

Dashboard ► My courses ► Computer Science 315 / Applied Maths 796 ► Assignments 1-4 ►
Assignment 3 Test - Part A/ Werkopdrag 3 Toets - Afdeling A

Started on	Tuesday, 23 April 2019, 12:00
-------------------	-------------------------------

State	Finished
--------------	----------

Completed on	Tuesday, 23 April 2019, 12:51
---------------------	-------------------------------

Time taken	50 mins 28 secs
-------------------	-----------------

Marks	5.67/8.00
--------------	-----------

Grade	6.38 out of 9.00 (71%)
--------------	--

Question **1**

Correct

Mark 1.00 out of
1.00

Re-order the following steps of the EM algorithm for fitting a Gaussian Mixture Model so that they appear in the correct order.

Herrangskik die volgende stappe van die VO algoritme om 'n Gaussiese Mengsel Model te pas sodat dit in die korrekte volgorde verskyn.

✓ Initialize the different coefficients (i.e. the model parameters) / Initialiseer die verskillende koëffisiente (d.i. die modelparameters)

✓ E-step: Evaluate the responsibilities using the current parameter estimates / V-stap: Bereken die verantwoordlikhede gegee die huidige geskatte parameterwaardes

✓ M-step: Estimate parameters using current responsibilities / O-stap: Beraam die parameters gegee die huidige verantwoordlikhede

✓ Check for convergence of either the parameters or the log-likelihood / Toets of óf die parameters óf die log-aanneemlikheid gekonvergeer het

✓ Keep iterating from step 2 until convergence / Hou aan itereer vanaf stap 2 tot konvergensie

Question **2**

Correct

Mark 1.00 out of
1.00

Re-order the following steps of the *K*-Means algorithm so that they appear in the correct order

Herrangskik die volgende stappe van die K-Gemiddeldes algoritme sodat die stappe in die korrekte volgorde verskyn

✓ Choose initial cluster means, i.e choose K random observations / Initialiseer die groepsgemiddeld van elke groepering, d.i kies K willekeurige waarnemings

✓ Assign class labels, i.e. assign each observation to the cluster mean closest to it / Ken klasetikette toe, d.i. ken elke waarneming toe aan die groepsgemiddeld wat die naaste is daaraan

✓ Update cluster means, i.e. compute cluster means using the labeled (assigned) observations / Oupdateer die groepsgemiddeldes, d.i. bereken die groepsgemiddelde deur die geëtiketeerde (toegekende) waarnemings te gebruik

✓ Repeat above three steps until convergence / Herhaal die vorige drie stappe totdat die algoritme konvergeer

Question **3**

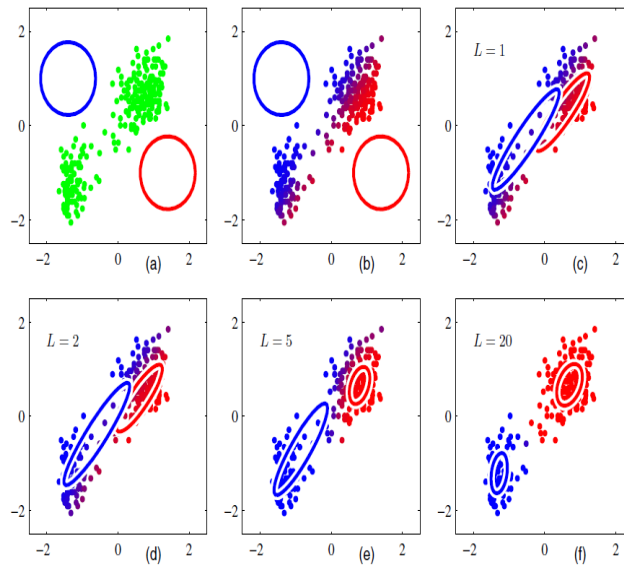
Correct

Mark 1.00 out of

Which of the following descriptions can be associated with the following figure.

1.00

Watter van die volgende beskrywings kan geassosieer word met die volgende figuur.



Select one or more:

- ☒ a. Figure (b) depicts the Expectation step of the Expectation Maximization Algorithm / *Figuur (b) beeld die Verwagtingstap van die Verwagtings Optimiseerings Algoritme uit* ✓
- ☐ b. K-Means / *K-Gemiddeldes*
- ☐ c. Supervised Classification/ *Gerigte Klassifikasie*
- ☒ d. Figure (c) depicts the Maximization Step of the Expectation Maximization Algorithm / *Figuur (c) beeld die Optimeringstap van die Verwagtings Optimiseerings Algoritme uit* ✓
- ☒ e. Two classification features, two clusters / *Twee klassifikasie kenmerke, twee groeperings* ✓

The correct answers are: Figure (b) depicts the Expectation step of the Expectation Maximization Algorithm / *Figuur (b) beeld die Verwagtingstap van die Verwagtings Optimiseerings Algoritme uit*, Figure (c) depicts the Maximization Step of the Expectation Maximization Algorithm / *Figuur (c) beeld die*

