Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

1

MLP

Section Id :	64065328986
Section Number :	12
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	22
Number of Questions to be attempted :	22
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 :	64065363354
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Number : 208 Question Id : 640653445653 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 PRACTICE"

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 :

6406531484909. ✔ Yes

6406531484910. ✖ No

Sub-Section Number : 2

Sub-Section Id : 64065363355

Question Shuffling Allowed : Yes

Is Section Default? : null

Question Number : 209 Question Id : 640653445654 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following options is the output of the following block of code?

```
from sklearn.linear_model import Perceptron
X=[(0,1),(0,2),(2,0),(3.5,3.5)]
y=[1,2,3,4]
clf = Perceptron()
clf.fit(X, y)
clf.score(X, y)
```

Options :

6406531484911. ✔ 1.0

6406531484912. ✖ 0.0

6406531484913. ✖ 0.5

6406531484914. ✖ 0.75

Question Number : 210 Question Id : 640653445659 Question Type : MCQ Is Question

Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 1

Question Label : Multiple Choice Question

If the data is expected to be already centered, is intercept estimation (fit _intercept) necessary for RidgeClassifier?

Options :

6406531484927. ✖ Yes, as intercepts are required to be modeled and fit, especially at this scenario, keep fit _intercept =True

6406531484928. ✔ No, as no intercept will be used in calculations: fit _intercept =False

Question Number : 211 Question Id : 640653445677 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 1

Question Label : Multiple Choice Question

In GradientBoostingRegressor, which of the following parameters is used to specify the fraction of samples to be used for fitting the individual base learners?

Options :

6406531484990. ✖ sample fraction

6406531484991. ✖ shrinkage

6406531484992. ✖ subspace

6406531484993. ✔ subsample

Sub-Section Number :	3
Sub-Section Id :	64065363356
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 212 Question Id : 640653445655 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2.5

Question Label : Multiple Choice Question

Which of the following options is the output of the following block of code?

```
import numpy as np
from sklearn.dummy import DummyClassifier
X = np.array([1,-512, 3.14j, 1])
y = np.array([0, 0, 0, 1])
dummy_clf = DummyClassifier(strategy="most_frequent")
dummy_clf.fit(X, y)
dummy_clf.predict(X)
```

Options :

6406531484915. ✖ array([0.5, 0.5, 0.5, 0.5])

6406531484916. ✖ SyntaxError:'j' is a complex number, hence DummmyClassifier failed

6406531484917. ✔ array([0, 0, 0, 0])

6406531484918. ✖ array([1, 1, 1, 1])

Question Number : 213 Question Id : 640653445664 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2.5

Question Label : Multiple Choice Question

Which of the following options represent the output of the following block of code?

```
import numpy as np
from sklearn.naive_bayes import CategoricalNB
rng = np.random.RandomState(1)
X = rng.randint(6, size=(3, 3))
y = np.array([1, 4, 6])
clf = CategoricalNB()
clf.fit(X, y)
print(clf.predict(X[2:3]))
#Output of the previous line, i.e., clf.predict(X[2:3])  is [6]
print(clf.score(X,y))
```

Options :

6406531484941. ✖ 0.5

6406531484942. ✖ 0.9

6406531484943. ✔ 1.0

6406531484944. ✖ 0.75

Question Number : 214 Question Id : 640653445665 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2.5

Question Label : Multiple Choice Question

Which of the following options is the correct output of the following block of code?

```
from sklearn.model_selection import LeaveOneOut
X = [0, 1, 1, 0]
loo = LeaveOneOut()
for train, test in loo.split(X):
    print("%s %s" % (train, test))
```

Options :

6406531484945. ✔
```
[1 2 3] [0]
[0 2 3] [1]
[0 1 3] [2]
[0 1 2] [3]
```

6406531484946. ✖
```
[1 2 3] [4]
[4 2 3] [1]
[4 1 3] [2]
[4 1 2] [3]
```

6406531484947. ✖
```
[1 1 0] [0]
[0 0 1] [1]
[0 1 0] [1]
[0 1 1] [0]
```

6406531484948. ✖ None of these

Question Number : 215 Question Id : 640653445666 Question Type : MCQ Is Question

Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2.5

Question Label : Multiple Choice Question

Which of the following options is the output of the possible block of code?

```
X = [[0], [1], [2], [3], [4]]
y = [0, 0, 1, 1, 1]

from sklearn.neighbors import KNeighborsClassifier
neigh = KNeighborsClassifier(n_neighbors=1)
neigh.fit(X, y)
print(neigh.predict([[3.1]]))
```

