NOTE: Enter your answer to the nearest integer.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

4

MLT

Number of Questions :	15
Section Marks :	50

Question Number : 99 Question Type : MCQ

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 :

A. ✔ Yes

B. ✖ No

Question Number : 100 Question Type : MCQ

Correct Marks : 0

Question Label : Multiple Choice Question

Instructions:

1. Read the questions carefully.
2. Answers should be up to correct decimal places.
3. Always add dummy feature in the beginning of feature matrix where necessary.
4. Answers are case sensitive.
5. Do NOT include single or double quotes in your typed answers.

Useful Data

Inference using logistic regression happens as follows. T is called the threshold and is some real number in the interval $(0, 1)$. $\hat{y}$ stands for the predicted label.

$$\hat{y} = \begin{cases} 1, & P(y = 1 \mid x) \geq T \\ 0, & \text{otherwise} \end{cases}$$

The threshold is usually set to 0.5 unless otherwise specified.

Options :

A. ✔ Useful Data has been mentioned above.

B. ✖ This data attachment is just for a reference & not for an evaluation.

Question Number : 101 Question Type : MCQ

Correct Marks : 2

Question Label : Multiple Choice Question

A logistic regression model is tested on a test set comprising of 100 datapoints. The label 1 corresponds to the positive class and 0 to the negative class. All evaluation metrics should be computed with respect to the positive class.

Consider the following table:

True Label	$P(y=1 \mid x)$	Number of points
0	$(0, 0.5)$	30
1	$(0, 0.5)$	5
0	$(0.5, 0.75)$	15
1	$(0.5, 0.75)$	15
0	$(0.75, 1)$	5
1	$(0.75, 1)$	30

The table is to be understood as follows. The first row shows that there are 30 points in the test set from class 0 for which the predicted probability falls in the range (0, 0.5). Consider two models based on these predictions, $M_{0.5}$ and $M_{0.75}$,which have thresholds of $T = 0.5$ and $T = 0.75$ respectively. Which of these two models has higher precision?

Options :

A. ✖ $M_{0.5}$

B. ✔ $M_{0.75}$

C. ✖ Both have the same precision

D. ✖ Insufficient information to compute the precision of both the models

Question Number : 102 Question Type : MCQ

Correct Marks : 2

Question Label : Multiple Choice Question

A metal detector can be seen as performing a binary classification task: presence of a metal corresponds to the positive class (label 1) and its absence corresponds to the negative class (label 0).

A really good metal detector should be able to correctly identify almost all objects that have metal content in them. If it incorrectly classifies even a single object that has metal content in it, that could have serious consequences. In this process it may classify some harmless, metal-free objects as belonging to the positive class. But that is a price we are willing to pay.

You have a logistic regression model trained for this task. What can you say about the choice of threshold for a really good metal detector?

Options :

A. ✔ The threshold should be a low value.

B. ✖ The threshold should be a high value.

C. ✖ The performance of the model is independent of the threshold

Question Number : 103 Question Type : MCQ

Correct Marks : 2

Question Label : Multiple Choice Question

A logistic regression model is trained on a dataset where each data-point has m + 1 features (including the dummy feature). The weight vector obtained at the end of the training process is $\mathbf{w_0}$ and has shape m+1. Note that the first component of the weight vector is the weight corresponding to the dummy feature. Let us call this model M_0.

Consider two different models M_1 and M_2 generated from M_0, with the following parameters:

$$\mathbf{w_1} = \frac{\mathbf{w_0}}{2}, \quad \text{for } M_1$$
$$\mathbf{w_2} = 2\mathbf{w_0}, \quad \text{for } M_2$$

All three models operate at the same threshold of 0.5. Which of the following statements is true regarding the accuracies A_0, A_1 and A_2 of the three models M_0, M_1 and M_2 on the same test set?

Options :

A. ✖ $A_0 > A_1 > A_2$

B. ✖ $A_1 = \frac{A_0}{2}$ and $A_2 = 2A_0$

C. ✖ $A_0 < A_1 < A_2$

D. ✔ $A_0 = A_1 = A_2$

Question Number : 104 Question Type : MCQ

Correct Marks : 2

Question Label : Multiple Choice Question

Consider a binary classification problem for text documents where the two classes are **spam** and **not-spam**. A Multinomial Naive Bayes model is fit on the training data. The resulting parameter estimates for some words in the dataset are given below:

Word	Class	$P(\text{Word} \mid \text{Class})$
want	spam	0.03
want	not-spam	0.04
more	spam	0.02
more	not-spam	0.09
money	spam	0.1
money	not-spam	0.01

What will be the model's predicted label for the message "want more more money"? Assume that the prior probability is equal to 0.5 for each class.

