

Time: 2 hour 30 minutes

Max. Marks: 80

Q.1	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks (2 marks each)																				
1.	If dimensionality reduction is performed on a record data matrix, the transformed data matrix_____																				
Option A:	has reduced number of rows																				
Option B:	has reduced number of columns																				
Option C:	has reduced number of both rows and columns																				
Option D:	has same number of rows and columns																				
2.	Consider the following data: 4, 8, 9, 15, 21, 21, 24, 25, 26, 28, 29, 34. Partition the given data with Bin size: 4. What is the output obtained after smoothing the data by Bin Boundaries.																				
Option A:	Bin 1: 4, 4, 4, 15	Bin 2: 21, 21, 25, 25	Bin 3: 26, 26, 26, 34																		
Option B:	Bin 1: 4, 4, 15, 15	Bin 2: 21, 21, 21, 25	Bin 3: 26, 26, 34, 34																		
Option C:	Bin 1: 4, 15, 15, 15	Bin 2: 21, 25, 25, 25	Bin 3: 26, 26, 26, 34																		
Option D:	Bin 1: 4, 4, 4, 15	Bin 2: 21, 25, 25, 25	Bin 3: 26, 26, 26, 34																		
3.	Knowledge discovery in databases is referred to																				
Option A:	Non Trivial process of choosing dataset																				
Option B:	Non Trivial process for identifying useful patterns in data																				
Option C:	Non Trivial process for identifying invalid patterns in data																				
Option D:	Non Trivial process of creating patterns in data																				
4.	For the given confusion matrix compute recall <div style="text-align: center;"><div>Predicted data</div><table><tr><td></td><td>Cancer Classes</td><td>Yes</td><td>No</td><td>Total</td></tr><tr><td rowspan="3">Actual data</td><td>Yes</td><td>90</td><td>210</td><td>300</td></tr><tr><td>No</td><td>140</td><td>9560</td><td>9700</td></tr><tr><td>Total</td><td>230</td><td>9770</td><td>10000</td></tr></table></div>				Cancer Classes	Yes	No	Total	Actual data	Yes	90	210	300	No	140	9560	9700	Total	230	9770	10000
	Cancer Classes	Yes	No	Total																	
Actual data	Yes	90	210	300																	
	No	140	9560	9700																	
	Total	230	9770	10000																	
Option A:	20%																				
Option B:	30%																				
Option C:	40%																				
Option D:	45%																				
5.	You are given reviews of food quality of few restaurants as Good, Average or Poor. Finding reviews of a new restaurant is an example of_____																				
Option A:	Classification																				
Option B:	Regression																				
Option C:	Clustering																				
Option D:	Association mining																				

