Feedback — Week 2 Quiz

Help Center

Thank you. Your submission for this quiz was received.

You submitted this quiz on **Sat 9 May 2015 10:35 AM PDT**. You got a score of **8.75** out of **10.00**.

Question 1

Given Eps and Minpts, if a point p is density-reachable from a point q, which of the following statements are correct?

Your Answer		Score	Explanation
☐ There exist at least one point, neither p nor q, that is density-reachable from both p and q	✔	0.25	
☐ Point p must be directly density-reachable from q	✔	0.25	
☑ Point q is density-connected to p	✔	0.25	
☐ Point q is also density-reachable from p	✔	0.25	
Total		1.00 / 1.00	

Question Explanation

The following statement is correct:

- Point q is density-connected to p. (Correct since both p and q are density-reachable from q).

The following statements are incorrect:

- Point q is also density-reachable from p. (Incorrect when p is not a core point.)
- Point p must be directly density-reachable from q. (Incorrect when the chain of points include points other than p and q.)
- There exist at least one point, neither p nor q, that is density-reachable from both p and q

(Incorrect since p may not be a core point.)

Question 2

Suppose Eps = 4cm and Minpts = 5. Randomly select two points p, q from the observed data points. We have dist(p, q) = 1cm. Which of the following statements are correct?

Your Answer		Score	Explanation
☐ Points p and q must be in the same cluster	✔	0.25	
☑ Points p and q may not belong to any clusters	✔	0.25	
☐ Points p and q must be in different clusters	✔	0.25	
☐ Points p and q may be in the different clusters	✘	0.00	
Total		0.75 / 1.00	

Question Explanation

The following statements are correct:

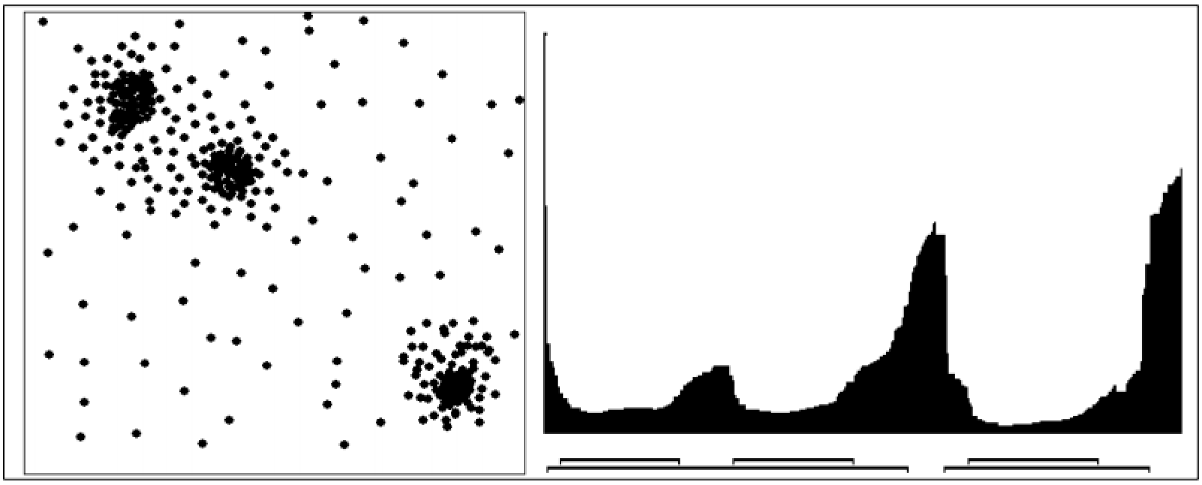
- Points p and q may not belong to any clusters. (Correct if neither p nor q is a core point.)
- Points p and q may be in the different clusters. (Correct if (i) neither p nor q is a core point, (ii) there exists a core point o1 such that o1 is only density-reachable to p but not density reachable to q, and (iii) there exists a core point o2 such that o2 is only density-reachable to q but not density reachable to p.)

The following statements are incorrect:

- Points p and q must be in different clusters; (Incorrect since p and q might be noise/outliers, and they do not belong to any clusters.)
- Points p and q must be in the same cluster; (Incorrect since p and q might be noise/outliers, and they do not belong to any clusters.)

