Week 1 : Assignment 1

The due date for submitting this assignment has passed.

Due on 2021-08-18, 23:59 IST.

As per our records you have not submitted this assignment.

1) Which of the following is not a type of supervised learning? ***2 points***

- A. Classification
- B. Regression
- C. Clustering
- D. None of the above

No, the answer is incorrect.
Score: 0
Accepted Answers:
C. Clustering

2) As the amount of training data increases ***2 points***

- A. Training error usually decreases and generalization error usually increases
- B. Training error usually decreases and generalization error usually decreases
- C. Training error usually increases and generalization error usually decreases
- D. Training error usually increases and generalization error usually increases

No, the answer is incorrect.
Score: 0
Accepted Answers:
C. Training error usually increases and generalization error usually decreases

3) Suppose I have 10,000 emails in my mailbox out of which 300 are spams. The spam detection system detects 150 mails as spams, out of which 50 are actually spams. What is the precision and recall of my spam detection system ? ***2 points***

A. Precision = 33.33%, Recall = 16.66%

B. Precision = 25%, Recall = 33.33%

C. Precision = 33.33%, Recall = 75%

D. Precision = 75%, Recall = 33.33%

No, the answer is incorrect.
Score: 0

Accepted Answers:
A. Precision = 33.33%, Recall = 16.66%

4) Which of the following are not classification tasks ? ***2 points***

A. Find the gender of a person by analyzing his writing style

B. Predict the price of a house based on floor area, number of rooms etc.

C. Predict whether there will be abnormally heavy rainfall next year

D. Detect Pneumonia from Chest X-ray images

No, the answer is incorrect.
Score: 0

Accepted Answers:
B. Predict the price of a house based on floor area, number of rooms etc.

5) Occam's razor is an example of: ***2 points***

A. Inductive bias

B. Preference bias

No, the answer is incorrect.
Score: 0

Accepted Answers:
A. Inductive bias

6) A feature F1 can take certain value: A, B, C, D, E, F and represents grade of students from a college. Which of the following statements is true in the following case? ***2 points***

A. Feature F1 is an example of a nominal variable.

B. Feature F1 is an example of ordinal variables.

C. It doesn't belong to any of the above categories.

D. Both of these

No, the answer is incorrect.
Score: 0

Accepted Answers:
B. Feature F1 is an example of ordinal variables.

7) Which of the following is a categorical feature? ***2 points***

A. Height of a person

B. Price of petroleum

C. Mother tongue of a person

D. Amount of rainfall in a day

No, the answer is incorrect.
Score: 0

Accepted Answers:
C. Mother tongue of a person

8) Which of the following tasks is NOT a suitable machine learning task? ***2 points***

- A. Finding the shortest path between a pair of nodes in a graph
- B. Predicting if a stock price will rise or fall
- C. Predicting the price of petroleum
- D. Grouping mails as spams or non-spams

No, the answer is incorrect.
Score: 0

Accepted Answers:
A. Finding the shortest path between a pair of nodes in a graph

9) Which of the following is correct for reinforcement learning? ***2 points***

- A. The algorithm plans a sequence of actions from the current state.
- B. The algorithm plans one action at each time step.
- C. The training instances contain examples of states and best actions of the states.
- D. The algorithm groups unseen data based on similarity.

No, the answer is incorrect.
Score: 0

Accepted Answers:
B. The algorithm plans one action at each time step.

10) What is the use of Validation dataset in Machine Learning? ***2 points***

- A. To train the machine learning model.
- B. To evaluate the performance of the machine learning model
- C. To tune the hyperparameters of the machine learning model
- D. None of the above.

No, the answer is incorrect.
Score: 0

Accepted Answers:
C. To tune the hyperparameters of the machine learning model

Week 2 : Assignment 2

The due date for submitting this assignment has passed.

Due on 2021-08-18, 23:59 IST.

As per our records you have not submitted this assignment.

1) Identify whether the following statement is true or false? ***2 points***

"Overfitting is more likely when the set of training data is small"

- A. True
- B. False

No, the answer is incorrect.
Score: 0

Accepted Answers:
A. True

2) Which of the following criteria is typically used for optimizing in linear regression. ***2 points***

- A. Maximize the number of points it touches.
- B. Minimize the number of points it touches.
- C. Minimize the squared distance from the points.
- D. Minimize the maximum distance of a point from a line.

No, the answer is incorrect.
Score: 0

Accepted Answers:
C. Minimize the squared distance from the points.

3) Which of the following is false? ***2 points***

- A. Bias is the true error of the best classifier in the concept class
- B. Bias is high if the concept class cannot model the true data distribution well
- C. High bias leads to overfitting

D. For high bias both train and test error will be high

No, the answer is incorrect.
Score: 0

Accepted Answers:
C. High bias leads to overfitting

4) The following dataset will be used to learn a decision tree for predicting whether a person is happy (H) or sad (S), based on the color of shoes, whether they wear a wig and the number of ears they have. ***2 points***

Color	Wig	Num. Ears	Emotion (Output)
G	Y	2	S
G	N	2	S
G	N	2	S
B	N	2	S
B	N	2	H
R	N	2	H
R	N	2	H
R	N	2	H
R	Y	3	H

Which attribute should you choose as the root of the decision tree?

A. Color

B. Wig

C. Number of ears

D. Any one of the previous three attributes

No, the answer is incorrect.
Score: 0

Accepted Answers:
A. Color

5) ***2 points***

Consider applying linear regression with the hypothesis as $h_\theta(x) = \theta_0 + \theta_1 x$.The training data is given in the table.

X	Y
6	7
5	4
10	9
3	4

The cost function is $J(\theta) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x_i) - y_i)^2$

What is the value of $J(\theta)$when $\boldsymbol{\theta} = (2, 1)$?

