6406531562203. ✖ Red, Red

6406531562204. ✖ Green, Green

MLT

Section Id :	64065330337
Section Number :	11
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	17
Number of Questions to be attempted :	17
Section Marks :	100
Display Number Panel :	Yes
Group All Questions :	No
Enable Mark as Answered Mark for Review and Clear Response :	Yes
Maximum Instruction Time :	0
Sub-Section Number :	1
Sub-Section Id :	64065367703
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Number : 169 Question Id : 640653470095 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 0

Question Label : Multiple Choice Question

THIS IS QUESTION PAPER FOR THE SUBJECT "DIPLOMA LEVEL : MACHINE LEARNING TECHNIQUES"

ARE YOU SURE YOU HAVE TO WRITE EXAM FOR THIS SUBJECT?

CROSS CHECK YOUR HALL TICKET TO CONFIRM THE SUBJECTS TO BE WRITTEN.

(IF IT IS NOT THE CORRECT SUBJECT, PLS CHECK THE SECTION AT THE TOP FOR THE SUBJECTS REGISTERED BY YOU)

Options :

6406531562217. ✔ YES

6406531562218. ✖ NO

Sub-Section Number :	2
Sub-Section Id :	64065367704
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 170 Question Id : 640653470096 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Multiple Choice Question

Standard-PCA is performed on a centered dataset in $\mathbb{R}^3$. Two principal components are given below:

$$\frac{1}{2}\cdot\begin{bmatrix}1\\ \sqrt{2}\\ -1\end{bmatrix},\quad \frac{1}{\sqrt{2}}\cdot\begin{bmatrix}1\\ 0\\ 1\end{bmatrix}$$

Which of the following could be the third?

Options :

6406531562219. ✔ $\frac{1}{2}\cdot\begin{bmatrix}1\\ -\sqrt{2}\\ -1\end{bmatrix}$

6406531562220. ✖ $\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$

6406531562221. ✖ $\frac{1}{\sqrt{2}} \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

6406531562222. ✖ $\frac{1}{\sqrt{3}} \cdot \begin{bmatrix} \sqrt{2} \\ -1 \\ 0 \end{bmatrix}$

Question Number : 171 Question Id : 640653470097 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Multiple Choice Question

Consider that 1000 data points belonging to d-dimensional space have a non-linear relationship. We apply kernel PCA to reduce the dimension of the data points and take the first k principal components. Can the value of k be larger than d?

Options :

6406531562223. ✔ Yes

6406531562224. ✖ No

Question Number : 172 Question Id : 640653470098 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Multiple Choice Question

Let X be the data matrix of shape $d \times n$ with $d > n$ for a centered dataset. The eigenvector corresponding to the largest eigenvalue λ of $X^T X$ is α_1. What will be the first principal component of the dataset?

Options :

6406531562225. ✖ α_1

6406531562226. ✖ $\frac{\alpha_1}{\sqrt{\lambda}}$

6406531562227. ✖ $X\alpha_1$

6406531562228. ✔ $\frac{X\alpha_1}{\sqrt{\lambda}}$

Question Number : 173 Question Id : 640653470099 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Multiple Choice Question

Is the following statement true or false?

For any desired transformation $\phi(x)$, we can design a kernel function $k(x_1, x_2)$ that will evaluate $\phi(x_1)^T\phi(x_2)$.

Options :

