Review Test Submission: Quiz 8: Clustering & dimension reduction

User	Fei Shen
Course	CS-584-Parent.17S
Test	Quiz 8: Clustering & dimension reduction
Started	4/16/17 7:34 PM
Submitted	4/16/17 7:47 PM
Due Date	4/28/17 11:59 PM
Status	Completed
Attempt Score	130 out of 130 points
Time Elapsed	13 minutes out of 2 hours
Results Displayed	All Answers, Submitted Answers, Correct Answers

Question 1

Use k-means to split four data points into two clusters. The data points are x1=(0,0), x2=(1,0), x3=(1,1), x4=(2,1). Suppose the initial centers are C1=(0,1), C2=(2,0). Now we do one iteration of the k-means algorithm (Pg. 8 in the lecture note "11-clustering.pdf") as follows:

Classify:

x1 should be assigned to center **[a]** (*please answer 1 or 2*),

x2 should be assigned to center **[b]** (*please answer 1 or 2*),

x3 should be assigned to center **[c]** (*please answer 1 or 2*),

x4 should be assigned to center **[d]** (*please answer 1 or 2*),

Recenter:

The new centroid of class 1 is (**[e]**,**[f]**) (*Please give the coordinates. Do NOT round*),

The new centroid of class 2 is (**[g]**,**[h]**) (*Please give the coordinates. Do NOT round*).

Specified Answer for: a ✓ 1

Specified Answer for: b ✓ 2

Specified Answer for: c ✓ 1

Specified Answer for: d ✓ 2

Specified Answer for: e ✓ 0.5

Specified Answer for: f ✓ 0.5

Specified Answer for: g 1.5

Specified Answer for: h 0.5

Correct Answers for: a		
Evaluation Method	**Correct Answer**	**Case Sensitivity**
Exact Match	1	
Correct Answers for: b		
Evaluation Method	**Correct Answer**	**Case Sensitivity**
Exact Match	2	
Correct Answers for: c		
Evaluation Method	**Correct Answer**	**Case Sensitivity**
Exact Match	1	
Correct Answers for: d		
Evaluation Method	**Correct Answer**	**Case Sensitivity**
Exact Match	2	
Correct Answers for: e		
Evaluation Method	**Correct Answer**	**Case Sensitivity**
Pattern Match	0?.50*	
Correct Answers for: f		
Evaluation Method	**Correct Answer**	**Case Sensitivity**
Pattern Match	0?.50*	
Correct Answers for: g		
Evaluation Method	**Correct Answer**	**Case Sensitivity**
Pattern Match	1.50*	
Correct Answers for: h		
Evaluation Method	**Correct Answer**	**Case Sensitivity**
Pattern Match	0?.50*	

Question 2

Given four data points, x1=(0,0), x2=(1,0), x3=(1,1), x4=(2,1), we want to use the E.M. algorithm to cluster them into two classes. Suppose initially the parameters of the two Gaussian components are: $\lambda = \{ \mu_1 = (0,0), \mu_2 = (1,1), \sigma_1^2 = \sigma_2^2 = \frac{1}{2} \}$, and initially P(y=1)=0.25, P(y=2)=0.75.

Now we do an E-step, please compute the probability that x4 is assigned to class 1, that is, P(y=1 | x4, λ) = ?.

The E.M. algorithm is given on Pg. 49 and Pg. 51 of the lecture note "11-clustering.pdf". Note the probabilities need to be normalized in E-step. Round your answer to 3 decimal places if necessary.

Selected Answer: 0.006

Correct Answer: 0.00605

Answer range +/- 0.001 (0.00505 - 0.00705)

Question 3

Given two data points x1=5, x2=2, suppose we want to use the E.M. algorithm to cluster them into two classes. In an E-step we have computed:

P(y=1 | x1,λ)=0.6, P(y=2 | x1,λ)=0.4,

P(y=1 | x2,λ)=0.2, P(y=2 | x2,λ)=0.8.

Now we need to do an M-step. Please compute the following parameters (probability, mean, variance). The E.M. algorithm is given on Pg. 49 and Pg. 51 of the lecture note "11-clustering.pdf". Note do NOT round your answers.

p_1=**[a]**

p_2=**[b]**

μ_1=**[c]**

μ_2=**[d]**

σ_1^2=**[e]**

σ_2^2=**[f]**

Specified Answer for: a ✓ 0.4

Specified Answer for: b ✓ 0.6

Specified Answer for: c ✓ 4.25

Specified Answer for: d ✓ 3

Specified Answer for: e ✓ 1.6875

Specified Answer for: f ✓ 2

Correct Answers for: a		
Evaluation Method	**Correct Answer**	**Case Sensitivity**
✓ *Pattern Match*	0?.40*	
Correct Answers for: b		
Evaluation Method	**Correct Answer**	**Case Sensitivity**
✓ *Pattern Match*	0?.60*	
Correct Answers for: c		
Evaluation Method	**Correct Answer**	**Case Sensitivity**
✓ *Pattern Match*	4.250?	
Correct Answers for: d		
Evaluation Method	**Correct Answer**	**Case Sensitivity**
✓ *Pattern Match*	3(.000)?	
Correct Answers for: e		
Evaluation Method	**Correct Answer**	**Case Sensitivity**

Pattern Match	(1.6875)|(2.848)| (2.8477)|2.84766|2.847656|2.8476563|2.84765625

Correct Answers for: f

Evaluation Method	Correct Answer	Case Sensitivity
Pattern Match	(2(.000)?)|(4(.000)?)	

Question 4

Given three 2-dimensional data pionts (1,2), (3,1), (4,3), we want to do principal component analysis (PCA). Please fill in the following blanks. Round your answers to three decimal places if necessary. You may refer to Pg 17,18,19 of Lecture note 12a.

The second principal component v2=(**[a]**, **[b]**)

Specified Answer for: a -0.578

Specified Answer for: b 0.816

Correct Answers for: a

Evaluation Method	Correct Answer	Case Sensitivity
Pattern Match	-0?.578	

Correct Answers for: b

Evaluation Method	Correct Answer	Case Sensitivity
Pattern Match	0?.816	

Saturday, April 29, 2017 3:57:31 PM CDT

← OK