- A. 0
- B. 1
- C. 2
- D. 2.5

No, the answer is incorrect.
Score: 0

Accepted Answers:
D. 2.5

6) In a binary classification problem, out of 64 data points 29 belong to class I and 35 belong to class II. What is the entropy of the data set? ***2 points***

- A. 0.97
- B. 0
- C. 1
- D. 0.99

No, the answer is incorrect.
Score: 0

Accepted Answers:
D. 0.99

7) Decision trees can be used for the following type of datasets: ***2 points***

I. The attributes are categorical
II. The attributes are numeric valued and continuous
III. The attributes are discrete valued numbers

- A. In case I only
- B. In case II only
- C. In cases II and III only
- D. In cases I, II and III

No, the answer is incorrect.
Score: 0

Accepted Answers:
D. In cases I, II and III

8) What is true for Stochastic Gradient Descent? ***2 points***

- A. In every iteration, model parameters are updated for multiple training samples
- B. In every iteration, model parameters are updated for one training sample
- C. In every iteration, model parameters are updated for all training samples
- D. None of the above

No, the answer is incorrect.
Score: 0

Accepted Answers:
B. In every iteration, model parameters are updated for one training sample

Week 3: Assignment 3

The due date for submitting this assignment has passed.

Due on 2021-08-25, 23:59 IST.

As per our records you have not submitted this assignment.

1) Suppose, you have given the following data where x and y are the 2 input variables and Class is the dependent variable. ***2 points***

X	Y	Class
-1	1	-
0	1	+
0	2	-
1	-1	-
1	0	+
1	2	+
2	2	-
2	3	+

Suppose, you want to predict the class of new data point x=1 and y=1 using euclidean distance in 7-NN. To which class the data point belongs to?

- A. + Class
- B. – Class
- C. Can't say

D. None of these

No, the answer is incorrect.
Score: 0

Accepted Answers:
B. – Class

2) Imagine you are dealing with 15 class classification problem. What is the maximum number of discriminant vectors that can be produced by LDA? ***2 points***

A. 20
B. 14
C. 21
D. 10

No, the answer is incorrect.
Score: 0

Accepted Answers:
B. 14

3) 'People who bought this, also bought…' recommendations seen on amazon is a result of which algorithm? ***2 points***

A. User based Collaborative filtering
B. Content based filtering
C. Item based Collaborative filtering
D. None of the above

No, the answer is incorrect.
Score: 0

Accepted Answers:
C. Item based Collaborative filtering

4) Which of the following is/are true about PCA? ***2 points***

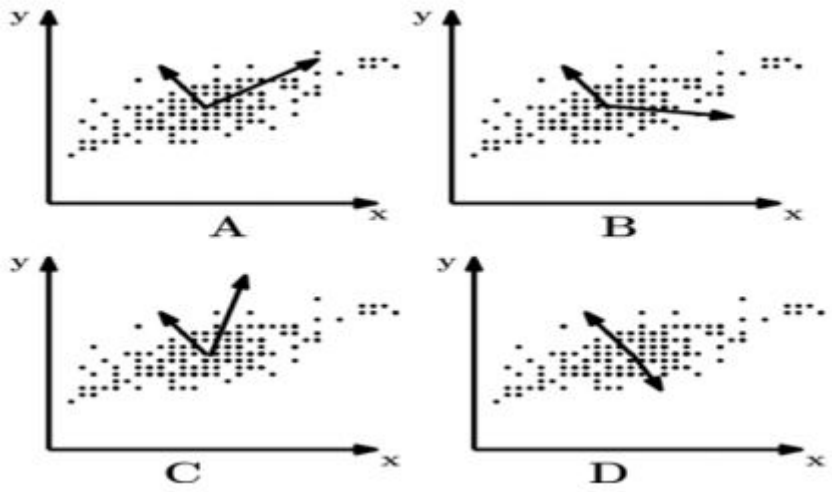
1. PCA is a supervised method
2. It identifies the directions that data have the largest variance
3. Maximum number of principal components <= number of features
4. All principal components are orthogonal to each other

A. Only 2
B. 1, 3 and 4
C. 1, 2 and 3
D. 2, 3 and 4

No, the answer is incorrect.
Score: 0

Accepted Answers:
D. 2, 3 and 4

5) Consider the figures below. Which figure shows the most probable PCA component directions for the data points? ***2 points***



- A. A
- B. B
- C. C
- D. D

No, the answer is incorrect.
Score: 0

Accepted Answers:
A. A

6) When there is noise in data, which of the following options would improve the performance of the KNN algorithm? ***2 points***

- A. Increase the value of k
- B. Decrease the value of k
- C. Changing value of k will not change the effect of the noise
- D. None of these

No, the answer is incorrect.
Score: 0

Accepted Answers:
A. Increase the value of k

7) Which of the following statements is True about the KNN algorithm? ***2 points***

- A. KNN algorithm does more computation on test time rather than train time.
- B. KNN algorithm does lesser computation on test time rather than train time.
- C. KNN algorithm does an equal amount of computation on test time and train time.
- D. None of these.

No, the answer is incorrect.
Score: 0

Accepted Answers:
A. KNN algorithm does more computation on test time rather than train time.

8) Find the value of the Pearson's correlation coefficient of X and Y from the data in the following table. ***2 points***

AGE (X)	GLUCOSE (Y)
43	99
21	65
25	79
42	75

- A. 0.47
- B. 0.68
- C. 1
- D. 0.33

No, the answer is incorrect.
Score: 0

Accepted Answers:
B. 0.68

Week 4 : Assignment 4

The due date for submitting this assignment has passed.

Due on 2021-09-01, 23:59 IST.

As per our records you have not submitted this assignment.

1) A spam filtering system has a probability of 0.95 to correctly classify a mail as spam and 0.10 probability of giving false positives. It is estimated that 1% of the mails are actual spam mails. Suppose that the system is now given a new mail to be classified as spam/ not-spam, what is the probability that the mail will be classified as spam? ***2 points***

- A. 0.89575
- B. 0.10425
- C. 0.1085
- D. 0.0995

No, the answer is incorrect.
Score: 0

Accepted Answers:
C. 0.1085

2) Bag I contains 4 white and 6 black balls while another Bag II contains 4 white and 3 black balls. One ball is drawn at random from one of the bags and it is found to be black. Find the probability that it was drawn from Bag I. ***2 points***

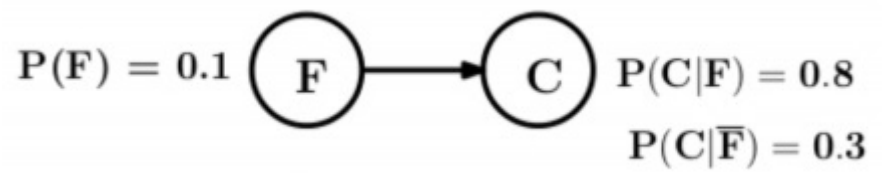
- A. 1/2
- B. 2/3
- C. 7/12
- D. 9/23

No, the answer is incorrect.
Score: 0

Accepted Answers:

C. 7/12

3) Consider the following Bayesian network, where F = having the flu and C = coughing: ***2 points***



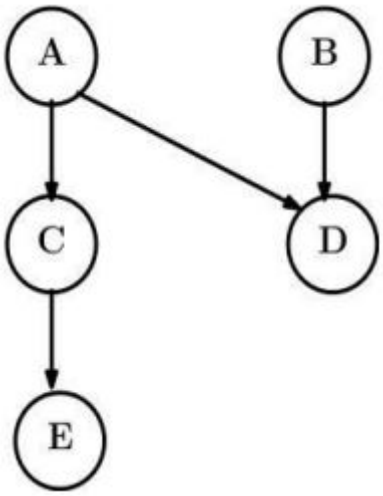
Find P(C) and P(F|C).

- A. 0.35, 0.23
- B. 0.35,0.77
- C. 0.24, 0.024
- D. 0.5, 0.23

No, the answer is incorrect.
Score: 0

Accepted Answers:
A. 0.35, 0.23

4) Consider the following Bayesian network. ***2 points***



Thus, the independence expressed in this Bayesian net are that
A and B are (absolutely) independent.
C is independent of B given A.
D is independent of C given A and B.
E is independent of A, B, and D given C.

Suppose that the net further records the following probabilities:

Suppose that the net further records the following probabilities:

$P(A) = 0.3$
$P(B) = 0.6$
$P(C|A) = 0.8$
$P(C|\overline{A}) = 0.4$
$P(D|A,B) = 0.7$
$P(D|A,\overline{B}) = 0.8$
$P(D|\overline{A},B) = 0.1$
$P(D|\overline{A},\overline{B}) = 0.2$
$P(E|C) = 0.7$
$P(E|C) = 0.7$

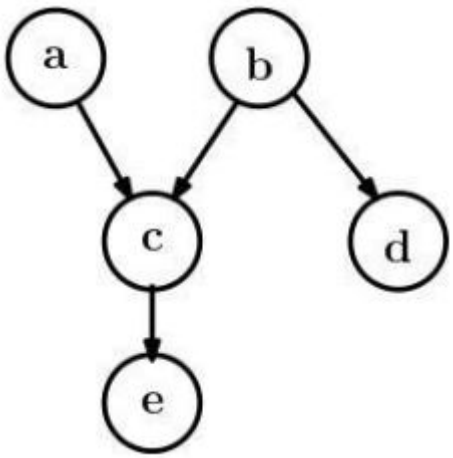
Find $P(D)$.

- A. 0.32
- B. 0.50
- C. 0.40
- D. 0.78

No, the answer is incorrect.
Score: 0

Accepted Answers:
A. 0.32

5) Consider the following graphical model, mark which of the following pair of random variables are independent given no evidence? ***2 points***



- A. a,b
- B. c,d
- C. e,d
- D. c,e

No, the answer is incorrect.
Score: 0

Accepted Answers:
A. a,b

6) In a Bayesian network a node with only outgoing edge(s) represents ***2 points***

- A. a variable conditionally independent of the other variables.

B. a variable dependent on its siblings.

C. a variable whose dependency is uncertain.

D. None of the above.

No, the answer is incorrect.
Score: 0

Accepted Answers:
A. a variable conditionally independent of the other variables.

7) ***2 points***

It is given that $P(A|B) = 2/3$ and $P(A|\overline{B}) = 1/3$. Compute the value of $P(B|A)$.

A. ½

B. ⅔

C. ¾

D. Not enough information.

No, the answer is incorrect.
Score: 0

Accepted Answers:
D. Not enough information.

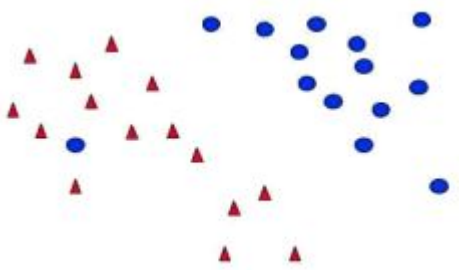
Week 5 : Assignment 5

The due date for submitting this assignment has passed.

Due on 2021-09-01, 23:59 IST.

As per our records you have not submitted this assignment.

1) What would be the ideal complexity of the curve which can be used for separating the two classes shown in the image below? ***2 points***



- A) Linear
- B) Quadratic
- C) Cubic
- D) insufficient data to draw conclusion

No, the answer is incorrect.
Score: 0

Accepted Answers:
A) Linear

2) I. Logistic Regression is used for regression purposes. ***2 points***
II. Logistic Regression is used for classification purposes.

- A) Only I is Correct
- B) Only II is Correct

C) Both I and II are Correct

D) Both I and II are Incorrect

No, the answer is incorrect.
Score: 0

Accepted Answers:
B) Only II is Correct

3) Which of the following methods do we use to best fit the data in Logistic Regression? ***2 points***

A) Least Square Error

B) Maximum Likelihood

C) Jaccard distance

D) Both A and B

No, the answer is incorrect.
Score: 0

Accepted Answers:
B) Maximum Likelihood

4) Consider a following model for logistic regression: $P(y=1|x,w)=g(w0+w1x)$ where g(z) is the logistic function. ***2 points***

In the above equation the $P(y=1|x; w)$, viewed as a function of x, that we can get by changing the parameters w.

What would be the range of P in such a case?

A) (-inf,0)

B) (0,1)

C) (-inf, inf)

D) (0,inf)

No, the answer is incorrect.
Score: 0

Accepted Answers:
B) (0,1)

5) **State whether True or False.** ***2 points***

After training an SVM, we can discard all examples which are not support vectors and can still classify new examples.

A) TRUE

B) FALSE

No, the answer is incorrect.
Score: 0

Accepted Answers:
A) TRUE

6) Suppose you are dealing with 3 class classification problem and you want to train a SVM model on the data for that you are using One-vs-all method. ***2 points***

How many times we need to train our SVM model in such case?