Options :

A. ✖ spam

B. ✔ not-spam

Question Number : 105 Question Type : MCQ

Correct Marks : 2

Question Label : Multiple Choice Question

Consider two feature vectors of a dataset $\mathbf{x}$ and $\mathbf{y}$ as follows:

$$\mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ 4 \\ 3 \\ 2 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} 3 \\ 3 \\ 1 \\ 2 \\ 2 \end{bmatrix}$$

Apply kernel trick on these two vectors and select the correct answer from the following options.

Options :

A. ✖ 23

B. ✔ 529

C. ✖ [5,4,5,6,4]

D. ✖ [13,10,17,13,8]

Question Number : 106 Question Type : SA

Correct Marks : 2

Question Label : Short Answer Question

The weight vector for a trained support vector machine model on a linearly separable dataset is **w** = [1, 2,−2, 1,−1,−3, 2,−3, 0,−1]. Compute the width of the margin and enter your answer up to two decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.3 to 0.4

Question Type : COMPREHENSION

Question Numbers : (107 to 108)

Question Label : Comprehension

A logistic regression model is used for a binary classification problem with labels 0 and 1 that has m features, excluding the dummy feature. The weights corresponding to the m features are represented by $\boldsymbol{w}$, a vector of size m. The weight corresponding to the dummy feature is given by b, a scalar. The sigmoid activation function is:

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 107 Question Type : MCQ

Correct Marks : 2

Question Label : Multiple Choice Question

Given a feature vector $\boldsymbol{x}$, which of the following is the correct expression for the probability: $P(y = 1 \mid \boldsymbol{x})$?

Options :

A. ✔ $\sigma(\boldsymbol{w}^T\boldsymbol{x} + b)$

B. ✖ $\sigma(\boldsymbol{w}^T\boldsymbol{x})$

C. ✖ $\boldsymbol{w}^T\boldsymbol{x} + b$

D. ✖ $\sigma(\boldsymbol{w}^T\boldsymbol{x}) + b$

Question Number : 108 Question Type : MCQ

Correct Marks : 2

Question Label : Multiple Choice Question

Inference using logistic regression happens as follows. T is called the threshold and is some real number in the interval $(0, 1)$. $\hat{y}$ stands for the predicted label.

$$\hat{y} = \begin{cases} 1, & P(y = 1 \mid x) \geq T \\ 0, & \text{otherwise} \end{cases}$$

Given this setup, what is the equation of the decision boundary? ln is the natural logarithm or $\log_e$.

Options :

A. ✖ $w^T x + b = 0$

B. ✔ $w^T x + b + \ln\left(\frac{1}{T} - 1\right) = 0$

C. ✖ $w^T x + b - \ln\left(\frac{1}{T} - 1\right) = 0$

D. ✖ $w^T x + b - \ln\left(1 - \frac{1}{T}\right) = 0$

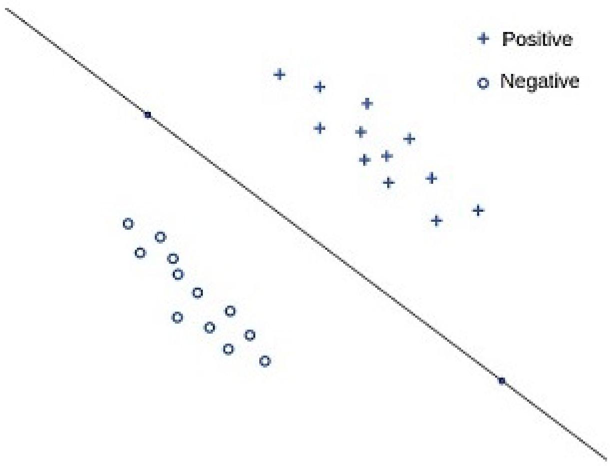
E. ✖ $w^T x + b + \ln\left(1 - \frac{1}{T}\right) = 0$

Question Type : COMPREHENSION

Question Numbers : (109 to 110)

Question Label : Comprehension

A logistic regression model has been trained using SGD for a linearly separable, binary classification problem, with labels 0 and 1. The decision boundary shown below is obtained after several epochs and is stable. That is, the weights no longer change with subsequent iterations of SGD. The learnt weights are stored in the vector $\boldsymbol{w}^*$. A threshold of 0.5 is used for all models.



Based on the above data, answer the given subquestions.