Question 3

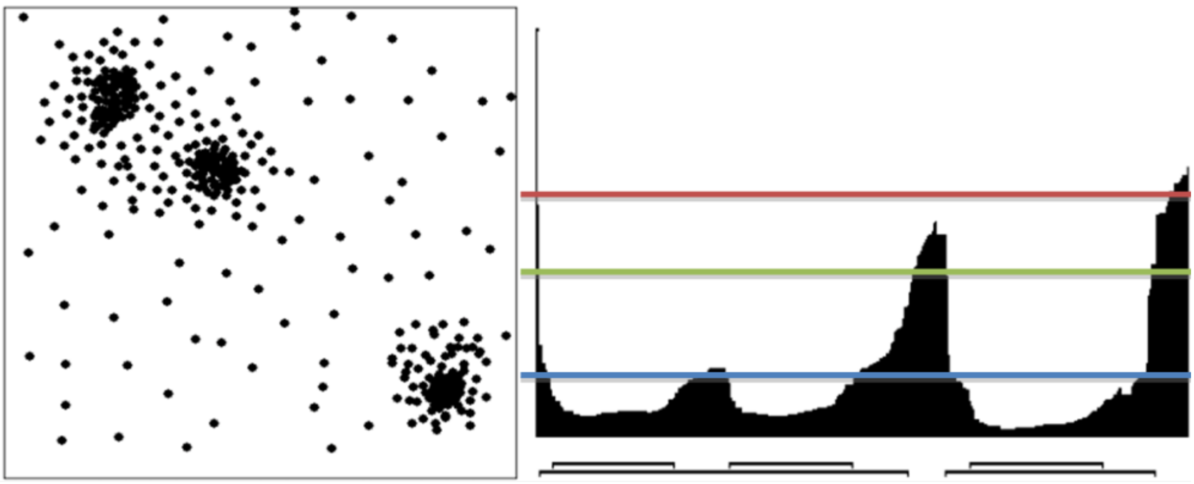
Given the following synthetic data set (left) and the reachability-plot (right), how many clusters might there be by setting different reachability thresholds?



Your Answer		Score	Explanation
1	✔	0.25	
more than 3	✔	0.25	
3	✔	0.25	
2	✔	0.25	
Total		1.00 / 1.00	

Question Explanation

There might be 1, 2, or 3 clusters, which correspond to the red, green and blue lines respectively. If we set Minpts = 1 and Eps to an extremely small value, then every point may form a cluster, i.e., there are more than 3 clusters.



Question 4

Considering the CLIQUE clustering algorithm, which of the following statements are correct?

Your Answer		Score	Explanation
☑ Any point that belongs to a cluster in a 1-D subspace A may not belong to any clusters (dense regions) in any 2-D subspaces that include A.	✔	0.25	
☐ Any point that belongs to a cluster in a 1-D subspace A must belong to some cluster (dense region) in any 2-D subspaces that includes A	✔	0.25	
☐ Any point that does not belongs to a cluster in a 1-D subspace A may belong to some cluster (dense region) in any 2-D subspaces that includes A	✔	0.25	
☑ Any point that does not belongs to a cluster in a 1-D subspace A must not belong to any clusters (dense regions) in any 2-D subspaces that include A.	✔	0.25	
Total		1.00 / 1.00	

Question Explanation

The following statements are correct:

- Any point that belongs to a cluster in a 1-D subspace A may not belong to any clusters (dense regions) in any 2-D subspaces that include A. (Correct since the other dimension in a 2-D subspace may belong to sparse regions, i.e., does not belong to any clusters.)
- Any point that does not belongs to a cluster in a 1-D subspace A must not belong to any clusters (dense regions) in any 2-D subspaces that include A. (Correct for any data point, as long there is one dimension that belongs to the sparse region, the data point will not belong to any dense region in higher dimensions that includes A.)