A) 1

B) 2

C) 3

D) 4

No, the answer is incorrect.
Score: 0

Accepted Answers:
C) 3

7) What is/are true about kernel in SVM? ***2 points***

1. Kernel function map low dimensional data to high dimensional space
2. It's a similarity function

A) 1

B) 2

C) 1 and 2

D) None of these.

No, the answer is incorrect.
Score: 0

Accepted Answers:
C) 1 and 2

8) Suppose you are using RBF kernel in SVM with high Gamma value. What doesthissignify? ***2 points***

A) The model would consider even far away points from hyperplane for modelling.

B) The model would consider only the points close to the hyperplane for modelling.

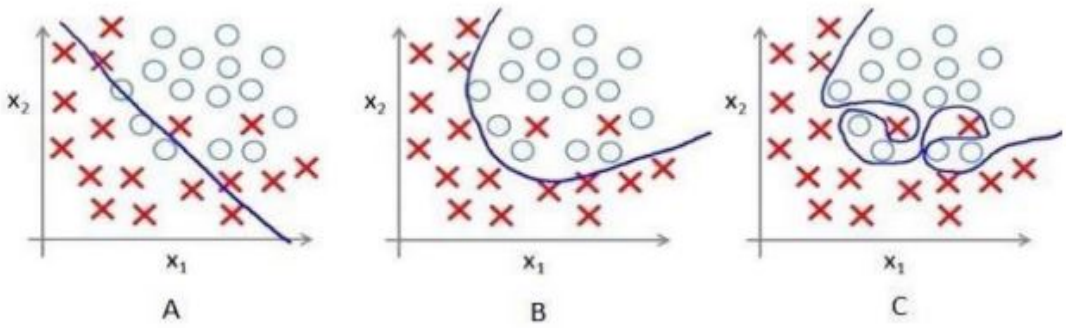
C) The model would not be affected by distance of points from hyperplane for modelling.

D) None of the above

No, the answer is incorrect.
Score: 0

Accepted Answers:
B) The model would consider only the points close to the hyperplane for modelling.

9) Below are the labelled instances of 2 classes and hand drawn decision boundaries for logistic regression. Which of the following figure demonstrates overfitting of the training data? ***2 points***



A) A

B) B

C) C

D) None of these

No, the answer is incorrect.
Score: 0

Accepted Answers:
C) C

10) What do you conclude after seeing the visualization in previous question? ***2 points***

C1. The training error in first plot is higher as compared to the second and third plot.
C2. The best model for this regression problem is the last (third) plot because it has minimum training error (zero).
C3. Out of the 3 models, the second model is expected to perform best on unseen data.
C4. All will perform similarly because we have not seen the test data.

A) C1 and C2

B) C1 and C3

C) C2 and C3

D) C4

No, the answer is incorrect.
Score: 0

Accepted Answers:
B) C1 and C3

Week 6 : Assignment 6

The due date for submitting this assignment has passed.

Due on 2021-09-08, 23:59 IST.

As per our records you have not submitted this assignment.

1) In training a neural network, we notice that the loss does not increase in the first few starting epochs: What is the reason for this? ***2 points***

- A) The learning Rate is low.
- B) Regularization Parameter is High.
- C) Stuck at the Local Minima.
- D) All of these could be the reason.

No, the answer is incorrect.
Score: 0
Accepted Answers:
D) All of these could be the reason.

2) What is the sequence of the following tasks in a perceptron? ***2 points***

I) Initialize the weights of the perceptron randomly.
II) Go to the next batch of data set.
III) If the prediction does not match the output, change the weights.
IV) For a sample input, compute an output.

- A) I, II, III, IV
- B) IV, III, II, I
- C) III, I, II, IV
- D) I, IV, III, II

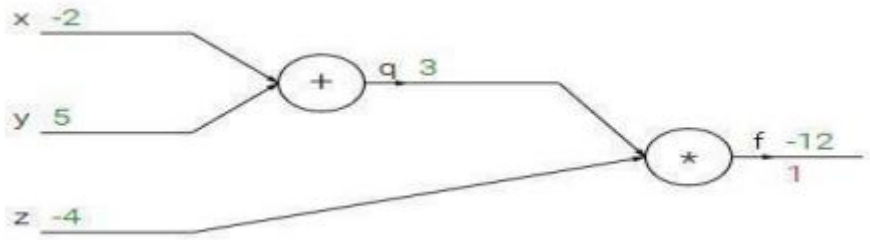
No, the answer is incorrect.
Score: 0
Accepted Answers:

D) I, IV, III, II

3) Suppose you have inputs as x, y, and z with values -2, 5, and -4 respectively. You have a neuron 'q' and neuron 'f' with functions: ***2 points***

q = x + y
f = q * z

Graphical representation of the functions is as follows:



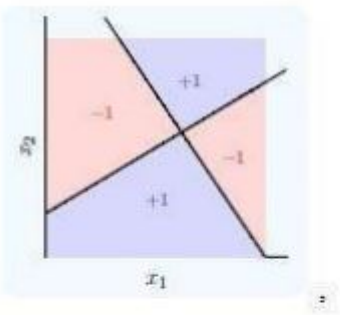
What is the gradient of F with respect to x, y, and z?

- A) (-3, 4, 4)
- B) (4, 4, 3)
- C) (-4, -4, 3)
- D) (3, -4, -4)

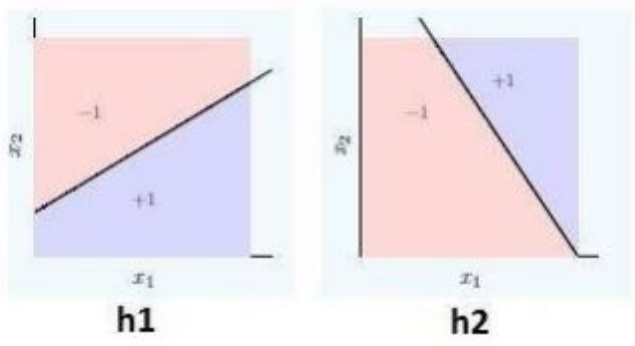
No, the answer is incorrect.
Score: 0

Accepted Answers:
C) (-4, -4, 3)

4) A neural network can be considered as multiple simple equations stacked together. Suppose we want to replicate the function for the below mentioned decision boundary. ***2 points***



Using two simple inputs h1 and h2,



What will be the final equation?