Sub questions

Question Number : 109 Question Type : MCQ

Correct Marks : 3

Question Label : Multiple Choice Question

If **x** is a single data-point that belongs to the negative class (label 0) and $P(y = 1 \mid \mathbf{x}) = y_{\text{prob}}$, then during the process of training which of the following is the correct expression for a single update of SGD with unit learning rate?

Options :

A. ✔ $w = w - y_{\text{prob}} x$

B. ✖ $w = w + y_{\text{prob}} x$

C. ✖ $w = w - (y_{\text{prob}} - 1) x$

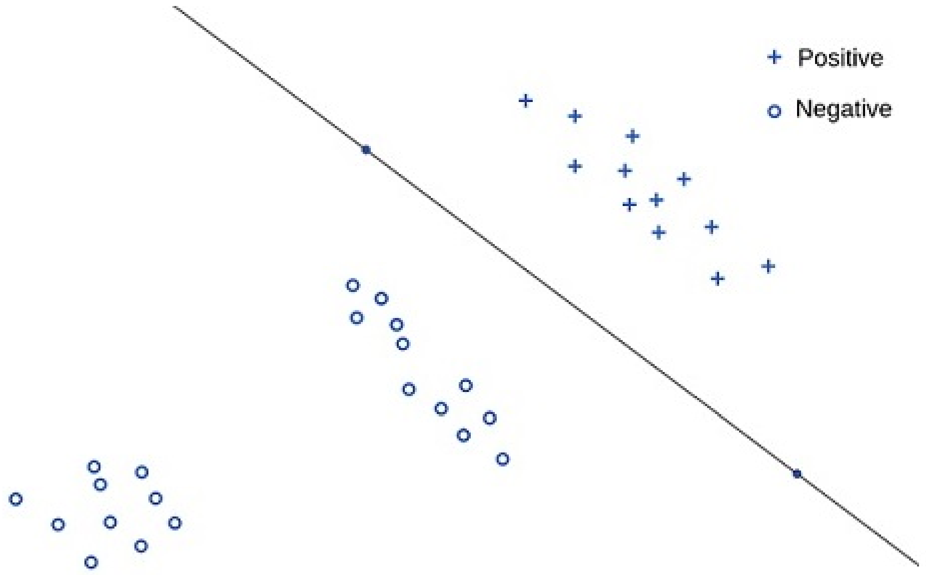
D. ✖ $w = w + (y_{\text{prob}} - 1)x$

Question Number : 110 Question Type : MCQ

Correct Marks : 2

Question Label : Multiple Choice Question

A new set of training data-points that belong to the negative-class, and that are **far away** from the existing decision boundary are introduced without disturbing the linear separability as shown below:



A new logistic regression model is trained on the entire dataset, which includes the new and the old points. However, before running SGD with unit learning rate, the weights are initialized to w^* rather than starting with a vector of zeros. We track the movement of the decision boundary through several epochs of SGD.

Select the most appropriate option.

Options :

A. ✖ The decision boundary will move a significant amount and will come to rest very close to the negative data-points.

B. ✖ The decision boundary will move a significant amount and will come to rest very close to the positive data-points.

C. ✔ There will be very little movement in the decision boundary and the final boundary will be extremely close to the initial one.

D. ✖ The decision boundary will keep oscillating between the positive and negative class and will never settle in a stable configuration.

Question Type : COMPREHENSION

Question Numbers : (111 to 112)

Question Label : Comprehension

Consider a regression problem where the labels lie in the range $(0, 1)$. We take the help of the sigmoid function to model this problem. Specifically, if the feature vector x is of size $m + 1$ (including the dummy feature) and the weight vector w has the same shape as x, then the model predicts a value $\hat{y}$ as follows:

$$\hat{y} = \sigma(w^T x)$$

We shall use the standard least square loss for regression. Use a normalizing factor of $\frac{1}{2}$ in the least square loss.

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 111 Question Type : MCQ

Correct Marks : 2

Question Label : Multiple Choice Question

If z is some real number find the derivative of $\sigma(z)$ with respect to z.

Options :

A. ✖ $\sigma(z) \cdot \sigma(1-z)$

B. ✖ $\sigma(z) \cdot \sigma(z)$

C. ✖ $\sigma(z) \cdot (1 + \sigma(z))$

D. ✔ $\sigma(z) \cdot (1 - \sigma(z))$

Question Number : 112 Question Type : MCQ

Correct Marks : 3

Question Label : Multiple Choice Question

Find the gradient of the loss with respect to the weight vector for a single data-point x with true label y. The model's prediction is $\hat{y}$.