*Optimeringstap van die Verwagtings
Optimiseerings Algoritme uit, Two classification
features, two clusters / Twee klassifikasie
kenmerke, twee groeperings*

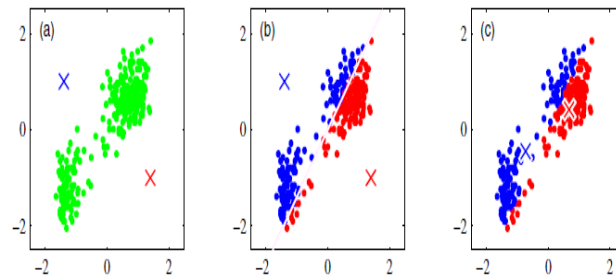
Question **4**

Partially correct

Mark 0.67 out of
1.00

Which of the following descriptions can be associated with the following figure.

*Watter van die volgende beskrywings kan
geassosieer word met die volgende figuur.*



Select one or more:

- ☐ a. New cluster means are computed In Figure (b) / Nuwe groepsgemiddeldes word in Figuur (b) bereken
- ☒ b. New class labels are computed In Figure (c) / Nuwe klasetikette word in Figuur (c) bereken ✗
- ☒ c. Unsupervised learning/ Ongerigte leer ✓
- ☒ d. K-means/ K-gemiddelde ✓
- ☐ e. Gaussian Mixture Models / Gaussiese Mengsel Modelle

The correct answers are: K-means/ K-gemiddelde, Unsupervised learning/ Ongerigte leer

Question **5**

Correct

Mark 1.00 out of
1.00

Assume that you have 4-dimensional data belonging to 2 classes. Moreover, assume that you are employing a 1-of-K coding scheme.

Neem aan dat u 4-dimensionele data voortgebring het uit 2 klasse. Neem verder aan dat jy gebruik maak van 1-uit-K kodering.

Which one of the following options is a valid encoded class label consistent with the information given above.

Watter een van die volgende is 'n geldige geënkodeerde klasetiket in ooreenstemming met die inligting hierbo.

Select one:

- ☐ a. 1
- ☒ b. $[0, 1]^T$ ✓
- ☐ c. $[1, 0, 0, 0]^T$
- ☐ d. None of the options are correct / *Nie een van die opsies is korrek nie.*
- ☐ e. $[1, 1, 0, 0]^T$
- ☐ f. $[1, 1]^T$

The correct answer is: $[0, 1]^T$

Question **6**

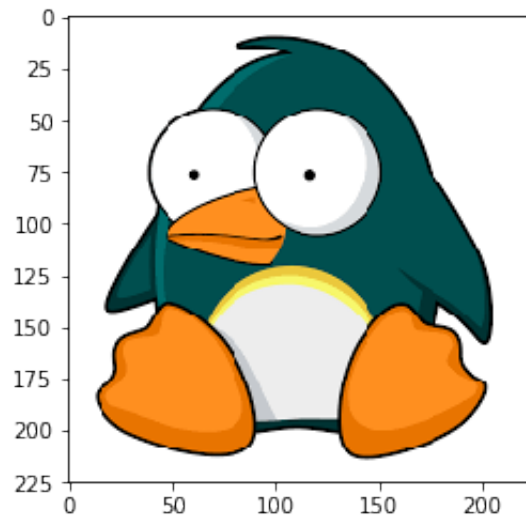
Incorrect

Mark 0.00 out of
1.00

In the assignment you were given the following image:

In die werksopdrag is jy die volgende beeld gegee:

i



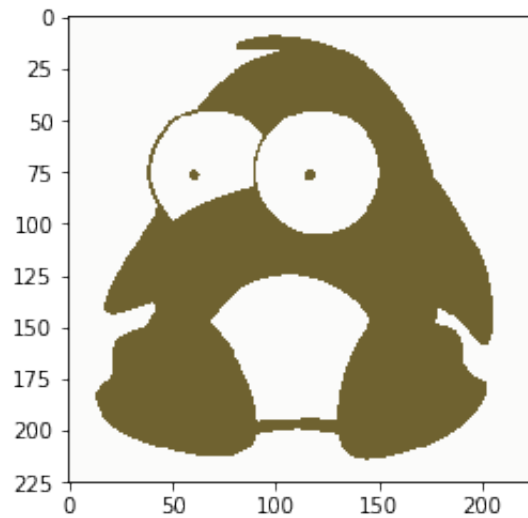
There are in general three important things one should know about a dataset when working with unsupervised classification algorithms: the number of features d , the number of samples n and the number of clusters K the data should be divided into.

In die algemeen is daar drie belangrike dinge wat mens aangaande 'n dataset moet weet wanneer mens werk met ongerigte klassifikasie algoritmes: die aantal kenmerke d , die aantal waarnemings n en die aantal groeperings K wat aanvaar moet word.

