
2	15	4	0

4) Which statement about the K-Means algorithm is true?

(a) K-Means deals well with overlapping clusters of different classes.

(b) K-Means can be improved by choosing initial cluster centers that are far apart.

(c) K-Means is guaranteed to converge on the best cluster split.

(d) K-Means works well when clusters are not spherical.

QUIZ 02

A	B	C	D
3	12	3	3

5) Which of the following statements about decision boundaries is false?

(a) Decision boundaries in decision trees are noisy because small changes in the dataset can lead to large changes in pivot choices.

(b) Linear regression creates a complicated decision boundary that is known to easily overfit to the individual data samples.

(c) The decision boundaries in a Gaussian Mixture Model cluster assignment consider the cluster variance.

(d) With most techniques, poor decision boundaries can be improved with additional data, especially if it is near the decision boundary

QUIZ 02

6) Consider the table of initial distances and figure below. K-Means is initalized with three clusters, with starting points (1,1), (1,2), and (2,1).

Cluster	1	2	3
Start	(1,1)	(1,2)	(2,1)

Distance	1,1	1,2	2,1	1,7	1,8	1,9	7,1	8,1	9,1
Cluster 1	0	1	1	6	7	8	6	7	8
Cluster 2	1	0	1.41	5	6	7	6.08	7.07	8.06
Cluster 3	1	1.41	0	6.08	7.07	8.06	5	6	7

6.1) Which points are intially assigned to each cluster?

Cluster 1: ____

Cluster 2: ____

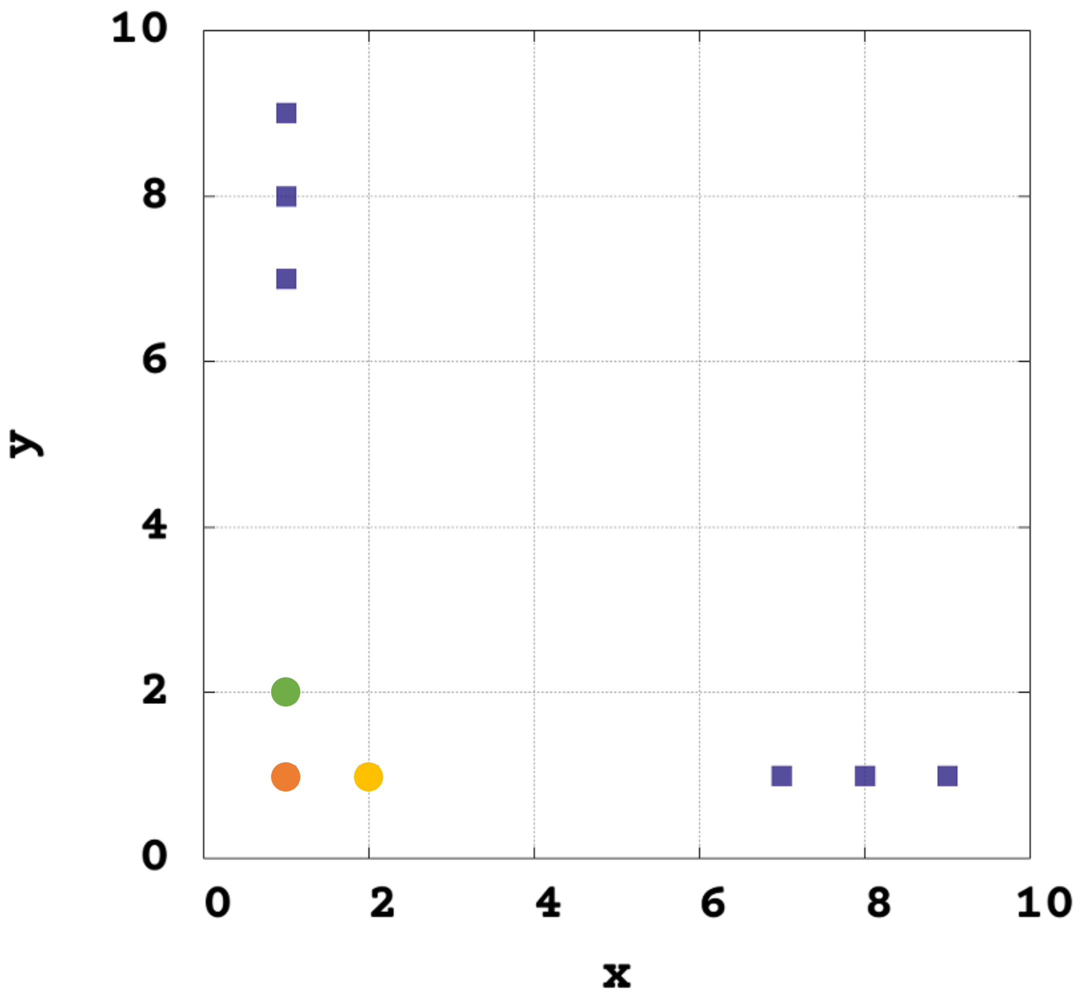
Cluster 3: ____

6.2) After cluster assignment, the cluster locations are updated. What are the new cluster locations? You may use fractional values.