Options :

A. ✔ $(\hat{y} - y) \cdot \hat{y} \cdot (1 - \hat{y}) \cdot \boldsymbol{x}$

B. ✖ $(\hat{y} - y) \cdot \boldsymbol{x}$

C. ✖ $(\hat{y} - y) \cdot (1 - \hat{y}) \cdot \boldsymbol{x}$

D. ✖ $\hat{y} \cdot (1 - \hat{y}) \cdot \boldsymbol{x}$

Question Type : COMPREHENSION

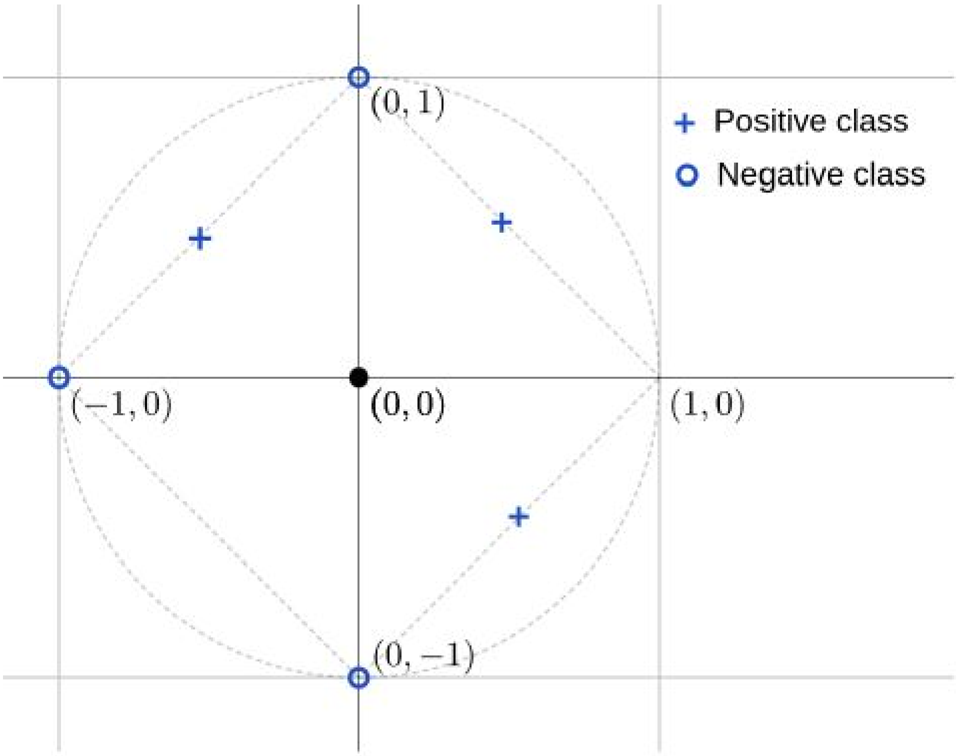
Question Numbers : (113 to 114)

Question Label : Comprehension

Consider a KNN classifier for a binary classification problem with k = 3. A test-point at (0, 0) has to be classified using this model. The training-dataset is as follows:

• The three negative class data-points are fixed and occupy three vertices of the diamond.

• The three positive class data-points can be anywhere on the edges of the diamond except the vertices



Based on the above data, answer the given subquestions.

Sub questions

Question Number : 113 Question Type : MCQ

Correct Marks : 3

Question Label : Multiple Choice Question

If the Manhattan distance metric is used, what is the predicted label of the test-point?

Options :

A. ✖ Positive class

B. ✖ Negative class

C. ✔ It could be either of the two classes. An exact decision requires information about how to break ties.

Question Number : 114 Question Type : MCQ

Correct Marks : 2

Question Label : Multiple Choice Question

If the Euclidean distance metric is used, what is the predicted label of the test-point?

Options :

A. ✔ Positive class

B. ✖ Negative class

C. ✖ It could be either of the two classes. An exact decision requires information about how to break ties.

Question Type : COMPREHENSION

Question Numbers : (115 to 116)

Question Label : Comprehension

Consider a Softmax regression model for a multi-class classification problem that has k classes. For an input vector $\mathbf{x}$, just before the Softmax, the linear combination is given by $\mathbf{z}$.

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 115 Question Type : MCQ

Correct Marks : 2

Question Label : Multiple Choice Question

$p_c = P(y = c \mid \mathbf{x})$ is the conditional probability that the point belongs to class c. Which of the following is the correct expression for p_c?