6.	BIRCH falls under which clustering approach												
Option A:	Partitioning approach												
Option B:	Hierarchical approach												
Option C:	Density-based approach												
Option D:	Distribution based approach												
7.	Given {2,4,3,10,11,12,20,25,30}, Assume $k=2$ and initial means are $m_1=4$, $m_2=11$. Apply k-means clustering technique and find its output after 1st iteration												
Option A:	$K_1 = \{2,3,4,10,11,12\}$ $K_2 = \{20,30,25\}$												
Option B:	$K_1 = \{2,3,4\}$ $K_2 = \{10,11,12,20,30,25\}$												
Option C:	$K_1 = \{2,3\}$ $K_2 = \{4,10,11,12,20,30,25\}$												
Option D:	$K_1 = \{2,3,4,10\}$ $K_2 = \{11,12,20,30,25\}$												
8.	In one of the frequent item-set examples, it is observed that if milk and bread are bought then eggs are also purchased by the customers. After generating an association rule among the given set of items, it is inferred												
Option A:	{Milk} is antecedent and {eggs} is consequent												
Option B:	{Milk} is antecedent and the item set {bread, eggs} is consequent												
Option C:	The item set {milk, bread} is consequent and {eggs} is antecedent												
Option D:	The item set {milk, bread} is antecedent and {eggs} is consequent												
9.	For the given transactional database compute confidence for the rule Milk \Rightarrow Beer <table border="1" data-bbox="408 1039 970 1361"> <thead> <tr> <th>TID</th><th>Items</th></tr> </thead> <tbody> <tr> <td>1</td><td>Bread, Milk</td></tr> <tr> <td>2</td><td>Bread, Diaper, Beer, Eggs</td></tr> <tr> <td>3</td><td>Milk, Diaper, Beer, Coke</td></tr> <tr> <td>4</td><td>Bread, Milk, Diaper, Beer</td></tr> <tr> <td>5</td><td>Bread, Milk, Diaper, Coke</td></tr> </tbody> </table>	TID	Items	1	Bread, Milk	2	Bread, Diaper, Beer, Eggs	3	Milk, Diaper, Beer, Coke	4	Bread, Milk, Diaper, Beer	5	Bread, Milk, Diaper, Coke
TID	Items												
1	Bread, Milk												
2	Bread, Diaper, Beer, Eggs												
3	Milk, Diaper, Beer, Coke												
4	Bread, Milk, Diaper, Beer												
5	Bread, Milk, Diaper, Coke												
Option A:	20%												
Option B:	50%												
Option C:	40%												
Option D:	60%												
10.	_____ is an interactive computer-based application that combines data and mathematical models to help decision makers solve complex problems faced in managing the public and private enterprises and organizations.												
Option A:	Data Mining												
Option B:	Data dredging												
Option C:	Decision support system												
Option D:	Artificial Intelligence system												

- A Define data warehouse. Describe different OLAP operations in detail 10
- B Apply Naive Bayes classifier algorithm to the dataset given below, and classify the unknown data sample? 10
- Given all the previous patients I've seen (below are their symptoms and their diagnosis)

chills	runny nose	headache	fever	flu ?
Y	N	Mild	Y	N
Y	Y	No	N	Y
Y	N	Strong	Y	Y
N	Y	Mild	Y	Y
N	N	No	N	N
N	Y	Strong	Y	Y
N	Y	Strong	N	N
Y	Y	Mild	Y	Y

Do I believe that patient with following symptoms has the flu?

chills	runny nose	headache	fever	flu ?
Y	N	Mild	Y	?

- C Explain multi-level and multidimensional association rules with example 10

Q.3 Solve any Two Questions out of Three

- A Suppose we have six objects with name A, B, C, D, E and F. Apply single linkage clustering and draw dendrogram for the given data. 10

	X	Y
A	1	1
B	1.5	1.5
C	5	5
D	3	4
E	4	4
F	3	3.5

- B Suppose the data for analysis includes the attribute age. The age values for data tuples are (in increasing order): 13,15,16,16,19,20,20,21,22,22,25,25,25,25,30,33,33,35,35,35,35,36,40,45,46,52,70 10
- What is mean of data? What is median of data?
 - What is mode of data? Comment on data's modality.
 - What is mid-range of data?
 - Give the five point summary of the data.
 - Show box plot of the data

- C What is Business Intelligence (BI)? Explain BI architecture in detail 10

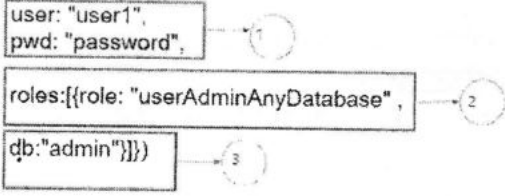
Q.4 Solve any Two Questions out of Three

- A** Briefly explain Bagging and Boosting of classifiers **10**
- B** For the table given, apply Apriori algorithm and show frequent item set and strong association rules. Assume Minimum Support of 30% and Minimum confidence of 70%. **10**

