TKM COLLEGE OF ENGINEERING, KOLLAM-5

Department of Computer Applications

III Semester MCA

Internal Assessment Retest (Offline) Mar 2022

Course with Code: 20MCA 201 DATA SCIENCE &MACHINE LEARNING

Time:2Hrs **Maximum Marks: 50**

Qn. No.	PART – A Answer all questions	Marks	BL	CO
1	Explain the need of data science.	3	L1	1
2	List the measures of central tendency for numeric as well as categorical data and explain the information about the dataset conveyed by it.	3	L2	1
3	Explain about Ensemble Modeling..	3	L1	1
4	Differentiate between supervised and unsupervised learning algorithm with example for each.	3	L2	2
5	Explain how to choose the value of k in k-NN algorithm.	3	L2	2
6	Explain why Laplace estimator is needed in classification using Bayes theorem.	3	L2	2
7	State Bayes' theorem in statistics.	3	L1	2
8	Write a note on Ordinary Least Square method in regression.	3	L1	3
9	Differentiate between Regression and Correlation.	3	L2	3
10	Differentiate between entropy and information gain.	3	L2	3
	PART – B			
	MODULE-I			
11 a	Explain various methods for visualizing univariate data.	5	L2	1
	OR			
b	Explain the data science classification and illustrate data science tasks.	5	L2	1
	MODULE II			
12 a	Based on the survey conducted in an institution, the students are classified based on their academic excellence, extracurricular and Co-curricular activities.	5	L3	2

Consider the data set given.

A (Academic Excellence)	B(Extracurricular achievement)	C (Co-curricular achievement)	Overall Performance
8	6	7	Excellent
5	6	4	Good
7	3	4	Good
6	9	8	Excellent

Find the category of student with A=9, B=5and C=5based on the data of trained samples using KNN algorithm.

OR

Find the probability to play golf on 15th day where conditions are, temperature=cool, humidity=high, wind=strong and outlook=sunny.

Day	Outlook	Temperature	Humidity	Wind	Play Golf
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes

5 | L3 | 2

MODULE-III

13a Construct a decision tree for the following set of training samples. 5 | L3 | 3

Instance	**Classification**	**m1**	**m2**
1	+	T	T
2	+	T	T
3	+	F	T
4	-	T	F
5	+	F	F

<table>
<tr><td>b</td><td>OR
Obtain a simple linear regression for the data given in the table below assuming that Y is the independent variable. Predict the value of Y, if X=75.
<table><tr><td>X</td><td>50</td><td>55</td><td>60</td><td>65</td><td>70</td></tr><tr><td>Y</td><td>42</td><td>44</td><td>46</td><td>48</td><td>50</td></tr></table></td><td>5</td><td>L3</td><td>3</td></tr>
<tr><td></td><td>MODULE-II &III</td><td></td><td></td><td></td></tr>
<tr><td>14 a</td><td>Explain how machines learn with suitable diagrams.</td><td>5</td><td>L2</td><td>2</td></tr>
<tr><td></td><td>OR</td><td></td><td></td><td></td></tr>
<tr><td>b</td><td>Explain how to construct classification rules from decision trees.</td><td>5</td><td>L2</td><td>3</td></tr>
<tr><td></td><td>*******</td><td></td><td></td><td></td></tr>
<tr><td>b</td><td></td><td></td><td></td><td></td></tr>
</table>

TKM COLLEGE OF ENGINEERING, KOLLAM-5
Department of Computer Applications
I Semester MCA
Internal Assessment (Offline) March 2021
Course with Code: 20MCA107 ADVANCED SOFTWARE ENGINEERING
Scheme of Valuation/Answer Key

Time:2Hrs **Maximum Marks: 50**

Qn.No	PART –A Answer all questions
1	Definition of custom assertion -1.5 mark,example-1.5 mark
2	Unit test explanation-1 mark, use of unit tests-2 mark
3	At least three difference between factory method and abstract factory method-1 mark

	each
4	Concept of Anti pattern – 2 mark, example -1 mark
5	Dataflow testing and its usage carries 1.5 marks each
6	Explanation of refactor method carries 3 mark
7	At least three characteristics of agility in agile frame work -1 mark each
8	Comparison between pair wise and state transition testing -2 mark, with example -1 mark
9	Usage of version control –at least three points – 1 mark each
10	Differentiating continuous delivery and continuous deployment with neat sketch –(2 mark,sketch-1 mark)
	PART -B
11a	Definition of structural design pattern- 1mark Types of structural design pattern-1 mark Explanation of any two with its structure -3 mark
11b	Explanation of assertion with example-2.5 mark Explanation of expected error test with example-2.5 mark
12 a	About SCRUM-2 mark Phases with explanation -3 mark
12 b	Any three testing methodologies -3.5 mark If explaining with example – 1.5 mark each
13 a	Use of xunit architecture- 1 mark Phases with explanation and sketch – 4 mark
13 b	Explanation of automated regression testing -2 mark Features carries 3 mark
14 a	Role of continuous integration in SCM- 2 mark Strategy for its implementation-3 mark
14 b	Explanation of deployment pipeline -1 mark Detailed sketch carries 1 mark ,stages with its explanation carries 3 mark