- A) (h1 AND NOT h2) OR (NOT h1 AND h2)
- B) (h1 OR NOT h2) AND (NOT h1 OR h2)
- C) (h1 AND h2) OR (h1 OR h2)
- D) None of these

No, the answer is incorrect.
Score: 0

Accepted Answers:
A) (h1 AND NOT h2) OR (NOT h1 AND h2)

5) Which of the following is true about model capacity (where model capacity means the ability of neural network to approximate complex functions)? ***2 points***

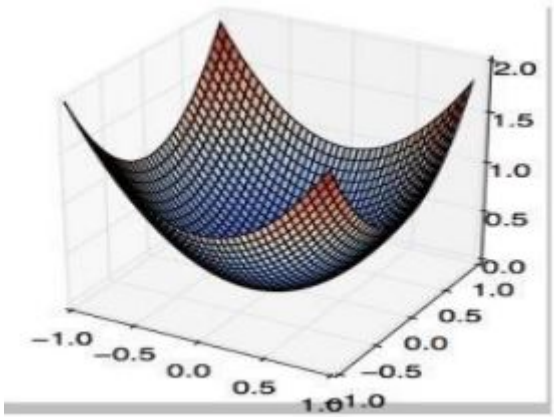
- A) As number of hidden layers increase, model capacity increases
- B) As dropout ratio increases, model capacity increases
- C) As learning rate increases, model capacity increases
- D) None of these

No, the answer is incorrect.
Score: 0

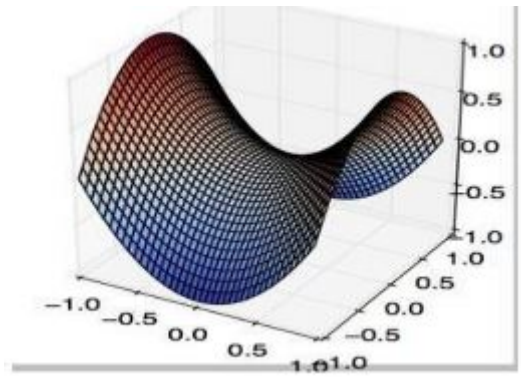
Accepted Answers:
A) As number of hidden layers increase, model capacity increases

6) First Order Gradient descent would not work correctly (i.e. may get stuck) in which of the following graphs? ***2 points***

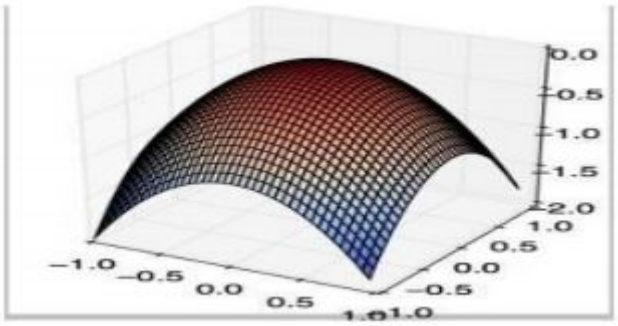
- A)



- B)



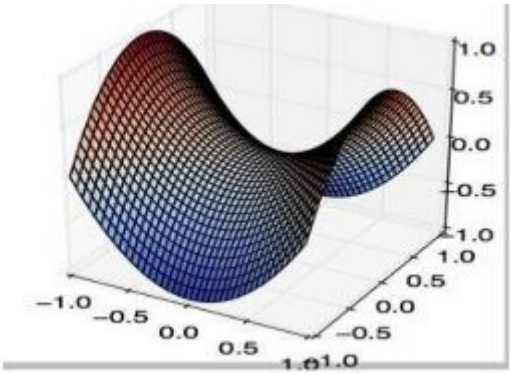
C)



D)None of These

No, the answer is incorrect.
Score: 0

Accepted Answers:
B)



7) Which of the following is true? ***2 points***

Single layer associative neural networks do not have the ability to

I) Perform pattern recognition
II) Find the parity of a picture
III) Determine whether two or more shapes in a picture are connected or not

A) II and III are true
B) II is true
C) All of the above
D) None of the above

No, the answer is incorrect.
Score: 0

Accepted Answers:
A) II and III are true

8) The network that involves backward links from outputs to the inputs and hidden layers is called as ***2 points***

A) Self-organizing Maps
B) Perceptron
C) Recurrent Neural Networks
D) Multi-Layered Perceptron

No, the answer is incorrect.
Score: 0

Accepted Answers:
C) Recurrent Neural Networks

Week 7 : Assignment 7

The due date for submitting this assignment has passed.

Due on 2021-09-15, 23:59 IST.

As per our records you have not submitted this assignment.

1) Which of the following option is / are correct regarding the benefits of ensemble model? ***2 points***

1. Better performance
2. More generalized model
3. Better interpretability

- A) 1 and 3
- B) 2 and 3
- C) 1 and 2
- D) 1, 2 and 3

No, the answer is incorrect.
Score: 0

Accepted Answers:
C) 1 and 2

2) In AdaBoost, we give more weights to points having been misclassified in previous iterations. Now, if we introduced a limit or cap on the weight that any point can take (for example, say we introduce a restriction that prevents any point's weight from exceeding a value of 10). Which among the following would be an effect of such a modification? ***2 points***

- A) We may observe the performance of the classifier reduce as the number of stages increase.
- B) It makes the final classifier robust to outliers.
- C) It may result in lower overall performance.
- D) None of these.

No, the answer is incorrect.
Score: 0

Accepted Answers:
B) It makes the final classifier robust to outliers.
C) It may result in lower overall performance.

3) Which among the following are some of the differences between bagging and boosting? ***2 points***

- [ ] A) In bagging we use the same classification algorithm for training on each sample of the data, whereas in boosting, we use different classification algorithms on the different training data samples.NPTEL Online Certification Courses Indian Institute of Technology Kharagpur
- [ ] B) Bagging is easy to parallelize whereas boosting is inherently a sequential process.
- [ ] C) In bagging we typically use sampling with replacement whereas in boosting, we typically use weighted sampling techniques.
- [ ] D) In comparison with the performance of a base classifier on a particular data set, bagging will generally not increase the error whereas as boosting may lead to an increase in the error.