TID	Items
01	1,3,4,6
02	2,3,5,7
03	1,2,3,5,8
04	2,5,9,10
05	1,4

- C** What is an outlier? Describe methods used for outlier analysis. **10**

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	In order to identify the users, web analytics tools need to report on?
Option A:	clickstream
Option B:	user sessions
Option C:	Unique users
Option D:	Page Views
2.	Which of the following is the valid arrow function?
Option A:	let sum = (x: number, y: number) =>x+y;
Option B:	let sum() : (x: number, y: number) : number
Option C:	let sum = (x: number, y: number) => return x+y;
Option D:	let sum(x: number, y: number) =>x+y;
3.	Which of the following is the correct syntax for writing AngularJS expressions?
Option A:	(expression)
Option B:	{{expression}}
Option C:	{{{expression}}}
Option D:	[expression]
4.	Which of the following directive is used to bind the application data to the HTML view in AngularJS?
Option A:	ng-app directive
Option B:	ng-model directive
Option C:	ng-bind directive
Option D:	ng-init directive
5.	— is a MongoDB method to find any document already inserted into the collection.
Option A:	findItem
Option B:	fineOne
Option C:	findDocument
Option D:	findCollection
6.	After starting the mongo shell, your session will use the ----- database by default.
Option A:	mongo
Option B:	master
Option C:	test
Option D:	primary
7.	The app.run function in Flask does NOT have this parameter.
Option A:	host
Option B:	port
Option C:	debug
Option D:	checks
8.	Which of the following internet standard not used by AJAX
Option A:	XML

Option B:	CSS
Option C:	XMLHttpRequest Object
Option D:	JSON
9.	<pre>db.createUser({ user: "user1", pwd: "password", roles: [{role: "userAdminAnyDatabase", db: "admin"}]})</pre>  <p>For given code choose write interpretation-</p>
Option A:	<ol style="list-style-type: none"> 1. Specifies user name and password. 2. Specifies user role as 'user1'. 3. Writes information to the 'admin' database
Option B:	<ol style="list-style-type: none"> 1. Specifies user name and password. 2. Specifies user role as an admin. 3. Writes information to the 'admin' database
Option C:	<ol style="list-style-type: none"> 1. Specifies user name and password. 2. Specifies user role as an admin. 3. Writes information to the database
Option D:	<ol style="list-style-type: none"> 1. Specifies user name and password. 2. Specifies user role as an admin. 3. Writes information to the 'local' database
10.	update role of "MSD" as "C and WK" from emp collection
Option A:	>db.empl.update({name:"MSD"},{\$set:{role:"c and WK"}})
Option B:	>db.emp.update({name:"MSD"},{\$set:{role:"c and WK"}})
Option C:	>db.emp.update({name:"MSD"},{\$set:{role:"c" role:" WK"}})
Option D:	>db.empl.update({name:"MSD"},{set:{role:"c and WK"}})

Q2 (20 Marks)	Solve any Two Questions out of Three	10 marks each
A	List and Explain factors for Measuring the success of a E-commerce website	
B	Compare and contrast Web 1.0, Web 2.0 and Web 3.0.	
C	Write code in typescript to explain concept of Inheritance.	

Q3. (20 Marks)	Solve any Two Questions out of Three	10 marks each
A	Illustrate the use of expressions in AngularJS with suitable example.	
B	Explain Routing using ng-Route, ng-Repeat, ng-style, ng-view with example	
C	List and explain mongo shell commands	

Q4. (20 Marks)	Solve any Two Questions out of Three	10 marks each
A	Write Flask script to display "Welcome to Flask framework!"	
B	Explain dynamic url building in flask	
C	Explain how AJAX works with a diagram? Describe 4 standards used by AJAX.	

Examination Summer 2022

Examination Commencement from 17th May to 31th May 2022

Program: T.E. (Information Technology)
(Choice Based) SEM VI (REV-2020-21) (C Scheme)
TE Semester VI