6406531562229. ✔ TRUE

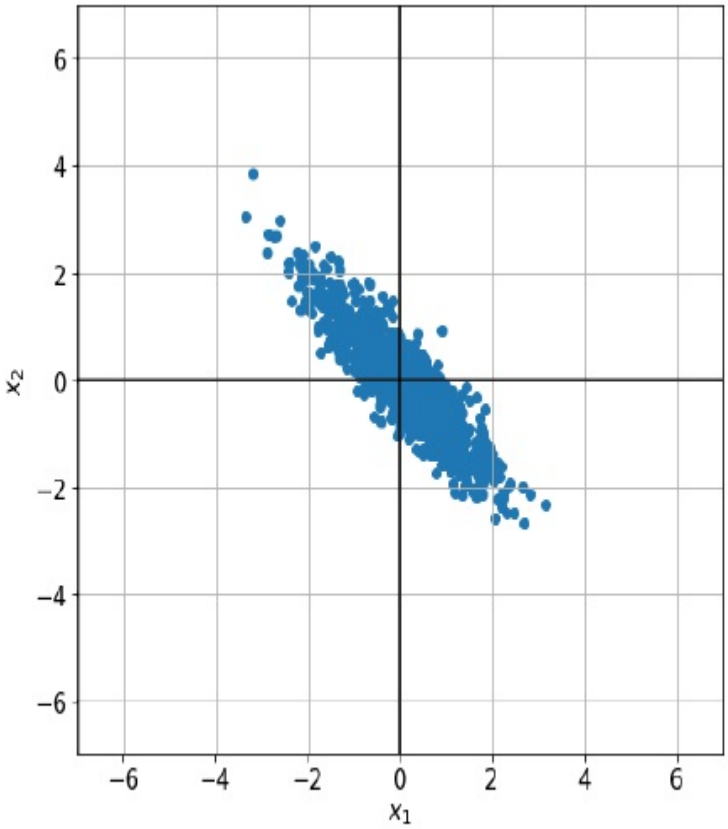
6406531562230. ✖ FALSE

Question Number : 174 Question Id : 640653470100 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Multiple Choice Question

Consider the following centered dataset in $\mathbb{R}^2$:



Which of the following could be the first principal component? Recall that the first P.C is the most important one.

Options :

6406531562231. ✖ $\begin{bmatrix}1\\0\end{bmatrix}$

6406531562232. ✖ $\begin{bmatrix}0\\1\end{bmatrix}$

6406531562233. ✖ $\frac{1}{\sqrt{2}} \cdot \begin{bmatrix}1\\1\end{bmatrix}$

6406531562234. ✔ $\frac{1}{\sqrt{2}} \cdot \begin{bmatrix}-1\\1\end{bmatrix}$

Sub-Section Number : 3

Sub-Section Id : 64065367705

Question Shuffling Allowed : Yes

Is Section Default? : null

Question Number : 175 Question Id : 640653470101 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5 Selectable Option : 0

Question Label : Multiple Select Question

In standard PCA, which of the following are correct formulations of the optimization problem? The dataset $\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$, is centered, each point lies in $\mathbb{R}^d$, and $\mathbf{C}$ is the covariance matrix. $\mathbf{w} \in \mathbb{R}^d$ is the variable that we are optimizing over.

Options :

6406531562235. ✔

$$\min_{\mathbf{w}} \quad \frac{1}{n} \cdot \sum_{i=1}^{n} ||\mathbf{x}_i - (\mathbf{x}_i^T \mathbf{w})\mathbf{w}||^2$$
$$\text{subject to} \quad ||\mathbf{w}|| = 1$$

6406531562236. ✔

$$\max_{\mathbf{w}} \quad \mathbf{w}^T \mathbf{C} \mathbf{w}$$
$$\text{subject to} \quad ||\mathbf{w}|| = 1$$

6406531562237. ✖

$$\max_{\mathbf{w}} \quad \frac{1}{n} \cdot \sum_{i=1}^{n} ||\mathbf{x}_i - (\mathbf{x}_i^T \mathbf{w})\mathbf{w}||^2$$
$$\text{subject to} \quad ||\mathbf{w}|| = 1$$

6406531562238. ✖

$$\min_{\mathbf{w}} \quad \mathbf{w}^T \mathbf{C} \mathbf{w}$$
$$\text{subject to} \quad ||\mathbf{w}|| = 1$$

Question Number : 176 Question Id : 640653470102 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5 Selectable Option : 0

Question Label : Multiple Select Question

Consider a GMM with 3 components that is used to model a dataset that has 100 points. The EM algorithm is run on this dataset to estimate the parameters of the GMM. After convergence, some of the values of λ_k^i for the specific point $\mathbf{x}_{10}$ are given below:

$$\lambda_1^{10} = 0.2, \quad \lambda_2^{10} = 0.5$$

Select all true statements.

Options :

6406531562239. ✖ If a point is picked randomly from the dataset of 100 points, then there is 20% chance that it comes from the first component.

6406531562240. ✔ $\lambda_3^{10} = 0.3$

6406531562241. ✖ $\sum_{i=1}^{100} \lambda_3^i = 1$

6406531562242. ✔ There is a 50% chance that the point $\mathbf{x}_{10}$ comes from the second component.

Question Number : 177 Question Id : 640653470103 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5 Selectable Option : 0

Question Label : Multiple Select Question

Consider a dataset with n points and a GMM with K components. If we fix λ and maximize for θ, what is our estimate for the mean of the k^{th} component?

Options :

6406531562243. ✔ It is the weighted mean of the n points, where the weight for point i in component k is given by λ_k^i

6406531562244. ✖ It is the mean of the n points.

