University of Mumbai

Examination 2021 under cluster KJSIEIT

Examinations Commencing from 22nd April 2021 to 30th April 2021

Program: Computer Engineering

Curriculum Scheme: Rev2019

Examination: MEAI Semester I

Course Code: MEAIC102 and Course Name: Machine Learning and Pattern Recognition

Time: 2 hour Max. Marks: 80

==========

Q1.	**Choose the correct option for following questions. All the Questions are compulsory and carry equal marks**
1.	____is a widely used and effective machine learning algorithm based on the idea of bagging
Option A:	Decision Tree
Option B:	Regression
Option C:	Classification
Option D:	Random Forest
2.	To find the minimum or the maximum of a function, we set the gradient to zero because
Option A:	The value of the gradient at extrema of a function is always zero
Option B:	Depends on the type of problem
Option C:	Both A &B
Option D:	None of the above
3.	Recalled output in pattern association problem depends on_____
Option A:	Nature of the input-output
Option B:	Design of network
Option C:	Both
Option D:	None of the above
4.	Machine learning is _____
Option A:	The autonomous acquisition of knowledge through the use of manual programs
Option B:	The selective acquisition of knowledge through the use of manual programs
Option C:	The autonomous acquisition of knowledge through the use of computer programs
Option D:	The selective acquisition of knowledge through the use of computer programs
5.	Which of the following is the correct way to preprocess the data?When performing regression or classification
Option A:	Normalize the data -> PCA -> normalize PCA output -> training
Option B:	PCA -> normalize PCA output -> training
Option C:	Normalize the data -> PCA -> training
Option D:	None
6.	The most widely used metrics and tools to assess a classification model are:

Option A:	Confusion matrix
Option B:	Cost-sensitive accuracy
Option C:	Area under the ROC curve
Option D:	All of the above
7.	Which of the following is true about Naive Bayes?
Option A:	Assumes that all the features in a dataset are independent
Option B:	Assumes that all the features in a dataset are equally important
Option C:	Both A and B
Option D:	None of the above
8.	In which of the following cases will K-means clustering fail to give good results? 1) Data points with outliers 2) Data points with different densities 3) Data points with nonconvex shapes
Option A:	1 and 2
Option B:	2 and 3
Option C:	1, 2, and 3
Option D:	1 and 3
9.	What is Reinforcement learning?
Option A:	All data is unlabelled and the algorithms learn to inherent structure from the input data
Option B:	Some data is llabelled but most of it is unlabelled and a mixture of supervised and unsupervised techniques can be used.
Option C:	All data is labeled and the algorithms learn to predict the output from the input data
Option D:	It is a framework for learning where an agent interacts with an environment and receives a reward for each interaction
10.	Supervised learning and unsupervised clustering both require at least one
Option A:	hidden attribute
Option B:	output attribute
Option C:	input attribute
Option D:	categorical attribute
11.	A regression model in which more than one independent variable is used to predict the dependent variable is called
Option A:	a simple linear regression model
Option B:	a multiple regression models
Option C:	an independent model
Option D:	None of the above
12.	A term used to describe the case when the independent variables in a multiple regression model are correlated is

Option A:	Regression
Option B:	correlation
Option C:	multicollinearity
Option D:	None of the above
13.	A multiple regression model has
Option A:	only one independent variable
Option B:	more than one dependent variable
Option C:	more than one independent variable
Option D:	None of the above
14.	Which of the following function is used for k-means clustering?
Option A:	k-means
Option B:	k-mean
Option C:	Heat map
Option D:	none of the mentioned
15.	Which of the following is the most appropriate strategy for data cleaning before performing clustering analysis, given less than desirable number of data points: Capping and flouring of variables Removal of outliers
Option A:	1 only
Option B:	2 only
Option C:	both
Option D:	None of the above
16.	Suppose you are dealing with 4 class classification problem and you want to train a SVM model on the data for that you are using One-vs-all method. How many times we need to train our SVM model in such case?
Option A:	1
Option B:	2
Option C:	3
Option D:	4
17.	Suppose you have trained an SVM with linear decision boundary after training SVM, you correctly infer that your SVM model is under fitting. Which of the following option would you more likely to consider iterating SVM next time?
Option A:	You want to increase your data points
Option B:	You want to decrease your data points
Option C:	You will try to calculate more variables
Option D:	You will try to reduce the features
18.	Which of the following statement is true about outliers in Linear regression?
Option A:	Linear regression is sensitive to outliers
Option B:	Linear regression is not sensitive to outliers
Option C:	Can't say
Option D:	None of the above
19.	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
20.	Machine learning techniques differ from statistical techniques in that machine learning methods
Option A:	typically assume an underlying distribution for the data.
Option B:	are better able to deal with missing and noisy data
Option C:	are not able to explain their behavior.
Option D:	have trouble with large-sized datasets.

Qu-2	**Solve any Four out of Six (5 marks each)**
A	Describe Linear regression with an example.
B	Enumerate the steps in the k-NN clustering algorithm
C	Write a short note on the following : a.)Pooling b.) Gradient Descent
D	Difference between supervised learning and Unsupervised learning
E	Explain the term Overfitting and Underfitting
F	Explain support vector machine algorithm

Qu-3.	**Solve any one Questions out of Two**
A	**Solve any Two 5 marks each**
i.	What are the important objectives of machine learning?
ii.	What are the basic design issues and approaches to machine learning?
B	**Solve any One** **10 marks each**
i.	A dealer has a warehouse that stores a variety of fruits and vegetables. When fruit is brought to the warehouse, various types of fruit may be mixed together. The dealer wants a model that will sort the fruit according to type. Justify with reasons how machine learning model is efficient compared to feature based classification technique.
ii.	Briefly explain unsupervised machine learning algorithm. Explain with the help of a suitable example where you will choose unsupervised learning over supervised learning?

University of Mumbai

Examination 2021 under cluster KJSIEIT

Examinations Commencing from 22nd April 2021 to 30th April 2021

Program: Computer Engineering

Curriculum Scheme: Rev2019

Examination: MEAI Semester I

Course Code: MEAIC102 and Course Name: Machine Learning and Pattern Recognition

Time: 2 hour Max. Marks: 80

==========

Question Number	Correct Option (Enter either 'A' or 'B' or 'C' or 'D')
Q1.	D
Q2.	A
Q3.	C
Q4	C
Q5	A
Q6	D
Q7	D
Q8.	C
Q9.	D
Q10.	D
Q11.	C
Q12.	C
Q13.	B
Q14.	A
Q15.	A
Q16.	D
Q17.	C
Q18.	C
Q19.	A
Q20.	D