Course Code ITC603

Time: 2 hours and 30 minutes

Max. Marks: 80

Q1.	Choose the correct option for the following questions. All the questions are compulsory and carry equal marks
1.	Which is better for avoiding jamming
Option A:	Direct sequence spread spectrum
Option B:	Frequency hopping spread spectrum
Option C:	Time hopping spread spectrum
Option D:	Code division hopping spread spectrum
2.	CDMA uses handoff
Option A:	Hard handoff
Option B:	Hard & Soft hand off
Option C:	Soft Handoff
Option D:	Orthogonal Handoff
3.	UMTS use which multiple access technique?
Option A:	FDMA
Option B:	TDMA
Option C:	CDMA
Option D:	WCDMA
4.	2.5G allows standard web pages to be viewed in compressed format on mobile
Option A:	SAP
Option B:	WEP
Option C:	RAP
Option D:	WAP
5.	Carrier spacing used in GSM
Option A:	290KHz
Option B:	200KHz
Option C:	490KHz
Option D:	450KHz
6.	Wi-Max uses
Option A:	Channel division multiplexing
Option B:	Space division multiplexing
Option C:	Time division multiplexing
Option D:	Orthogonal frequency division multiplexing
7.	L2CAP layer in Bluetooth
Option A:	Handles data encryption
Option B:	Administration and control of all link actions
Option C:	Provides connection-oriented and connectionless data services to upper-layer protocols
Option D:	To join piconet

8.	What is the frequency range of IEEE 802.11a standard?
Option A:	2.4 Gbps
Option B:	5 Gbps
Option C:	2.4 Ghz
Option D:	5 GHz
9.	In Bluetooth the operating frequency band is
Option A:	2.4GHz
Option B:	9.4GHz
Option C:	8.4GHz
Option D:	9.4GHz
10.	WPA uses _____ Algorithm to check integrity of the packets.
Option A:	TKIP
Option B:	SAP
Option C:	DOA
Option D:	RAP

Q2 (20 Marks Each)	Solve any Two Questions out of Three	10 marks each
A	Explain the LTE network architecture.	
B	i. Compare FDMA, TDMA and CDMA ii. Explain the protocol architecture of IEEE802.11	
C	Draw and explain the GSM Time slot hierarchy	

Q3 (20 Marks Each)	Solve any Two Questions out of Three	10 marks each
A	Explain spread spectrum and briefly outline DSSS and FHSS	
B	Draw and explain WSN protocol layer architecture. Give the applications of WSN	
C	State the features of WiMAX. Draw and explain the architecture of WiMAX	

Q4 (20 Marks Each)	Solve any Two Questions out of Three	10 marks each
A	i. State the characteristics of MANET. ii. Give the Features of VANET and E-VANET	
B	i. Write a Short Note on UMTS security ii. Compare WEP and WPA2 protocol.	
C	State the features of Cisco Unified Wireless Network. Draw and explain Cisco Unified Wireless Network Architecture	

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	_____ are the action making parts of an agent that takes in the input for the user.
Option A:	Actuators
Option B:	Sensors
Option C:	Environments
Option D:	Performance
2.	_____ is optimal search algorithm in terms of heuristics
Option A:	Min Max Algorithm
Option B:	Depth Limited Search
Option C:	Hill Climbing Algorithm
Option D:	A* Algorithm
3.	P in PEAS stands for
Option A:	Performance Criteria
Option B:	Performance Evaluation
Option C:	Performance Measure
Option D:	Performance Environment
4.	_____ is called as greedy local search
Option A:	Hill Climbing
Option B:	DFS
Option C:	BFS
Option D:	Uniform cost
5.	Backward Chaining and Forward Chaining in AI is
Option A:	Goal-driven and Data-driven approach respectively
Option B:	Bottom -Up and Top-down Approach respectively
Option C:	Goes from fact to result and goes from result to fact respectively.
Option D:	Uses "BFS" and "DFS" respectively
6.	Identify the one which is not a type of learning
Option A:	Reinforcement Learning
Option B:	Semi Unsupervised Learning
Option C:	Supervised Learning
Option D:	Unsupervised Learning
7.	Machine learning is a subset of which of the following.
Option A:	Artificial Intelligence
Option B:	Deep Learning
Option C:	Data Learning
Option D:	Statistics
8.	Which of the following is not a univariate graphical EDA technique?
Option A:	Histograms

Option B:	Box Plots
Option C:	Stem and Leaf plots
Option D:	Pair plots
9.	Which statistical tool should be used to test the equality of 3 or more population means?
Option A:	ANOVA
Option B:	T-test
Option C:	Chi-square test
Option D:	Interval Estimation
10.	Which is NOT the correct statement about the InterQuartile Range.
Option A:	The interquartile range tells you the spread of the middle half of your distribution.
Option B:	$IQR = Q3 - Q1$
Option C:	In boxplot upper whisker indicates $Q3$
Option D:	In boxplot IQR is indicated by the edges of the rectangle

