27 MCQs

**Weight 2 Mark**

1) Which of the following option is true about k-NN algorithm?

A) It can be used for classification  
B) It can be used for regression  
**C) It can be used in both classification and regression**

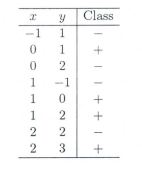
**Weight 1 Mark**

2) Which of the following machine learning algorithm can be used for imputing missing values of both categorical and continuous variables?

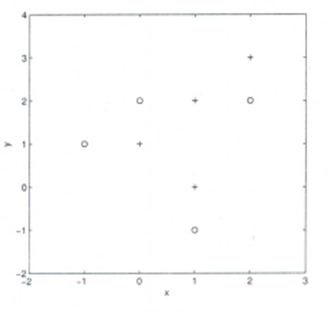
**A) K-NN**  
B) Linear Regression  
C) Logistic Regression

**Context: for question number 3 and 4**

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

[](https://cdn.analyticsvidhya.com/wp-content/uploads/2017/08/01184933/Pic_C_11_12.jpg)

Below is a scatter plot which shows the above data in 2D space.

[](https://cdn.analyticsvidhya.com/wp-content/uploads/2017/08/01185059/Pic_C_11_12_b.JPG_.jpg)

**Weight 2 Mark**

3) Suppose, you want to predict the class of new data point x=1 and y=1 using eucludian distance in 3-NN. In which class this data point belong to?

A) **+ Class**

B) – Class

C) Can’t say

D) None of these

Solution: A

All three nearest point are of +class so this point will be classified as +class.

**Weight 3 Mark**

4) In the previous question, you are now want use 7-NN instead of 3-KNN which of the following x=1 and y=1 will belong to?

A) + Class

B) **– Class**

C) Can’t say

Solution: B

Now this point will be classified as – class because there are 4 – class and 3 +class point are in nearest circle.

**Weight 1 Mark**

5) Movie Recommendation systems are an example of:

a)Classification

**b)Clustering**

c)Reinforcement Learning

d)Regression

**Weight 1 Mark**

6). Is it possible that Assignment of observations to clusters does not change between successive iterations in K-Means

**A. Yes**

B. No

C. Can’t say

D. None of these

**Weight 3 Mark**

7)Three clusters in KMeans C1, C2, C3 has following points:

C1: {(2,2), (4,4), (6,6)}

C2: {(0,4), (4,0)}

C3: {(5,5), (9,9)}

What will be the clusters (centroids)

A. **C1: (4,4), C2: (2,2), C3: (7,7)**

B. C1: (6,6), C2: (4,4), C3: (9,9)

C. C1: (2,2), C2: (0,0), C3: (5,5)

D. None of these

Solution: (A)

Finding centroid for data points in cluster C1 = ((2+4+6)/3, (2+4+6)/3) = (4, 4)

**Weight 1 Mark**

Q8). Which of the following method is used for finding optimal of cluster in K-Mean algorithm?

A. **Elbow method**

B. Manhattan method

C. Ecludian mehthod

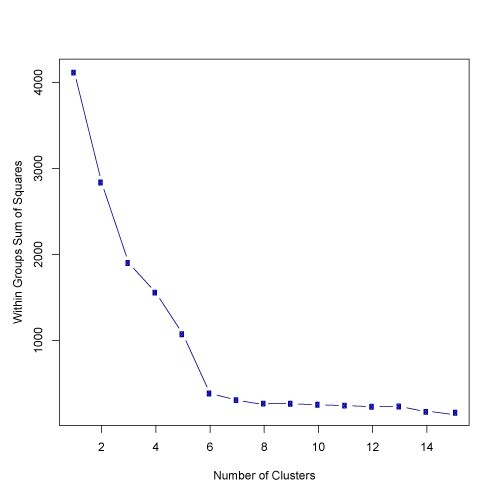
D. All of the above

E. None of these

Solution: (A)

**Weight 1 Mark**

Q9). What should be the best choice for number of clusters based on the following results:



A. 5

**B. 6**

C. 14

D. Greater than 14

Solution: (B)

**Weight 1 Mark**

10) A \_\_\_\_\_\_\_\_\_ is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility.