Options :

A. ✔ $\dfrac{e^{z_c}}{\sum_{i=1}^{k} e^{z_i}}$

B. ✖ $\dfrac{z_c}{\sum_{i=1}^{k} z_i}$

C. ✖ $\dfrac{z_c}{\sum_{i=1}^{k} e^{z_i}}$

D. ✖ $\dfrac{\sum_{i=1}^{k} e^{z_i}}{e^{z_c}}$

Question Number : 116 Question Type : MCQ

Correct Marks : 3

Question Label : Multiple Choice Question

Assume that we have a new input vector $\boldsymbol{x}^*$ for which the linear combination just before Softmax turns out to be $2\boldsymbol{z}$. Which of the following is the correct expression for $P(y = c \mid \boldsymbol{x}^*)$? The value of p_c is the one computed in the previous question.

Options :

A. ✖ p_c

B. ✖ $2p_c$

C. ✖ $\dfrac{p_c}{\sum_{i=1}^{k} p_i}$

D. ✔ $\dfrac{(p_c)^2}{\sum_{i=1}^{k} (p_i)^2}$

Question Type : COMPREHENSION

Question Numbers : (117 to 120)

Question Label : Comprehension

Consider a binary classification problem. The training-data has several features out of which we have access to only two binary features (x_1, x_2). The labels are 1 and 2 for the two classes. The training data set has the following distribution of points:

Feature	Number of points	True label
$(0, 0)$	20	1
$(0, 0)$	5	2
$(0, 1)$	1	1
$(0, 1)$	24	2
$(1, 0)$	18	1
$(1, 0)$	7	2
$(1, 1)$	5	1
$(1, 1)$	20	2

The table is to be parsed as follows. The first row of the table states that there are 20 points from class 1 that have $x_1 = 0, x_2 = 0$.

A Bernoulli Naive Bayes model is fit for this data with the following matrix of probabilities:

$$\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}$$

Each entry in this matrix can be understood as follows.

For $i, j \in \{1, 2\}$:

$$w_{ij} = P(x_i = 1 \mid y = j)$$

You can ignore smoothing. For all questions, report the answer up to two decimal places. Do not round-up or round-down the answer.

For example, if you get a value of 0.37914, just report 0.37.

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 117 Question Type : SA

Correct Marks : 1.5

Question Label : Short Answer Question

What is the value of w_{11}?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.51 to 0.53

Question Number : 118 Question Type : SA

Correct Marks : 1.5

Question Label : Short Answer Question

What is the value of w_{12}?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.47 to 0.49

Question Number : 119 Question Type : SA

Correct Marks : 1.5

Question Label : Short Answer Question

What is the value of w_{21}?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.12 to 0.14

Question Number : 120 Question Type : SA

Correct Marks : 1.5

Question Label : Short Answer Question

What is the value of w_{22}?

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

0.77 to 0.79

Question Type : COMPREHENSION

Question Numbers : (121 to 124)

Question Label : Comprehension

There is a linearly separable data set of 8 samples with 2 features as given in the below table.

Index	Data Points	Labels
a	(1,1)	1
b	(2,1.5)	-1
c	(1.5,1)	-1
d	(-0.5,1.5)	1
e	(-0.5,0.5)	1
f	(2,2)	-1
g	(2.5,1)	-1
h	(0.5,1)	1

Use support vector machine algorithm to develop a machine learning model that discriminates the two classes. By using visual inspection method identify the support vectors and use them to compute the weight vector ($\mathbf{w} = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$) and bias (b) of the separating hyperplane.

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 121 Question Type : SA

Correct Marks : 2

Question Label : Short Answer Question

Enter the value of $\mathbf{w_0}$ up to two decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

-5 to -3

Question Number : 122 Question Type : SA

Correct Marks : 2

Question Label : Short Answer Question

Enter the value of $\mathbf{w_1}$ up to two decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

2 to 3

Question Number : 123 Question Type : SA

Correct Marks : 2

Question Label : Short Answer Question

Enter the value of **b** up to two decimal places.

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Range

Text Areas : PlainText

Possible Answers :

2 to 3

Question Number : 124 Question Type : MCQ

Correct Marks : 2

Question Label : Multiple Choice Question

If there is a new point [1,−1.5], what is the predicted class label given by the SVM model you have developed?

Options :

A. Positive class

B. Negative class

AppDev 1

Number of Questions : 17

Section Marks : 50

Question Number : 125 Question Type : MCQ

Correct Marks : 0

Question Label : Multiple Choice Question

THIS IS QUESTION PAPER FOR THE SUBJECT "DIPLOMA LEVEL: MODERN APPLICATION DEVELOPMENT 1"

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 :

A. YES

B. NO

Question Number : 126 Question Type : MCQ

Correct Marks : 3

Question Label : Multiple Choice Question