Q2	10 marks each
A	<p>Solve Resolution:</p> <ol style="list-style-type: none"> 1. All people that are not poor and are smart are happy. 2. Those people that read are not stupid. 3. John can read and is wealthy. 4. Happy people have exiting lives. <p>Can anyone be found with an exciting life?</p>
B	What do you mean by EDA ? Explain different categorizations of EDA. For each type of EDA explain 1 technique that belongs to it in detail.

Q3	10 marks each
A	Elaborate in detail the steps in developing a Machine Learning application with architectural diagram.
B	<ol style="list-style-type: none"> 1. Illustrate with diagram how Goal based agent works. 2. Describe PEAS and also write down the PEAS representations for Automated car driver

Q4	10 marks each																													
A	Compare min max and alpha Bea pruning algorithms.																													
B	Consider you are performing ML for predicting housing prices you have trained three models and following data summarizes the predicted house price by each model for 5 different trial runs. <table><tr><th rowspan="2">Model Code</th><th colspan="5">House Price Predicted (Lakh Rs)</th></tr><tr><th>Trial 1</th><th>Trial 2</th><th>Trial 3</th><th>Trial 4</th><th>Trial 5</th></tr><tr><td>A</td><td>3.5</td><td>3.4</td><td>3.8</td><td>3.5</td><td>3.4</td></tr><tr><td>B</td><td>3.9</td><td>3.8</td><td>3.7</td><td>3.9</td><td>3.6</td></tr><tr><td>C</td><td>3.5</td><td>3.3</td><td>3.6</td><td>3.5</td><td>3.8</td></tr></table>	Model Code	House Price Predicted (Lakh Rs)					Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	A	3.5	3.4	3.8	3.5	3.4	B	3.9	3.8	3.7	3.9	3.6	C	3.5	3.3	3.6	3.5	3.8
Model Code	House Price Predicted (Lakh Rs)																													
	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5																									
A	3.5	3.4	3.8	3.5	3.4																									
B	3.9	3.8	3.7	3.9	3.6																									
C	3.5	3.3	3.6	3.5	3.8																									
Perform One way ANOVA F Test on this data and comment on whether the mean house price predicted by models A, B, C are same with level of significance 0.05. (Use of F Table is allowed)																														

University of Mumbai
Examinations summer 2022

AI and DS1 SEM VI IT 27/05/22

Corrections

Q1 is of 20 marks. Each subquestion is of 2 marks.

Answer any 2 in questions 2,3 and 4

Q2 C	Compare Linear Regression Vs Logistics Regression with suitable diagrams and formulas.																																																
Q3 C	What do you mean by covariance and correlation ? Explain what the range of coefficients of correlation and covariance suggest. Calculate COV(Age, Strength) and CORR(Age, Strength) for following data. How do you interpret these values?																																																
	<table><tr><td>Subject ID</td><td>Age</td><td>Strength</td></tr><tr><td>A</td><td>38</td><td>20</td></tr><tr><td>B</td><td>62</td><td>15</td></tr><tr><td>C</td><td>22</td><td>30</td></tr><tr><td>D</td><td>38</td><td>21</td></tr><tr><td>E</td><td>45</td><td>18</td></tr><tr><td>F</td><td>69</td><td>12</td></tr><tr><td>G</td><td>75</td><td>14</td></tr><tr><td>H</td><td>38</td><td>28</td></tr><tr><td>I</td><td>80</td><td>9</td></tr><tr><td>J</td><td>32</td><td>22</td></tr><tr><td>K</td><td>51</td><td>20</td></tr><tr><td>L</td><td>56</td><td>19</td></tr><tr><td>M</td><td>21</td><td>28</td></tr><tr><td>N</td><td>34</td><td>23</td></tr><tr><td>O</td><td>76</td><td>14</td></tr></table>	Subject ID	Age	Strength	A	38	20	B	62	15	C	22	30	D	38	21	E	45	18	F	69	12	G	75	14	H	38	28	I	80	9	J	32	22	K	51	20	L	56	19	M	21	28	N	34	23	O	76	14
	Subject ID	Age	Strength																																														
	A	38	20																																														
	B	62	15																																														
	C	22	30																																														
	D	38	21																																														
	E	45	18																																														
	F	69	12																																														
	G	75	14																																														
	H	38	28																																														
	I	80	9																																														
	J	32	22																																														
	K	51	20																																														
	L	56	19																																														
M	21	28																																															
N	34	23																																															
O	76	14																																															
Q4 C	<p>i. Explain forward chaining and backward chaining algorithm with the help of example.</p> <p>ii. What is heuristic function? Which search algorithm types use it?</p>																																																