6406531562245. ✖ $$\hat{\mu}_k = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$$

6406531562246. ✔ $$\hat{\mu}_k = \frac{\sum_{i=1}^{n} \lambda_k^i x_i}{\sum_{i=1}^{n} \lambda_k^i}$$

Question Number : 178 Question Id : 640653470104 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5 Selectable Option : 0

Question Label : Multiple Select Question

With respect to the Lloyd's algorithm, choose the correct statements:

Options :

6406531562247. ✔ The partition configurations cannot repeat themselves.

6406531562248. ✖ After doing the reassignments (consider at least one point reassigned to the new cluster), we might get the same means for all clusters.

6406531562249. ✔ Objective function after making the re-assignments strictly reduces.

6406531562250. ✖ Objective function after making the re-assignments may increase.

6406531562251. ✔ A change in the objective function's value indicates that the partition configuration has changed.

6406531562252. ✖ For partitioning n data points across k partitions, Lloyd's algorithm takes k^n iterations to converge.

Sub-Section Number :	4
Sub-Section Id :	64065367706
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 179 Question Id : 640653470105 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Short Answer Question

Consider a dataset of 20 points where the i^{th} data-point is given by:

$$\mathbf{x}_i = a_i \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix} + b_i \cdot \begin{bmatrix} -1 \\ 1 \\ 3 \\ 0 \end{bmatrix}$$

where, a_i and b_i are real numbers such that $\sum_{i=1}^{20} a_i = \sum_{i=1}^{20} b_i = 0$. Standard PCA is performed on this dataset. If the top two principal components are retained and used to reconstruct the dataset, what is the reconstruction error?

Hint: Think about what happens in $\mathbb{R}^2$ or $\mathbb{R}^3$ for a similar situation and extend this idea to $\mathbb{R}^4$.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

0

Question Number : 180 Question Id : 640653470106 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Short Answer Question

Consider the following prior for the parameter p of a Bernoulli distribution:

$$p \sim \text{Beta}(3, 2)$$

The dataset observed is as follows:

$$\{1, \quad 0, \quad 1, \quad 0, \quad 1, \quad 1, \quad 1, \quad 0, \quad 0, \quad 0, \quad 1, \quad 1, \quad 1\}$$

What is $\hat{p}$, a point estimate for the parameter of the Bernoulli distribution, if we use the expectation of the posterior as the method of estimation? Enter your answer correct to three decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.60 to 0.62

Question Number : 181 Question Id : 640653470107 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Short Answer Question

A dataset containing 200 examples in three-dimensional space has been transformed into a higher-dimensional space using a polynomial kernel of degree two. What will be the dimension of the transformed feature space?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

10

Sub-Section Number : 5

Sub-Section Id : 64065367707

Question Shuffling Allowed : No

Is Section Default? : null

Question Id : 640653470108 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (182 to 183)

Question Label : Comprehension

Consider a centered dataset in $\mathbb{R}^{100}$ that has 1000 data-points. Call this dataset D_1.
Standard PCA is performed on D_1 and the **scalar** projections of D_1
on the top 5 principal components are computed. Call the resulting dataset D_2.

NOTE: D_2 is only made up of the scalar projections. All the principal components are thrown out after computing the scalar projections and are not a part of D_2.

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 182 Question Id : 640653470109 Question Type : SA Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2.5

Question Label : Short Answer Question

Each data-point in D_2
belongs to $\mathbb{R}^k$. What is
the value of k?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

5

Question Number : 183 Question Id : 640653470110 Question Type : SA Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2.5

Question Label : Short Answer Question

If each real number occupies unit storage space in the memory, compute:

$$\frac{\text{size}(D_1)}{\text{size}(D_2)}$$

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

20

Sub-Section Number :	6
Sub-Section Id :	64065367708
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Id : 640653470111 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (184 to 185)

Question Label : Comprehension

Assume that you have a dataset of five points $\{x_1, x_2, x_3, x_4, x_5\}$, all of which are non-negative. You hypothesise that the data points are iid random variables with the following density:

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 184 Question Id : 640653470112 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Multiple Choice Question

What is the log-likelihood of this dataset under this distribution? ln represents the natural logarithm or $\log_e$.