Options :

6406531484949. ✖ Output is [0]

6406531484950. ✔ Output is [1]

6406531484951. ✖ Output is [3]

6406531484952. ✖ Output is [3.1]

Question Number : 216 Question Id : 640653445667 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2.5

Question Label : Multiple Choice Question

Which of the following options is the correct output of the following block of code?

```
from sklearn.metrics import mean_squared_error
y_true = [3, -0.5, 2, 7]
y_pred = [2.5, 0.0, 2, 8]
mean_squared_error(y_true, y_pred, squared=False)
```

Options :

6406531484953. ✖ 0.75

6406531484954. ✖ 0.375

6406531484955. ✔ 0.612

6406531484956. ✖ 2.839

Question Number : 217 Question Id : 640653445668 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2.5

Question Label : Multiple Choice Question

Consider the block of code given below:

```
from sklearn.metrics import confusion_matrix
y_true = ["cat", "ant", "cat", "cat", "ant", "bird"]
y_pred = ["ant", "ant", "cat", "cat", "ant", "cat"]
confusion_matrix(y_true, y_pred, labels=["ant", "bird", "cat"])
```

The resulting confusion matrix from this code is following:

```
array([[2, 0, 0],
       [0, 0, 1],
       [1, 0, 2]])
```

What will be the precision value for the class “ant”?

Options :

6406531484957. ✖ 0.75

6406531484958. ✖ 0.0

6406531484959. ✖ 1.00

6406531484960. ✔ 0.67

Question Number : 218 Question Id : 640653445670 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2.5

Question Label : Multiple Choice Question

Which of the following options is the output of the following block of code?

```
import numpy as np
X = np.array([[0, 0], [1, 1], [2, 1]])
y = np.array([1, 3, 3])
from sklearn.svm import SVC
clf = SVC()
clf.fit(X, y)
print(clf.predict([[2, 3]]))
```

Options :

6406531484965. ✔ [3]

6406531484966. ✖ [1]

6406531484967. ✖ [2]

6406531484968. ✖ [0]

Question Number : 219 Question Id : 640653445672 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2.5

Question Label : Multiple Choice Question

In the Figure 1, different classification boundaries are shown with various regularization parameters. If regularization parameters for Figures 1(A), 1(B) and 1(C) are C_A, C_B, C_C correspondingly, then find out which of the following option represent the relation between C_A, C_B, C_C?