The following statements are incorrect:

- Any point that belongs to a cluster in a 1-D subspace A must belong to some cluster (dense region) in any 2-D subspaces that includes A; (Incorrect since the other dimension in a 2-D subspace may belong to sparse regions, i.e., does not belong to any clusters.)
- Any point that does not belongs to a cluster in a 1-D subspace A may belong to some cluster (dense region) in any 2-D subspaces that includes A; . (Incorrect for any data point, as long there is one dimension that belong to the sparse region, the data point will not belong to any dense region in higher dimension that includes A.)

Question 5

In Gaussian Mixture Model, the random variables are univariate. For i-th cluster, where i = 1… k, the data points are drawn from Gaussian distribution,

, $x \sim N(\mu_i, \sigma_i)$, i.e., $P(x|C_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}$. Suppose we have learned two clusters from some dataset with the parameter μ_1=0,σ_1=1,μ_2=-2,σ_2=3,P(C_1)=0.25,P(C_2)=0.75. For the new data point x=1, what is the probability that it belongs to the first cluster, i.e. P(C_1 |x)?

Your Answer		Score	Explanation
0.75			
0.5	✔	1.00	
0.25			
0.8			
Total		1.00 / 1.00	

Question Explanation

Please note the texts of this question have been updated.

$P(C_1|x) \propto P(x|C_1 = N(\mu_1, \sigma_1))P(C_1) = 0.25 * \frac{1}{\sqrt{2\pi * 1}} * e^{-\frac{(1-0)^2}{2*1^2}}$. Similarly, $P(C_2|x) \propto 0.75 * \frac{1}{\sqrt{2\pi * 3}} * e^{-\frac{(1-(-2))^2}{2*3^2}}$. Because of $\mathrm{P(C_1|x) + P(C_2|x) = 1}$, we have $\mathrm{P(C_1|x)} = \frac{0.25}{0.25+0.25} = \frac{1}{2} = 0.5$.

Question 6

Which of the following statements about Gaussian Mixture Models (GMM) are correct?

Your Answer		Score	Explanation
☐ GMM usually converges much faster than K-means.	✔	0.25	
☐ In GMM, for different initializations, we will always have exactly same the clustering result.	✔	0.25	

☑ GMM assumes the data points are generated by some Gaussian distributions.	✔	0.25
☑ In GMM, we have to specify the number of clusters.	✔	0.25
Total		1.00 / 1.00

Question Explanation

Please note the texts of this question have been updated.

GMM is sensitive to its initialization. That is, different initializations may lead to different results. GMM usually converges slower than K-means because the parameters evolve more smoothly, as shown in the example in the slides.

Question 7

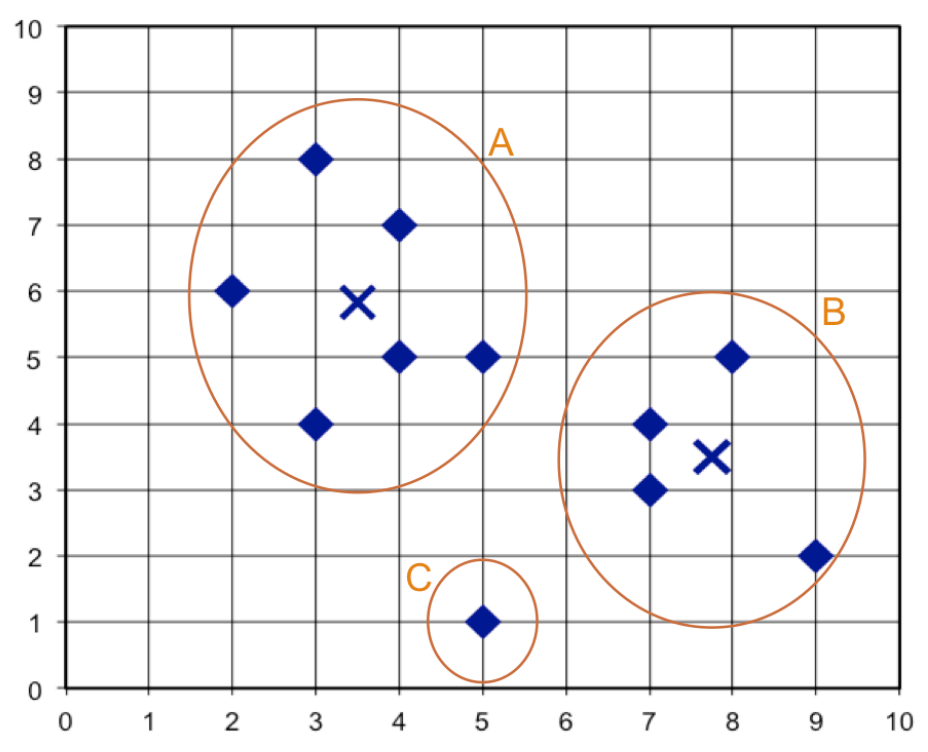
In which ONE of the following scenarios may Gaussian Mixture Models (GMM) not work properly?