No, the answer is incorrect.
Score: 0

Accepted Answers:
B) Bagging is easy to parallelize whereas boosting is inherently a sequential process.
C) In bagging we typically use sampling with replacement whereas in boosting, we typically use weighted sampling techniques.
D) In comparison with the performance of a base classifier on a particular data set, bagging will generally not increase the error whereas as boosting may lead to an increase in the error.

4) What is the VC-dimension of the class of circle in a 4-dimensional plane? ***2 points***

- ( ) A) 3
- ( ) B) 4
- ( ) C) 5
- ( ) D) 6

No, the answer is incorrect.
Score: 0

Accepted Answers:
C) 5

5) Considering the AdaBoost algorithm, which among the following statements is true? ***2 points***

- [ ] A) In each stage, we try to train a classifier which makes accurate predictions on any subset of the data points where the subset size is at least half the size of the data set.
- [ ] B) In each stage, we try to train a classifier which makes accurate predictions on a subset of the data points where the subset contains more of the data points which were misclassified in earlier stages.
- [ ] C) The weight assigned to an individual classifier depends upon the number of data points correctly classified by the classifier.
- [ ] D) The weight assigned to an individual classifier depends upon the weighted sum error of misclassified points for that classifier.

No, the answer is incorrect.
Score: 0

Accepted Answers:
B) In each stage, we try to train a classifier which makes accurate predictions on a subset of the data points where the subset contains more of the data points which were misclassified in earlier stages.
D) The weight assigned to an individual classifier depends upon the weighted sum error of misclassified points for that classifier.

6) Suppose the VC dimension of a hypothesis space is 6. Which of the following are true? ***0 points***

- [ ] A) At least one set of 6 points can be shattered by the hypothesis space.
- [ ] B) No sets of 6 points can be shattered by the hypothesis space.
- [ ] C) All sets of 6 points can be shattered by the hypothesis space.
- [ ] D) No set of 6 points can be shattered by the hypothesis space.

No, the answer is incorrect.
Score: 0

Accepted Answers:
A) At least one set of 6 points can be shattered by the hypothesis space.
D) No set of 6 points can be shattered by the hypothesis space.

7) Ensembles will yield bad results when there is a significant diversity among the models. Write True or False. ***2 points***

- ( ) A) True
- ( ) B) False

No, the answer is incorrect.
Score: 0

Accepted Answers:
B) False

8) Which of the following algorithms are not an ensemble learning algorithm? ***2 points***

- ( ) A) Random Forest
- ( ) B) Adaboost
- ( ) C) Gradient Boosting
- ( ) D) Decision Tress

No, the answer is incorrect.
Score: 0

Accepted Answers:
D) Decision Tress

Week 8 : Assignment 8

The due date for submitting this assignment has passed.

Due on 2021-09-22, 23:59 IST.

As per our records you have not submitted this assignment.

1) For two runs of K-Mean clustering is it expected to get same clustering results? ***2 points***

- A) Yes
- B) No

No, the answer is incorrect.
Score: 0

Accepted Answers:
B) No

2) Which of the following can act as possible termination conditions in K-Means? ***2 points***

I. For a fixed number of iterations.
II. Assignment of observations to clusters does not change between iterations. Except for cases with a bad local minimum.
III. Centroids do not change between successive iterations.
IV. Terminate when RSS falls below a threshold

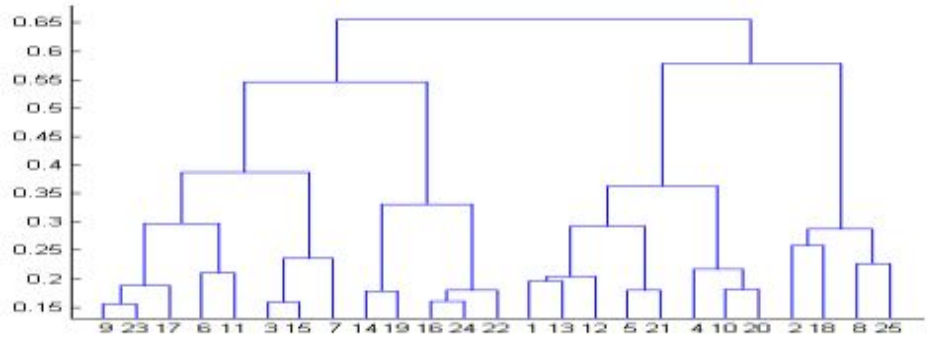
- A) I, III and IV
- B) I, II and III
- C) I, II and IV
- D) All of the above

No, the answer is incorrect.
Score: 0

Accepted Answers:
D) All of the above

Lecture 42: Kmeans Clustering (unit? unit=68&lesson=70)
Lecture 43: Agglomerative Hierarchical Clustering (unit? unit=68&lesson=71)
Lecture 44: Python Exereise on Kmeans Clustering (unit? unit=68&lesson=72)
Tutorial 8 (unit? unit=68&lesson=73)
Week 8 - Lecture Notes (unit? unit=68&lesson=74)
Quiz: Week 8 : Assignment 8 (assessment? name=110)
Feedback For Week 8 (unit? unit=68&lesson=75)
Assignment Solution
Download Videos

3) After performing K-Means Clustering analysis on a dataset, you observed the following dendrogram. Which of the following conclusion can be drawn from the dendrogram? ***2 points***

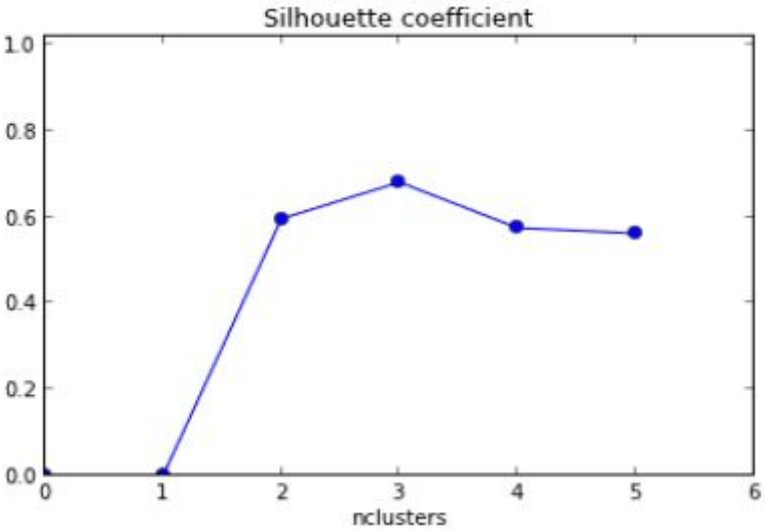


A) There were 28 data points in clustering analysis.
B) The best no. of clusters for the analysed data points is 4.
C) The proximity function used is Average-link clustering.
D) The above dendrogram interpretation is not possible for K-Means clustering analysis.

