These are sample MCQs to indicate pattern, may or may not appear in examination

University of Mumbai
Examination 2020

Program: BE Computer Engineering
Curriculum Scheme : Revised 2016
Examination: Third Year Semester : VI
Course Code: __CSDLO6021___and Course Name: ____Machine Learning

Time: 1hour	Max. Marks: 50
Note to the students:- All the Questions are compulsory and carry equal marks .	

Q1.	What type of problem is robot navigation is?
Option A:	Regression
Option B:	Classification
Option C:	Clustering
Option D:	Reinforcement Learning

Q2.	What is mean by linear regression?
Option A:	Modeling the relationship between independent variable with one or more dependent variables
Option B:	Modeling the relationship between dependent variable with one or more independent variables
Option C:	Identifying to which of a set of categories a new observation belongs
Option D:	system is used to store and manipulate knowledge to interpret information in a useful way.

Q3.	Given a large dataset of medical records from patients suffering from heart disease, try to learn whether there might be different clusters of such patients for which we might tailor separate treatments.?What type problem is?
Option A:	Supervised Learning problem
Option B:	Unsupervised Learning problem
Option C:	Reinforcement Learning problem
Option D:	Classification Problem

Q4.	What is function of dendrite in human brain?
Option A:	receives signals from other neurons
Option B:	point of connection to other neurons

Option C:	processes the information
Option D:	transmits the output of this neuron

Q5.	What is function of synapse in human brain?
Option A:	receives signals from other neurons
Option B:	point of connection to other neurons
Option C:	processes the information
Option D:	transmits the output of this neuron

Q6.	What is use of activation function in artificial neural network?
Option A:	Activation function decides, whether a neuron should be activated or not by calculating weighted sum and further adding bias with it.
Option B:	It is used to add weightage to input layer
Option C:	It is used to calculate cost fuction
Option D:	It is used in hidden layer

Q7.	What is perceptron?
Option A:	a single layer feed-forward neural network with pre-processing
Option B:	an auto-associative neural network
Option C:	a double layer auto-associative neural network
Option D:	a neural network that contains feedback

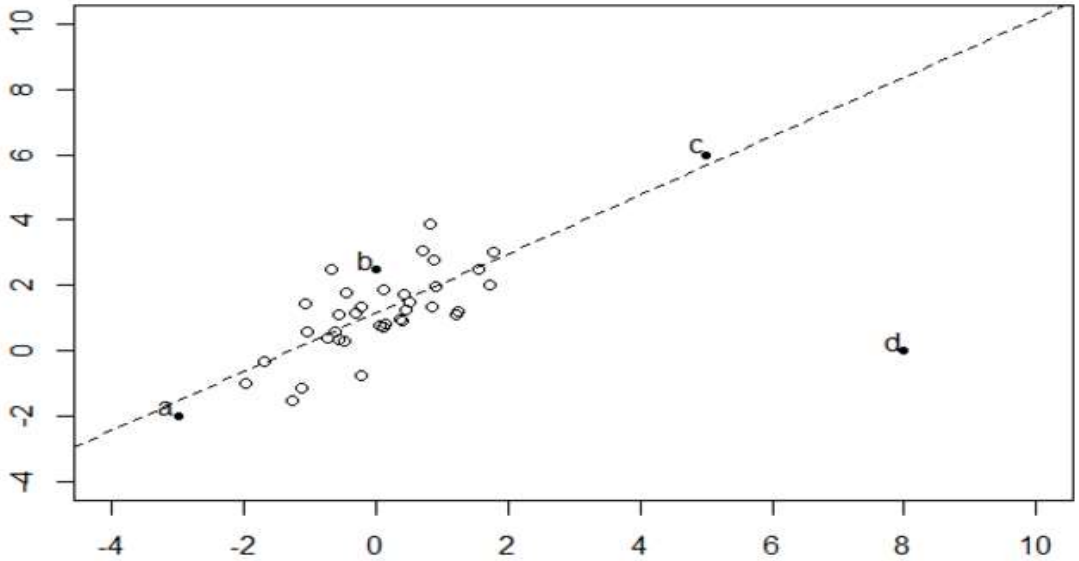
Q8.	Newton Method is of which type of optimization techniques?
Option A:	Derivative Free
Option B:	Derivative Based
Option C:	Both Derivative based and Derivative free
Option D:	Neither Derivative based nor derivative free