a) **Decision tree**

b) Graphs

c) Trees

d) Neural Networks

**Weight 1 Mark**

11) Decision Nodes are represented by \_\_\_\_ \_\_\_\_\_\_\_\_

a) Disks

b) Squares

c**) Circles**

d) Triangles

**Weight 1 Mark**

**12) What is back propagation?**

**a) It is another name given to the curvy function in the perceptron**

**b) It is the transmission of error back through the network to adjust the inputs**

**c) It is the transmission of error back through the network to allow weights to be adjusted so that the network can learn**

**d) None of the mentioned**

**Weight 1 Mark**

**13)A perceptron adds up all the weighted inputs it receives, and if it exceeds a certain value, it outputs a 1, otherwise it just outputs a 0.**

**a) True**

**b) False**

**c) Sometimes – it can also output intermediate values as well**

**d) Can’t say**

**Weight 2 Mark**

14) The amount of output of one unit received by another unit depends on what?

a) output unit

b) input unit

c) activation value

**d) weight**

Answer: d

Explanation: Activation is sum of wieghted sum of inputs, which gives desired output..hence output depends on weights.

**Weight 1 Mark**

15). The process of adjusting the weight is known as?

a) activation

b) synchronisation

c**) learning**

d) none of the mentioned

Answer: c

Explanation: Basic definition of learning in neural nets .

**Weight 1 Mark**

16). The procedure to incrementally update each of weights in neural is referred to as?

a) synchronisation

b) learning law

c) learning algorithm

d) **both learning algorithm & law**

Answer: d

Explanation: Basic definition of learning law in neural.

**Weight 1 Mark**

17). In what ways can output be determined from activation value?

a) deterministically

b) stochastically

c) **both deterministically & stochastically**

d) none of the mentioned

Answer: c

Explanation: This is the most important trait of input processing & output determination in neural networks.

In regression analysis, the variable that is being predicted is;

a) the independent variable

b) the dependent variable

c) usually denoted by x

d) usually denoted by r

**Weight 1 Mark**

18.) In regression analysis, the variable that is being predicted is;

a) the independent variable

**b) the dependent variable**

c) usually denoted by x

d) usually denoted by r

**Weight 1 Mark**

19)

) In the regression equation y = b

o

+ b

1

x, b

o

is the;

a) slope of the line

b) independent variable

c) y intercept

d) coefficient of determination

In the regression equation y = bo + b1x, bo is the;

a) slope of the line

b) independent variable

**c) y intercept**

d) coefficient of determination

**Weight 1 Mark**

20) If the slope of the regression equation y = bo + b1x is positive, then;

a) as x increases y decreases

b**) as x increases so does y**

c) Either a or b is correct

d) as x decreases y increases

**Weight 1 Mark**

21) Least square method calculates the best-fitting line for the observed data by minimizing the sum of the squares of the \_\_\_\_\_\_\_ deviations.

a) **Vertical**

b) Horizontal

c) Both of these

d) None of these

**Weight 1.5 Mark**

22) Which one is the least square method formula;

a) min ∑(yi – yi^) 2

b) min ∑( yi^ -yi)

**c) min ∑(yi - yi^ ) 2**

d) min ∑(yi - yi^ )

**Weight 2 Mark**

23) If the regression equation is equal to y=23.6−54.2x, then 23.6 is the \_\_\_\_\_ while -54.2 is the \_\_\_\_ of the regression line.

a) Slope, intercept

b) Slope, regression coefficient

c) **Intercept, slope**

d) Radius, intercept

**Weight 1 Mark**

24)The confusion matrix is used to: (Multiple answers possible)

a)Understand how attributes are related to each other

b)How data is spread on each dimension

**c)Evaluate performance of classification algorithms within each class**

**d)It is a visual representation of the actual distribution of predicted values of target labels in context of the actual values of the target labels**

**Weight 2.5 Mark**

**25)For the below confusion matrix, what is the accuracy?**

**Predicted**

**No Yes**

**Groud Truth No 53272 1307**

**Yes 1077 4344**

**a)95%**

**b)90%**

**c)96%**

**d)98%**

**Weight 2.5 Mark**

**26)** **A model makes predictions and predicts 70 examples for the first minority class, where 50 are correct and 20 are incorrect. It predicts 150 for the second class with 99 correct and 51 incorrect. Precision will be**

**a) 96.6**

**b)67.7**

**c)77.2**

**d)88.4**

**Weight 2.5 Mark**

* **27).** Recall = TruePositives / (TruePositives + FalseNegatives)

A model makes predictions and predicts 90 of the positive class predictions correctly and 10 incorrectly. Recall will be

* 1. 67
  2. 60
  3. **90**
  4. 80