What is the value of d , n and K in the case of Figure **A**? Note Figure **A** was generated by applying the K-means algorithm to Figure **i**.

*Wat is die waarde van d , n en K in die geval van Figuur **A**? Neem kennis dat Figuur **A** gegenereer is deur die K-Gemiddeldes algoritme toe te pas op Figuur **i**.*

A



Select one:

- ☐ a. $d = 1, n = 50625, K = 2$
- ☐ b. $d = 3, n = 50625, K = 1$
- ☐ c. $d = 1, n = 225, K = 1$
- ☐ d. $d = 3, n = 225, K = 2$
- ☐ e. $d = 3, n = 50625, K = 2$
- ☒ f. None of the options are correct / Nie een van die opsies is korrek nie. ✖

.

The correct answer is: $d = 3, n = 50625, K = 2$

Question 7

Correct

Mark 1.00 out of 1.00

Suppose a K -means model using $K = 2$ for 2-dimensional data has means at $A(-1, 2)$ and $B(3, -1)$. Assume the axis labels x_1 [horizontal] and x_2 [vertical].

Veronderstel 'n K -gemiddelde model met $K = 2$ vir 2-dimensionele data het gemiddelde by $A(-1, 2)$ en $B(3, -1)$. Aanvaar die volgende as-etiket: x_1 [horisontaal] en x_2 [vertikaal].

Compute the equation for the decision boundary separating the regions corresponding to the two means, in the form $x_2 = mx_1 + c$.

Bereken die vergelyking vir die skeidingslyn tussen die gebiede wat met die twee gemiddeldes ooreenstem, in die vorm $x_2 = mx_1 + c$.


Select one:

☐ a.

$$x_2 = -\frac{3}{4}x_1 - \frac{5}{6}$$

☒ b.

$$x_2 = \frac{4}{3}x_1 - \frac{5}{6}$$

 ☐ c. None of the options are correct/ Nie een van die opsies is korrek nie.

☐ d.

$$x_2 = -\frac{3}{4}x_1 - \frac{6}{5}$$

☐ e.

$$x_2 = \frac{4}{3}x_1 - \frac{6}{5}$$

Your answer is correct.

The correct answer is:

$$x_2 = \frac{4}{3}x_1 - \frac{5}{6}$$

Question **8**

Incorrect

Mark 0.00 out of
1.00

Let d and N denote the dimensionality of your data and the number of observations, respectively.

Laat d en N die dimensionaliteit en die aantal waarnemings van u data respektiewelik voorstel.

Find the data matrix \mathbf{X} ($d \times N$) below:

Vind die data matriks \mathbf{X} ($d \times N$) hieronder:

$$\begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}$$

Let $\mathbf{u}_1 = [1, 1]^T$.

Laat $\mathbf{u}_1 = [1, 1]^T$.

The above expressions can be realized with the following Python code:

Die bostaande uitdrukkings kan gerealiseer word met die volgende Python kode:

```
import numpy as np
X = np.array([[0,1],[2,3]])
u = np.array([1,1])
```

The following covariance update rule is associated with the Expectation Maximization algorithm for Gaussian Mixture Models:

Die volgende kovariansie opdateringsvergelyking kan geassosieer word met die Verwagting Optimering algoritme vir Gaussiese Mengsel Modelle:

$$\Sigma_j = \frac{1}{N_j} \sum_{n=1}^N \gamma(z_{nj})(\mathbf{x}_n - \mathbf{u}_j)(\mathbf{x}_n - \mathbf{u}_j)^T$$

where:

waar:

$$N_j = \sum_{n=1}^N \gamma_{nj}.$$

Let $\gamma(z_{11}) = 0.3$ and $\gamma(z_{21}) = 0.4$. Compute N_1 and Σ_1 and then choose the closest answer for the relevant quantities from the given drop-

down lists. (The notation $[\mathbf{X}]_{ij}$ refers to the ij -th element of \mathbf{X} .)

Laat $\gamma(z_{11}) = 0.3$ en $\gamma(z_{21}) = 0.4$. Bereken N_1 en Σ_1 en kies dan die naaste antwoorde vir die relevante waarde vanuit die gegewe kieserlyste. (Die notasie $[\mathbf{X}]_{ij}$ verwys na die ij -de element van \mathbf{X} .)

$[\Sigma_1]_{12}$ ✖

$[\Sigma_1]_{11}$ ✖

N_1 ✖

$[\Sigma_1]_{22}$ ✖

$[\Sigma_1]_{21}$ ✖

The correct answer is: $[\Sigma_1]_{12} \rightarrow -0.42857143$,
 $[\Sigma_1]_{11} \rightarrow 0.42857143$, $N_1 \rightarrow 0.7$, $[\Sigma_1]_{22} \rightarrow$
 2.71428571 , $[\Sigma_1]_{21} \rightarrow -0.42857143$