Q9.	Random search is of which type of optimization techniques?
Option A:	Derivative Free
Option B:	Derivative Based
Option C:	Both Derivative based and Derivative free
Option D:	Neither Derivative based nor derivative free

Q10.	How Down hill simplex works?
Option A:	It is a method to find the minimum of a function in more than one independent variable.
Option B:	It is a method to find the minimum of a function in more than one independent variable.
Option C:	It is a method to find the minimum and maximum of a function in more than one independent variable.
Option D:	It Computes gradient using the whole Training sample

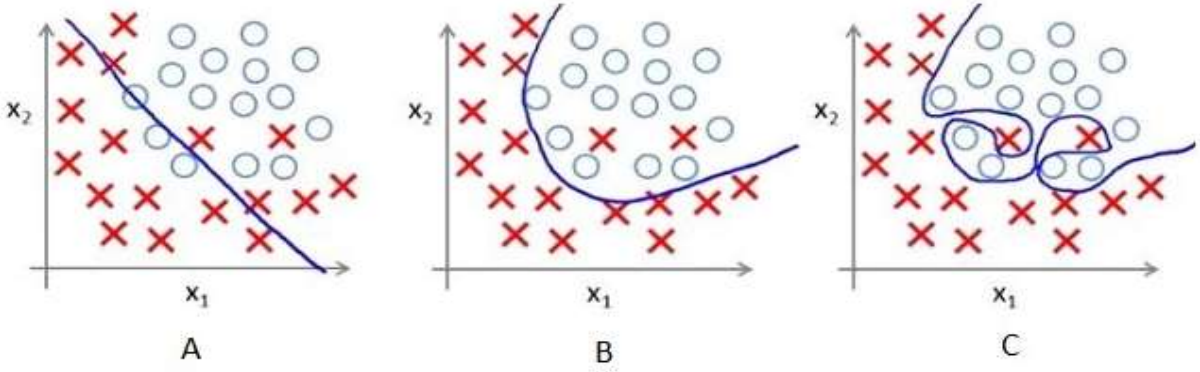
Q11.	How steepest descent optimization work?
Option A:	It's based on a convex function and tweaks its parameters iteratively to minimize a given function to its local minimum.
Option B:	It's based on a concave function and tweaks its parameters iteratively to minimize a given function to its local minimum.
Option C:	It's based on a convex function and tweaks its parameters iteratively to maximize a given function to its local minimum.
Option D:	It's based on a concave function and tweaks its parameters iteratively to maximize a given function to its local minimum.

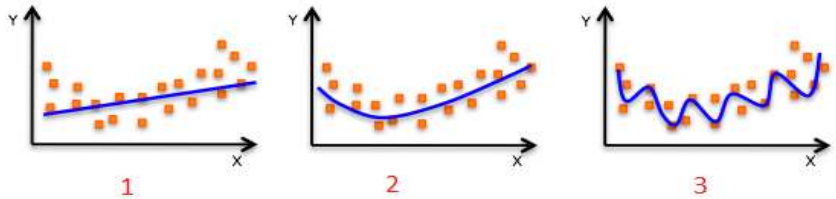
Q12.	Which of the following methods do we use to find the best fit line for data in Linear Regression?
Option A:	Least Square Error
Option B:	Maximum Likelihood
Option C:	Logarithmic Loss
Option D:	Both A and B

Q13.	**Consider the following dataset. Which bold point, if removed will have the largest effect on fitted regression line as shown in above figure(dashed)?** 
Option A:	a
Option B:	b
Option C:	c
Option D:	d

Q14.	**How does number of observations influence overfitting? Choose the correct answer(s). Note: Rest all parameters are same.** **1. In case of fewer observations, it is easy to overfit the data.** **2. In case of fewer observations, it is hard to overfit the data.** **3. In case of more observations, it is easy to overfit the data.** **4. In case of more observations, it is hard to overfit the data.**
Option A:	1 and 4