Correction in Q.2A) Q.4A)

Q2 A	4. Happy people have existing life.
Q4 A	Compare min-max and alpha-beta pruning algorithms

Program No. & Name of the Examination: 1T01236 // T.E.(Information Technology Engineering) (SEM-VI)(Choice Base Credit Grading System) (R- 19) (C Scheme)

Subject (Paper Code) : 89388 // Ethical Hacking and Forensic (DLOC)

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
Q1.	Definition of cybercrime is
Option A:	a criminal activity involving a computer, networked device or a network
Option B:	criminal activity involving a computer
Option C:	criminal activity involving a networked device
Option D:	any criminal activity that involves a network
Q2.	A copy which includes all necessary parts of evidence, which is closely related to the original evidence.
Option A:	Digital Evidence
Option B:	Best Evidence
Option C:	Original Evidence
Option D:	Complete Evidence
Q3.	CSIRT stands for
Option A:	Computer security incident response team
Option B:	Computer software incident resource team
Option C:	Common security incident resolution team
Option D:	Computer security incident resource team
Q4.	Forensic Duplication is necessary
Option A:	as it preserves original digital evidence & allows recreation of the duplicate image
Option B:	as it creates restored image
Option C:	as it creates and stores mirror image
Option D:	as it helps in live system duplication
Q5.	Analyzing data collected from different sites, Firewalls and IDS is called as..?
Option A:	Computer Forensics
Option B:	Network Forensics
Option C:	Mobile Devices Forensics
Option D:	Memory Forensics
Q6.	Which one among the following statements is not a goal of good forensic report writing
Option A:	describe accurately the details of the incident
Option B:	Be understandable to decision makers
Option C:	Be able to withstand the legal scrutiny
Option D:	Cannot be easily referenced
Q7.	Which among the following is not an example of cyber crime
Option A:	SQL injection
Option B:	Identity theft
Option C:	Hacking
Option D:	Designing antivirus
Q8.	If there ought to be no doubt about the reality of the specialist's decision, then the evidence is said to be...?
Option A:	Authentic
Option B:	Admissible
Option C:	Believable

Option D:	Reliable
Q9.	In central incident response team how many teams handle incidents occurring in whole organization?
Option A:	1
Option B:	2
Option C:	3
Option D:	4
Q10.	Restoration Process involves
Option A:	blind sector to sector copy of the duplicate file
Option B:	collection of Digital Evidence
Option C:	creation of response toolkit
Option D:	check the dependencies

Q2 (20 Marks)	Solve any Two Questions out of Three	10 marks each
A	Compare active attacks vs Passive attacks. Classify the cybercrimes and explain any one briefly	
B	Explain the phases of incident response Methodology with neat diagram	
C	Explain Volatile Data Collection from Windows system	

Q3 (20 Marks)	Solve any Two Questions out of Three	10 marks each
A	What are possible investigation phase carried out in Data Collection and Analysis	
B	What do you understand by social engineering? Give classification	
C	Briefly explain Types of digital Evidence with examples.	

Q4 (20 Marks)	Solve any Two Questions out of Three	10 marks each
A	Explain importance of forensic duplication and its methods.	
B	Explain various guidelines for digital forensic report writing along with its goals.	
C	Discuss the techniques of tracing an email message.	