Options :

6406531562258. ✖ $\prod_{i=1}^{5} \lambda e^{-\lambda x_i}$

6406531562259. ✖ $\sum_{i=1}^{5} \lambda e^{-\lambda x_i}$

6406531562260. ✔ $\sum_{i=1}^{5} [\ln(\lambda) - \lambda x_i]$

6406531562261. ✖ $\prod_{i=1}^{5} [\ln(\lambda) - \lambda x_i]$

6406531562262. ✖ $\sum_{i=1}^{5} [\ln(\lambda) - \lambda \ln(x_i)]$

Question Number : 185 Question Id : 640653470113 Question Type : SA Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 5

Question Label : Short Answer Question

You are given the actual values of these observations:

$x_1 = 1, \quad x_2 = 2, \quad x_3 = 3, \quad x_4 = 4, \quad x_5 = 5$

What is the maximum likelihood estimate for λ?
Enter your answer correct to three decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.32 to 0.34

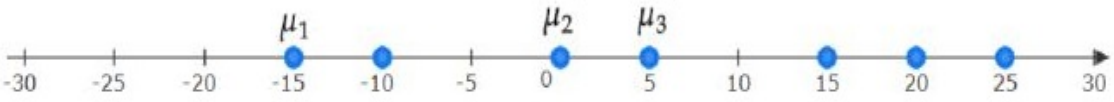
Sub-Section Number :	7
Sub-Section Id :	64065367709
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Id : 640653470114 Question Type : COMPREHENSION Sub Question Shuffling Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (186 to 190)

Question Label : Comprehension

Consider the following one-dimensional dataset of seven points that are distributed as follows:



k-means algorithm with $k = 3$ was run on the given data points. μ_1, μ_2, and μ_3 are the initial cluster means.

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 186 Question Id : 640653470115 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Short Answer Question

How many points belong to cluster 3 (mean μ_3) for the initial clusters?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

4

Question Number : 187 Question Id : 640653470116 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Short Answer Question

What will be the mean μ_3 after 1^{st} iteration? Enter your answer correct to two decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

16.23 to 16.27

Question Number : 188 Question Id : 640653470117 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 3

Question Label : Short Answer Question

How many points belong to cluster 2 (mean μ_2) after the 1^{st} iteration?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

2

Question Number : 189 Question Id : 640653470118 Question Type : SA Calculator : None

Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 4

Question Label : Short Answer Question

How many iterations will it take for the algorithm to converge with the given initial clusters?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

2

Question Number : 190 Question Id : 640653470119 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 4

Question Label : Multiple Choice Question

What will be the final cluster means?

Options :

6406531562268. ✖ $\mu_1 = -15,\ \mu_2 = 0,\ \mu_3 = 5$

6406531562269. ✖ $\mu_1 = -10,\ \mu_2 = 0,\ \mu_3 = 15$

6406531562270. ✖ $\mu_1 = -12.5,\ \mu_2 = 0,\ \mu_3 = 16.25$

6406531562271. ✖ $\mu_1 = -12.5,\ \mu_2 = 2.5,\ \mu_3 = 16.25$

6406531562272. ✔ $\mu_1 = -12.5,\ \mu_2 = 2.5,\ \mu_3 = 20$

Sub-Section Number :	8
Sub-Section Id :	64065367710
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Id : 640653470120 Question Type : COMPREHENSION Sub Question Shuffling

Allowed : No Group Comprehension Questions : No Question Pattern Type : NonMatrix Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Question Numbers : (191 to 192)

Question Label : Comprehension

A k-means++ algorithm with k = 3 was applied to following 2D points:

(0,0),(1,2),(3,1),(4,7),(-1,9),(4,-2)

(0,0) is chosen as the first cluster mean.

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 191 Question Id : 640653470121 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 4

Question Label : Multiple Choice Question

Which point has the highest probability of being chosen as the 2^{nd} cluster mean? Use the manhattan distance to compute the distances.

Options :

6406531562273. ✖ (3,1)

6406531562274. ✔ (4,7)

6406531562275. ✖ (-1,9)

6406531562276. ✖ (4,-2)

Question Number : 192 Question Id : 640653470122 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 4

Question Label : Multiple Choice Question

If the point with the highest score is chosen as the 2^{nd} cluster mean(answer from previous

question), Which point has the highest probability of being chosen as 3rd the cluster mean? Use the manhattan distance to compute the distances.

Options :

6406531562277. ✖ (3,1)

6406531562278. ✖ (4,7)

6406531562279. ✔ (-1,9)

6406531562280. ✖ (4,-2)

MLP

Section Id :	64065330338
Section Number :	12
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	24
Number of Questions to be attempted :	24
Section Marks :	50
Display Number Panel :	Yes
Group All Questions :	No
Enable Mark as Answered Mark for Review and Clear Response :	Yes
Maximum Instruction Time :	0
Sub-Section Number :	1
Sub-Section Id :	64065367711
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Number : 193 Question Id : 640653470123 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0