No, the answer is incorrect.
Score: 0

Accepted Answers:
D) The above dendrogram interpretation is not possible for K-Means clustering analysis.

4) What should be the best choice of no. of clusters based on the following results: ***2 points***



A) 1
B) 2
C) 3
D) 4

No, the answer is incorrect.
Score: 0

Accepted Answers:
C) 3

5) Given, six points with the following attributes: ***2 points***

point	x coordinate	y coordinate
p1	0.4005	0.5306
p2	0.2148	0.3854
p3	0.3457	0.3156
p4	0.2652	0.1875
p5	0.0789	0.4139
p6	0.4548	0.3022

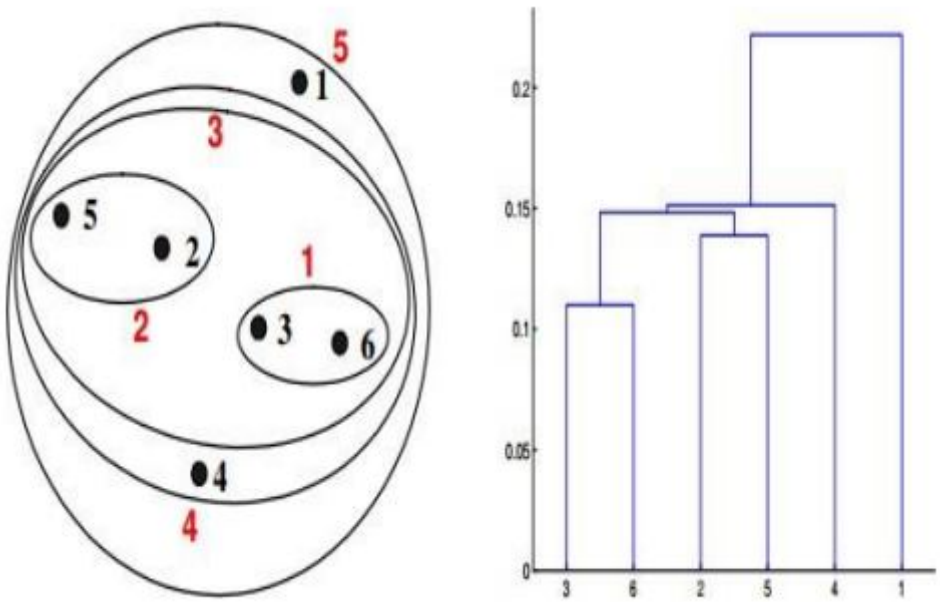
Table : X-Y coordinates of six points.

	p1	p2	p3	p4	p5	p6
p1	0.0000	0.2357	0.2218	0.3688	0.3421	0.2347
p2	0.2357	0.0000	0.1483	0.2042	0.1388	0.2540
p3	0.2218	0.1483	0.0000	0.1513	0.2843	0.1100
p4	0.3688	0.2042	0.1513	0.0000	0.2932	0.2216
p5	0.3421	0.1388	0.2843	0.2932	0.0000	0.3921
p6	0.2347	0.2540	0.1100	0.2216	0.3921	0.0000

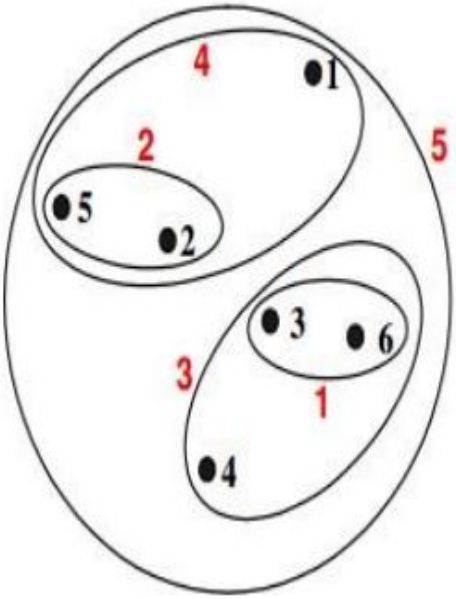
Table : Distance Matrix for Six Points

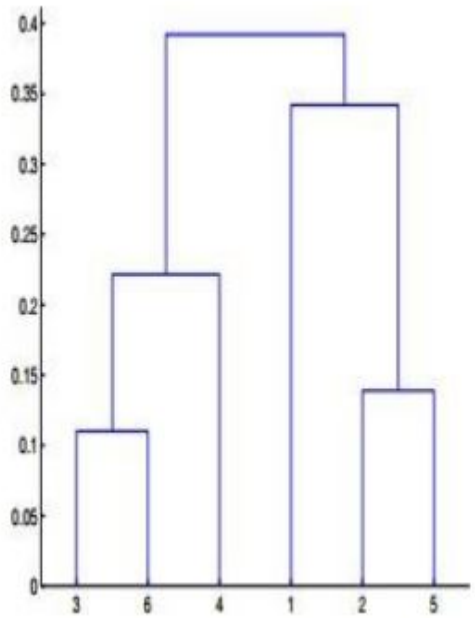
Which of the following clustering representations and dendrogram depicts the use of MIN or Single link proximity function in hierarchical clustering:

A)