Your Answer		Score	Explanation
◉ Build 10 clusters from 20 10-dimensional data points randomly generated from 10 Gaussian distributions.	✔	1.00	
○ Build 10 clusters from 2000 10-dimensional data points randomly generated from 10 Gaussian distributions.			
Total		1.00 / 1.00	

Question Explanation

GMM needs large datasets to fit its parameters.

Question 8



Consider the three clusters A, B, and C shown in Figure 1. Using Euclidean distance as the similarity measure, which two clusters would be merged first in agglomerative clustering using complete link (diameter)?

Your Answer		Score	Explanation
All three options above are tied			
A and B	✖	0.00	
B and C			
A and C			
Total		0.00 / 1.00	

Question Explanation

B and C. The farthest point in A to B is (3,8). The farthest point in A to C is (3,8). The farthest point in B to A is (9, 2). The farthest point in B to C is (8, 5). Complete link (diameter) between clusters is defined as the similarity between the most dissimilar members. We want to merge two clusters that produce the smallest diameter. Since (5,1) and (8,5) have the smallest distance between them of all possible pairs, B and C should be merged first.

Question 9

Recall from Lecture 4-8 that the objective of learning generative models is to find the parameters that maximize the likelihood of the observed data. Suppose we have a set of points D drawn from Gaussian distribution. For D = {-1, 0, 1}, which of the following set of parameters (μ, σ) produces the maximum L(N(μ, σ2): D)?

Your Answer		Score	Explanation
◉ μ = 0, σ = 1	✔	1.00	
○ μ = 0, σ = 2			
○ μ = 1, σ = 0			
○ μ = 2, σ = 0			
Total		1.00 / 1.00	

Question Explanation

μ = 0, σ = 1. The maximum likelihood estimator for μ in a Gaussian distribution is the sample mean, and the maximum likelihood estimator for σ2 is the sample variance. Something similar in spirit is done in the M-step of the k-means algorithm, where the average of all points in a cluster becomes the new centroid in order to minimize the sum of distance.

Question 10

Consider the three hierarchical clustering algorithms introduced in Lecture 4, BIRCH, CURE, and CHAMELEON. Which of the following statements about these algorithms is TRUE? (Select all that apply)

Your Answer		Score	Explanation
☐ CHAMELEON requires a graph as the input.	✔	0.25	
☑ BIRCH and CHAMELEON both use a divisive method to partition the objects into small groups first before merging them back to form the final clusters.	✔	0.25	

☑ CHAMELEON and CURE are better at capturing irregular shaped clusters than BIRCH	✔	0.25
☐ All three algorithms can only work with Euclidean distance as the similarity metric.	✔	0.25
Total		1.00 / 1.00

Question Explanation

CHAMELEON and CURE are better at capturing irregular shaped clusters than BIRCH (True. By merging graphlets, CHAMELEON is able to capture complex shapes. CURE forms complex shapes by merging small disks. BIRCH tends to produce spherical clusters as limited by the diameter and radius parameters.)

BIRCH and CHAMELEON both use a divisive method to partition the objects into small groups first before merging them back to form the final clusters. (True. BIRCH uses the CF Tree to divide the dataset into micro-clusters, whereas CHAMELEON breaks up the kNN graph into small graphlets. Both then merge these small units into larger clusters.)

All three algorithms can only work with Euclidean distance as the similarity metric. (False. They can work with any similarity metric.)

CHAMELEON requires a graph as the input. (False. CHAMELEON constructs the kNN graph from a set of objects by measuring the distance between objects and linking each to the k nearest neighbors.)