Option B:	2 and 3
Option C:	1 and 3
Option D:	All of the above

Q15.	Below are the three scatter plot(A,B,C left to right) and hand drawn decision boundaries for logistic regression. Suppose, above decision boundaries were generated for the different value of regularization. Which of the above decision boundary shows the maximum regularization? 
Option A:	A
Option B:	B
Option C:	C
Option D:	All of the above

Q16.	**The following visualization shows the fit of three different models (in blue line) on same training data. What can you conclude from these visualizations?**  1. The training error in first model is higher when compared to second and third model. 2. The best model for this regression problem is the last (third) model, because it has minimum training error. 3. The second model is more robust than first and third because it will perform better on unseen data. 4. The third model is overfitting data as compared to first and second model. 5. All models will perform same because we have not seen the test data.
Option A:	1 and 2
Option B:	1 and 3
Option C:	1, 3 and 4
Option D:	only 5

Q17.	Which one will be selected as root node of the decision tree for the following data given in table. What is the information gain of the root node.

Training Examples

Day	Outlook	Temp	Humidity	Wind	Tennis?
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Option A:	Outlook with information gain = 0.247
Option B:	Outlook with information gain = 0.563
Option C:	Humidity with information gain = 0.247
Option D:	Humidity with information gain = 0.563

Q18.	**The most popularly used dimensionality reduction algorithm is Principal Component Analysis (PCA). Which of the following is/are true about PCA?** **1. PCA is an unsupervised method** **2. It searches for 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**
Option A:	1 and 2
Option B:	2 and 3
Option C:	3 and 4
Option D:	All of the above

Q19.	**Which of the following is true about LDA?** **A. LDA aims to maximize the distance between class and minimize the within class distance** **B. LDA aims to minimize both distance between class and distance within class** **C. LDA aims to minimize the distance between class and maximize the distance within class** **D. LDA aims to maximize both distance between class and distance within class**
Option A:	A
Option B:	B
Option C:	C
Option D:	D

Q20.	Find SVD for following matrix, i.e U , Σ and V $$A = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$
Option A:	$\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 16 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$
Option B:	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
Option C:	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

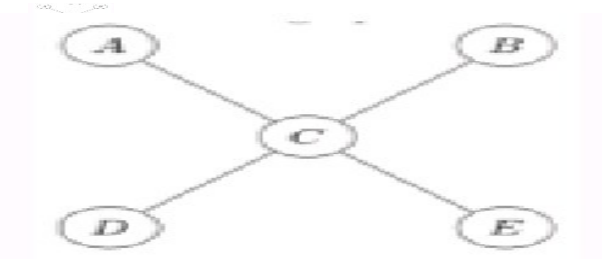
Option D:	None of these

Q21.	Find the Eigen Values of $\begin{bmatrix} 7 & 3 \\ 3 & -1 \end{bmatrix}$
Option A:	10 and 2
Option B:	8 and -2
Option C:	8 and 2
Option D:	None of these

Q22.	Instead of representing knowledge in a relatively declarative, static way (as a bunch of things that are true), rule-based system represent knowledge in terms of_____ that tell you what you should do or what you could conclude in different situations.
Option A:	Raw Text
Option B:	A bunch of rules
Option C:	Summarized Text
Option D:	Collection of various Texts

Q23.	Rule based classification algorithms generate _____ rule to perform the classification.
Option A:	if-then
Option B:	while.
Option C:	do while.
Option D:	Switch

Q24.	Which of the following property is false in case of Bayesain network ?
Option A:	Edges are directed
Option B:	Contains cycles
Option C:	Represents conditional independence relations among random variables
Option D:	All of the above

Q25.	 Consider the Markov network shown above. Two students Student A and Student B make the following claims -- Student A claims P(A|B)=P(A) Student B claims P(A|B,C)=P(A|C)
Option A:	Student A and Student B are correct
Option B:	Student A is Correct and Student B is incorrect
Option C:	Student A is incorrect and Student B is correct
Option D:	Both are incorrect