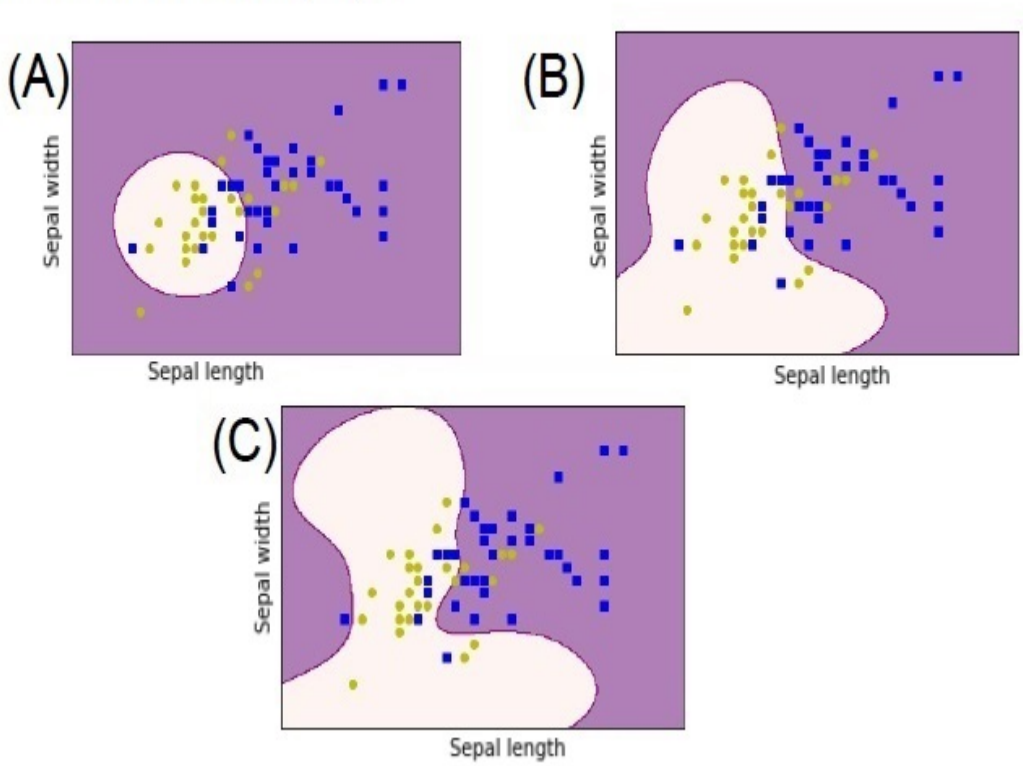

Figure 1

Options :

6406531484973. ✔ $C_A < C_B < C_C$

6406531484974. ✖ $C_A > C_B < C_C$

6406531484975. ✖ $C_A < C_B > C_C$

6406531484976. ✖ $C_A > C_B > C_C$

Sub-Section Number :	4
Sub-Section Id :	64065363357
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 220 Question Id : 640653445669 Question Type : MCQ Is Question

Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

In *HashingVectorizer* class, *build_analyzer()* method is used to

Options :

6406531484961. ✖ Transform a sequence of documents to a document-term matrix.

6406531484962. ✔ Return a callable to process input data.

6406531484963. ✖ Return a function to preprocess the text before tokenization.

6406531484964. ✖ Return a function that splits a string into a sequence of tokens.

Question Number : 221 Question Id : 640653445678 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

Choose the correct statements:

1. Bagging decreases bias.
2. Bagging increases variance.
3. Bagging increases bias.
4. Bagging decreases variance

Options :

6406531484994. ✖ Only 1 and 2 are correct.

6406531484995. ✔ Only 3 and 4 are correct.

6406531484996. ✖ Only 1 and 3 are correct.

6406531484997. ✖ Only 2 and 4 are correct.

Question Number : 222 Question Id : 640653445679 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

Which of the following statements is correct?

Options :

6406531484998. ✔ In both boosting and bagging techniques, a set of weak learners whose predictive power is slightly more than guessing is used for training.

6406531484999. ✖ Boosting and Bagging techniques cannot be deployed on any other machine learning model except Decision Trees.

6406531485000. ✖ Boosting and Bagging perform better than Artificial Neural Network in all circumstances.

6406531485001. ✖ Boosting and Bagging are as interpretative as Decision Trees are.

Sub-Section Number :	5
Sub-Section Id :	64065363358
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 223 Question Id : 640653445674 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2 Selectable Option : 0

Question Label : Multiple Select Question

Consider following code snippet:

```
import pandas as pd
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import train_test_split

df = pd.read_csv('balance-scale.data')

# take last column as label, and rest columns as features
y = df[df.columns[-1]]
X = df[df.columns[1:]]

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2)

clf = DecisionTreeClassifier(random_state = 0)
clf.fit(X_train,y_train)
```

Which of the following statements can be used to compute the height of the trained decision tree model?

Options :

6406531484981. ✔ `clf.get_depth()`

6406531484982. ✔ `clf.tree_.max_depth`

6406531484983. ✖ `clf.tree_.get_depth()`

6406531484984. ✖ `clf.tree.get_depth`

Question Number : 224 Question Id : 640653445675 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2 Selectable Option : 0

Question Label : Multiple Select Question

Which of the following are correct statements regarding cost complexity pruning?

Options :

6406531484985. ✔ It is useful in reducing overfitting in a decision tree.