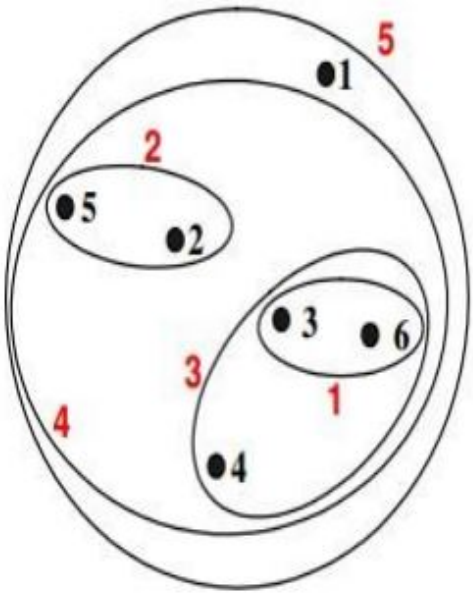


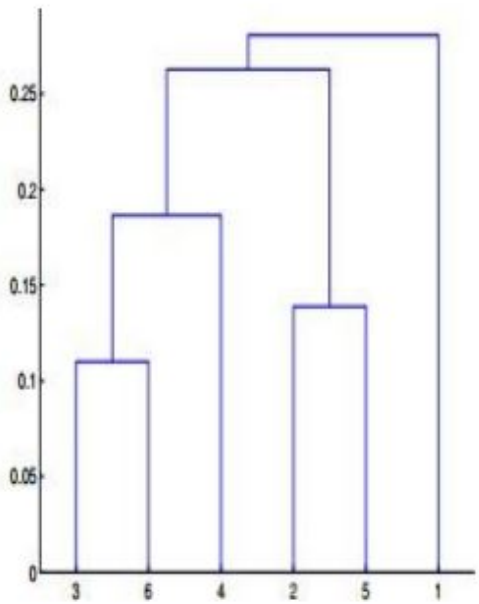
B)


4
1
2
5
5
2
3
6
3
1
4

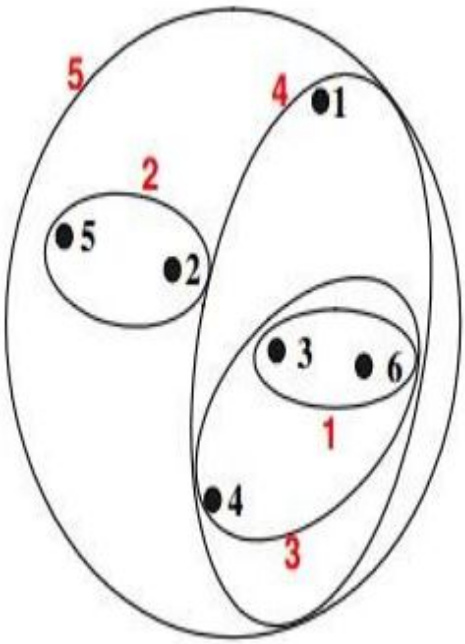

0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
3
6
4
1
2
5

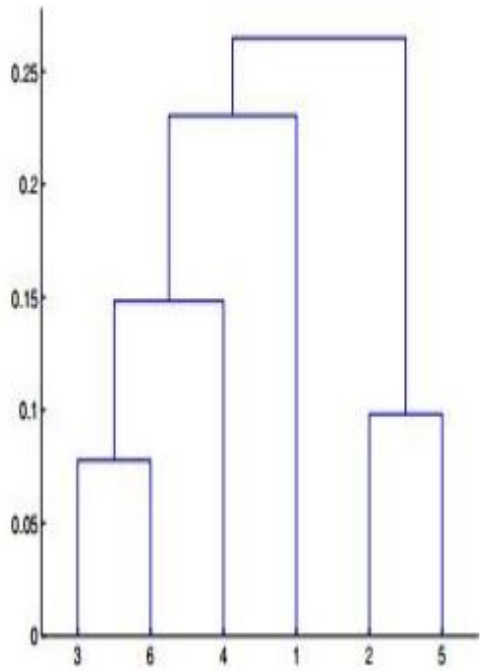
C)


5
1
2
5
2
3
6
3
1
4
4


0.25
0.2
0.15
0.1
0.05
0
3
6
4
2
5
1

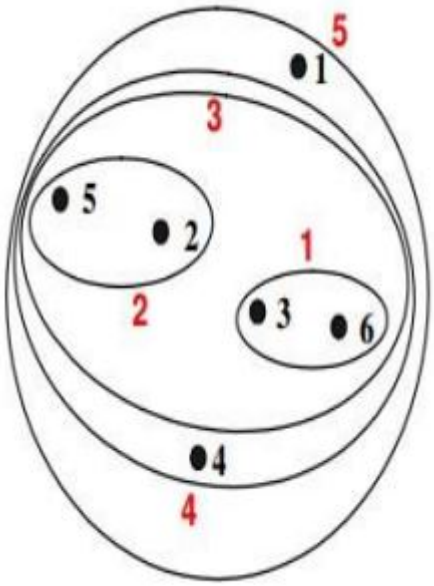
D)

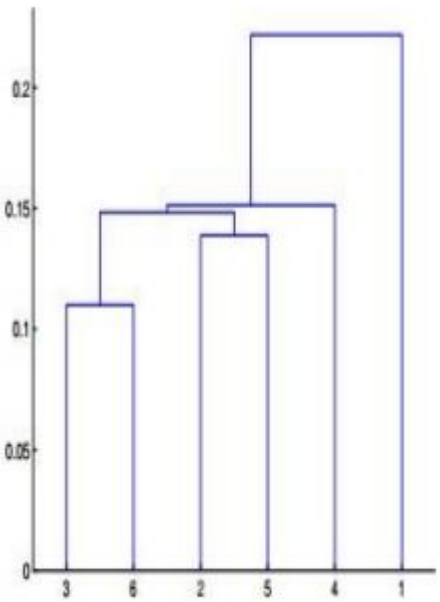




No, the answer is incorrect.
Score: 0

Accepted Answers:
A)





6) Which of the following algorithms are most sensitive to outliersWhatWh? ***2 points***

A) K-means clustering

B) K-medians clustering

C) K-modes clustering

D) K -medoids clustering

No, the answer is incorrect.
Score: 0

Accepted Answers:
A) K-means clustering

7) What is the possible reason(s) for producing two different dendograms using agglomerative clustering for the same data set? ***2 points***

- A) Proximity function
- B) No. of data points
- C) Variables used
- D) All of these

No, the answer is incorrect.
Score: 0

Accepted Answers:
D) All of these