Cluster 1: ____

Cluster 2: ____

Cluster 3: ____

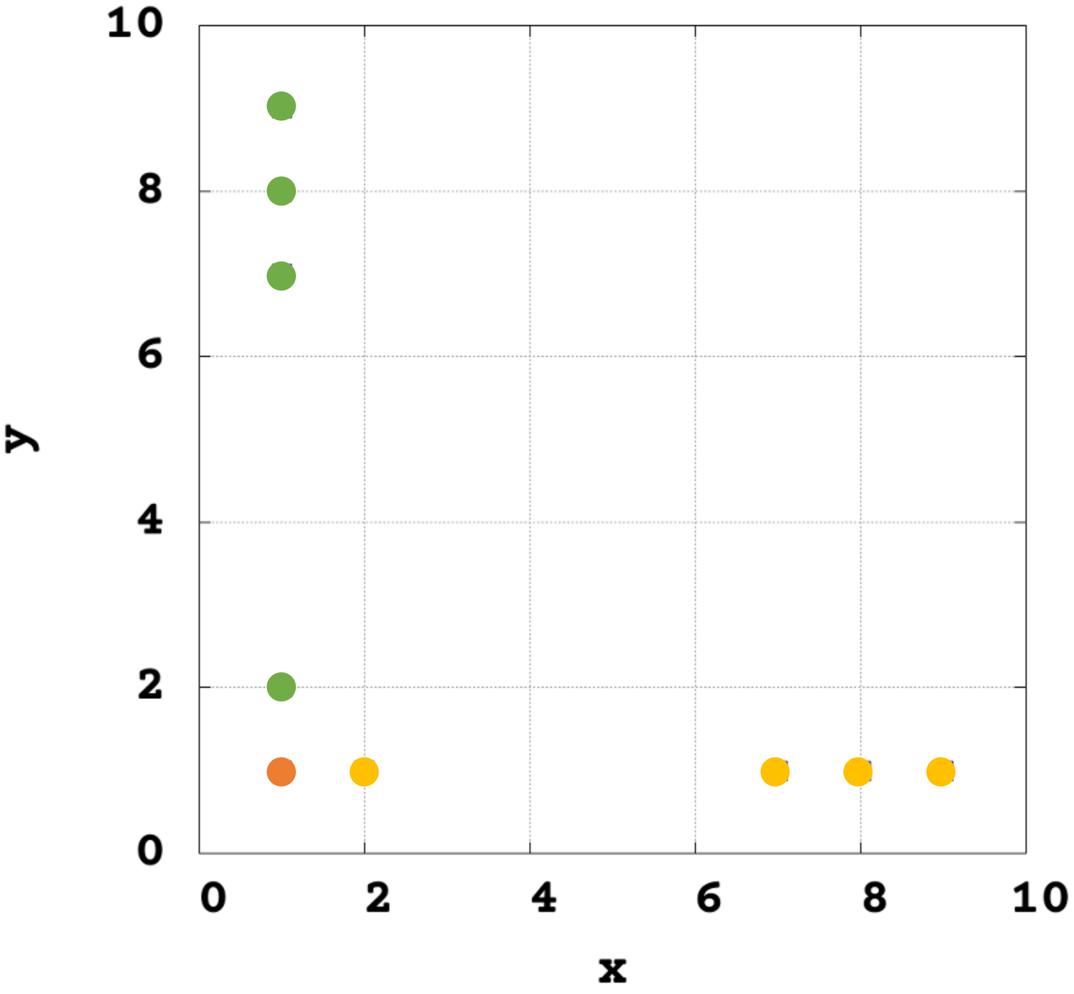


6.1) Which points are intially assigned to each cluster?

Cluster 1: (1,1) ____________________

Cluster 2: (1,2) (1,7) (1,8) (1,9) ____________________

Cluster 3: (2,1) (7,1) (8,1) (9,1) ____________________



6.2) After cluster assignment, the cluster locations are updated. What are the new cluster locations? You may use fractional values.