6406531484986. ✖ It is useful in reducing underfitting in a decision tree.

6406531484987. ✔ This technique allows the decision tree to grow as much as possible, then it clips parts of the tree.

6406531484988. ✖ This technique doesn't allow the decision tree to grow beyond a certain depth.

Sub-Section Number :	6
Sub-Section Id :	64065363359
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 225 Question Id : 640653445671 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2.5 Selectable Option : 0

Question Label : Multiple Select Question

Which of the following options are correct?

Options :

6406531484969. ✔ NuSVC implements the "one-versus-one" approach for multi-class classification.

6406531484970. ✔ LinearSVC implements "one-vs-the-rest" approach for multi-class classification.

6406531484971. ✔ SVC implements the "one-versus-one" approach for multi-class classification.

6406531484972. ✖ NuSVC implements the "one-versus-the-rest" approach for multi-class classification.

Question Number : 226 Question Id : 640653445673 Question Type : MSQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2.5 Selectable Option : 0

Question Label : Multiple Select Question

Mention which of the following statements are correct:

Options :

6406531484977. ✔ If we have a small number of hyperparameters and sufficient training time, we should apply GridSearchCV.

6406531484978. ✔ For a very large number of hyperparameters and less training time scenarios, RandomizedSearchCV will be more appropriate.

6406531484979. ✖ GridSearchCV takes smaller training time than RandomizedSearchCV if the number of hyperparameters is very large.

6406531484980. ✖ Number of hyperparameters and training time of the dataset, do not influence the selection of RandomizedSearchCV and GridSearchCV.

Sub-Section Number : 7

Sub-Section Id : 64065363360

Question Shuffling Allowed : No

Is Section Default? : null

Question Id : 640653445656 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 : (227 to 228)

Question Label : Comprehension

Based on the following block of code, prepare the confusion matrix and answer the given subquestions

```
from sklearn.metrics import confusion_matrix
x = [2, 0, 2, 2, 0, 2]
y = [0, 0, 2, 2, 0, 2]
confusion_matrix(x, y)
```

Sub questions

Question Number : 227 Question Id : 640653445657 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

Which of the following options represent the precision score?

Options :

6406531484919. ✔ 1.00

6406531484920. ✖ 0.75

6406531484921. ✖ 0.90

6406531484922. ✖ 0.00

Question Number : 228 Question Id : 640653445658 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

Which of the following options represent the recall value?

Options :

6406531484923. ✖ 1.00

6406531484924. ✔ 0.75

6406531484925. ✖ 0.90

6406531484926. ✖ 0.00

Sub-Section Number :	8
Sub-Section Id :	64065363361
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Id : 640653445660 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 : (229 to 231)

Question Label : Comprehension

Sunita wrote a code for logistic regression as shown below:

```
import numpy as np
from sklearn.linear_model import LogisticRegression
x = np.arange(10).reshape(-1, 1)
y = np.array([0, 0, 0, 0, 1, 1, 1, 1, 1, 1])
model = LogisticRegression(solver='liblinear', random_state=0).fit(x,y)
model.predict_proba(x)
```

The output of the code was following:

```
array([[0.74002157, 0.25997843],
       [0.62975524, 0.37024476],
       [0.5040632 , 0.4959368 ],
       [0.37785549, 0.62214451],
       [0.26628093, 0.73371907],
       [0.17821501, 0.82178499],
       [0.11472079, 0.88527921],
       [0.07186982, 0.92813018],
       [0.04422513, 0.95577487],
       [0.02690569, 0.97309431]])
```

Based on the above data, answer the given subquestions.

Sub questions

Question Number : 229 Question Id : 640653445661 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

Which of the following options represents the output if Sunita writes the following line of code after the given code block?

model.predict(x)

Options :

6406531484929. ✔ array([0, 0, 0, 1, 1, 1, 1, 1, 1, 1])

6406531484930. ✖ array([0, 0, 1, 0, 1, 1, 1, 1, 1, 1])

6406531484931. ✖ array([1, 1, 1, 0, 0, 0, 0, 0, 0, 0])

6406531484932. ✖ array([0, 0, 1, 1, 1, 1, 1, 1, 1, 1])

Question Number : 230 Question Id : 640653445662 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 1

Question Label : Multiple Choice Question

Which of the following options is the output of the following line of code?

model.classes_

Options :

6406531484933. ✖ Sparse matrix:[[0],[1]]

6406531484934. ✔ array([0, 1])

6406531484935. ✖ array([0, 1, 2, 3, 4, 5, 6, 7, 8, 9])

6406531484936. ✖ None of these

Question Number : 231 Question Id : 640653445663 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 2

Question Label : Multiple Choice Question

Which of the following options is the output of the following block of code?

```
print(model.score(x,y))
```

Options :

6406531484937. ✖ 1.0

6406531484938. ✔ 0.9

6406531484939. ✖ 0.5

6406531484940. ✖ 0.0

Sub-Section Number :	9
Sub-Section Id :	64065363362
Question Shuffling Allowed :	Yes
Is Section Default? :	null

Question Number : 232 Question Id : 640653445676 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

Consider following code snippet:

```
import pandas as pd
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import train_test_split

df = pd.read_csv('some-dataset.data')

# take last column as label, and rest columns as features
y = df[df.columns[-1]]
X = df[df.columns[1:]]

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2)

train_scores = []
test_scores = []

for i_depth in range(1,16):
  i_clf = DecisionTreeClassifier(max_depth = i_depth).fit(X_train, y_train)
  train_scores.append(i_clf.score(X_train, y_train))
  test_scores.append(i_clf.score(X_test, y_test))

plt.figure(figsize=(10,6))
plt.plot(range(1, 16), train_scores,c='b', label='Training')
plt.scatter(range(1, 16), train_scores,c='b')
plt.plot(range(1, 16), test_scores,c='r',label='Test')
plt.scatter(range(1, 16), test_scores,c='r')
plt.xlabel('Max Depth');plt.ylabel('Score')
plt.legend()
plt.xticks(list(range(1,16)),list(range(1,16)))
plt.grid(True);plt.show()
```

Above code block produces the following chart displayed below, Figure 2, what is the optimal value of tree depth based on this chart?

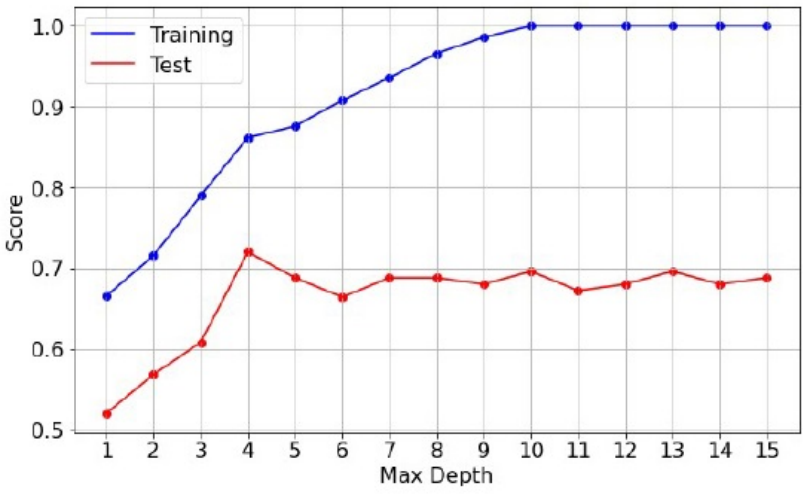

Figure 2

Response Type : Numeric

Evaluation Required For SA : Yes

Show Word Count : Yes

Answers Type : Equal

Text Areas : PlainText

Possible Answers :

4

BDM

Section Id :	64065328987
Section Number :	13
Section type :	Online
Mandatory or Optional :	Mandatory
Number of Questions :	13
Number of Questions to be attempted :	13
Section Marks :	16
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 :	64065363363
Question Shuffling Allowed :	No
Is Section Default? :	null

Question Number : 233 Question Id : 640653445680 Question Type : MCQ Is Question Mandatory : No Calculator : None Response Time : N.A Think Time : N.A Minimum Instruction Time : 0

Correct Marks : 0