Cluster 1: ____________________

Cluster 2: ____________________

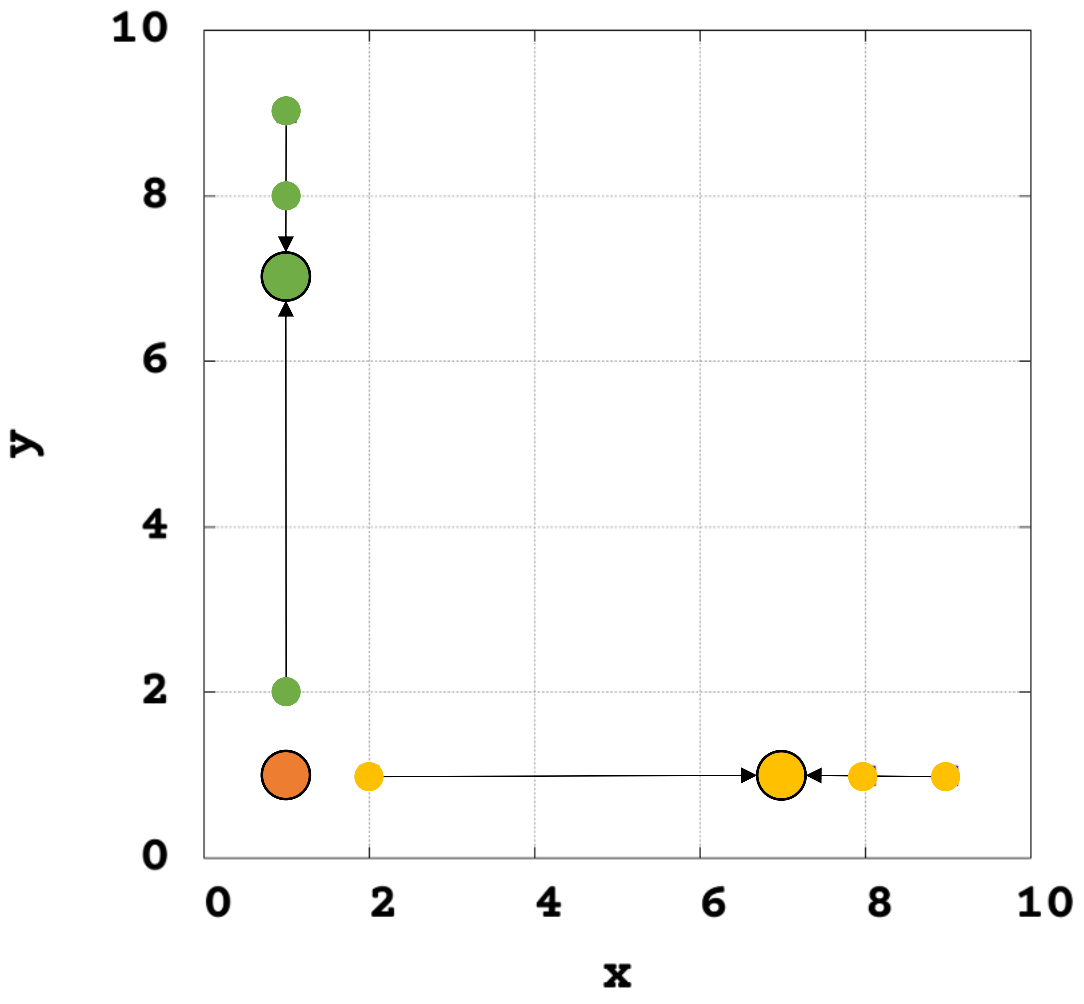
Cluster 3: ____________________

6.1) Which points are intially assigned to each cluster?

Cluster 1: (1,1)

Cluster 2: (1,2) (1,7) (1,8) (1,9)

Cluster 3: (2,1) (7,1) (8,1) (9,1)



6.2) After cluster assignment, the cluster locations are updated. What are the new cluster locations? You may use fractional values.

Cluster 1: (1,1)

Cluster 2: $\left(\frac{1+1+1+1}{4}, \frac{2+7+8+9}{4}\right) \rightarrow \left(1, \frac{13}{2}\right)$

Cluster 3: $\left(\frac{2+7+8+9}{4}, \frac{1+1+1+1}{4}\right) \rightarrow \left(\frac{13}{2}, 1\right)$

Q&A