

# UOM Exam Second half 2021\_ Question paper\_R2019/Comp/CSDLO5013 - Advance Database Management System/Sem-V

Dear Student,

Please note before you attempt this section of examination:

1. Q1, Q2, Q3 and Q4 carry 20 marks each.
2. This paper contains 20 Marks MCQ and 60 marks subjective section for 150 minutes duration.
3. It is mandatory for all the students to upload their answer papers in a single PDF format only.
4. You have to write Date of Examination, Seat number, Program, Scheme and semester, Subject name, Signature on EVERY PAGE.
5. Remain in the meet with your camera on and you in clear view throughout the duration of the exam.

---

\* Required

1. Email \*

---

2. Student Name (As per exam form filled) \*

---

3. Seat No \*

Refer Hall ticket

---

4. Class \*

*Mark only one oval.*

TE3

TE4

TE9

5. Roll Number \*

---

Solve Questions as per the instructions given separately.

- Please upload a single PDF for Q1 to Q4
- For MCQs Question write Question number & correct option with complete text in option.
- Q2 to Q4 are subjective questions - Solve Questions as per the instructions and marks allotted.

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	New lock acquired at _____
Option A:	Shrinking Phase
Option B:	Release phase
Option C:	Commit phase
Option D:	Growing Phase
2.	Which of the following code will return a valid JSON object?
Option A:	<code>JSON.parse("{ 'FirstName': 'John', 'LastName': 'Doe' }");</code>
Option B:	<code>JSON.parse({"FirstName": "John", "LastName": "Doe"});</code>
Option C:	<code>JSON.parse("({ 'FirstName': 'John', 'LastName': 'Doe' })");</code>
Option D:	<code>JSON.parse({ "FirstName": "John", "LastName": "Doe" });</code>
3.	Which of the following will happen after the following sequence of events is completed as per Two Phase Commit (2PC) protocol? TC sends <Prepare T1> to P1 and P2. P1 sends <Ready T1> to TC. P2 sends <Abort T1> to TC.
Option A:	TC sends <Commit T1> to P1 and P2
Option B:	TC sends <Abort T1> to P1 and P2.
Option C:	TC sends <Commit T1>, only to participant P1.
Option D:	TC sends <Abort T1>, only to participant P2.
4.	_____ is used to give comments in XML Document.
Option A:	<code>&lt;!-- --!&gt;</code>
Option B:	<code>&lt;/-- --&gt;</code>
Option C:	<code>&lt;!-- --&gt;</code>
Option D:	<code>&lt;?-- ?</code>
5.	A distributed database is a collection of data which belong _____ to the same system but are spread over the _____ of the network.
Option A:	Logically, sites
Option B:	Physically, sites
Option C:	Database, DBMS
Option D:	DBMS, Database
6.	Column in NoSQL is represented by
Option A:	Field
Option B:	Document
Option C:	Database
Option D:	Collection

7.	Which of the following are the simplest NoSQL databases?
Option A:	Wide-column
Option B:	Document
Option C:	Key-value
Option D:	JSON
8.	Which clause is used to search the data with a specified pattern?
Option A:	OPTIONAL MATCH
Option B:	MATCH
Option C:	WHERE
Option D:	FIND
9.	In Neo4J joins are represent as?
Option A:	Constraints
Option B:	Nodes
Option C:	Relationships
Option D:	Traversal
10.	Which of the following is true concerning a global transaction?
Option A:	The required data are at one local site and the distributed DBMS routes requests as necessary.
Option B:	The required data are located in at least one nonlocal site and the distributed DBMS routes requests as necessary.
Option C:	The required data are at one local site and the distributed DBMS passes the request to only the local DBMS.
Option D:	The required data are located in at least one nonlocal site and the distributed DBMS passes the request to only the local DBMS.

<b>Q2</b>	<b>Solve any Two Questions out of Three</b>	<b>10 marks each</b>
A	Explain different types of NOSQL Database.	
B	Mention the Datatypes in MongoDB. Which function is used to save and update document in <u>MogoDB</u> ?	
C	Describe in detail Graph Database with Neo4J	

<b>Q3</b>	<b>Solve any Two Questions out of Three</b>	<b>10 marks each</b>
A	Explain in detail Distributed DBMS Architecture.	
B	Explain 3PC in detail.	
C	What is XML? Explain XML Schema document with example.	

<b>Q4</b>	<b>Solve any Two Questions out of Three</b>	<b>10 marks each</b>
A	Explain the different methods of concurrency control in distributed Database.	
B	Explain JSON data types with suitable examples.	
C	Explain Data distribution model in NoSQL	

6. Please Upload complete scanned answer copy in a single PDF file. \*

Files submitted:

7. Have you uploaded correct scanned copy of the answer sheets. \*

*Mark only one oval.*

YES

---

This content is neither created nor endorsed by Google.

Google Forms

# UOM Exam Second half 2021\_ Question paper\_R2019/Comp/CSC503 - Computer Network/Sem-V

Dear Student,

Please note before you attempt this section of examination:

1. Q1, Q2, Q3 and Q4 carry 20 marks each.
2. This paper contains 20 Marks MCQ and 60 marks subjective section for 150 minutes duration.
3. It is mandatory for all the students to upload their answer papers in a single PDF format only.
4. You have to write Date of Examination, Seat number, Program, Scheme and semester, Subject name, Signature on EVERY PAGE.
5. Remain in the meet with your camera on and you in clear view throughout the duration of the exam.

---

\* Required

1. Email \*

---

2. Student Name (As per exam form filled) \*

---

3. Class \*

*Mark only one oval.*

TE3

TE4

TE9

4. Roll Number \*

---

5. Seat No \*

Refer Hall ticket

---

Solve Questions as per the instructions given separately.

- Please upload a single PDF for Q1 to Q4
- For MCQs Question write Question number & correct option with complete text in option.
- Q2 to Q4 are subjective questions - Solve Questions as per the instructions and marks allotted.

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	Transmission media are usually categorized as _____
Option A:	metallic or non-metallic
Option B:	determinate or indeterminate
Option C:	guided or unguided
Option D:	fixed or unfixed
2.	Automatic repeat request error management mechanism is provided by
Option A:	media access control sublayer
Option B:	logical link control sublayer
Option C:	application access control sublayer
Option D:	network interface control sublayer
3.	Which of the following device is used to connect two systems, especially if the systems use different protocols?
Option A:	Gateway
Option B:	Repeater
Option C:	Bridge
Option D:	Hub
4.	If a class B network on the Internet has a subnet mask of 255.255.248.0, what is the maximum number of hosts per subnet?
Option A:	512
Option B:	1024
Option C:	1022
Option D:	2046
5.	What is the format of IPV6 address?
Option A:	128
Option B:	64
Option C:	32
Option D:	16
6.	The root of the DNS tree is _____
Option A:	a string of 64 characters
Option B:	a string of 32 characters
Option C:	a string of 63 characters
Option D:	an empty string



7.	Which one of the following protocols is NOT used to resolve one form of address to another one?
Option A:	DHCP
Option B:	SNMP
Option C:	FTP
Option D:	ARP
8.	An endpoint of an inter-process communication flow across a computer network is called _____
Option A:	Port
Option B:	Socket
Option C:	Pipe
Option D:	Machine
9.	What is the maximum number of IP addresses that can be assigned to hosts on a local subnet that uses the 255.255.255.224 subnet mask?
Option A:	8
Option B:	16
Option C:	32
Option D:	30
10.	Which of the following is the multiple access protocol for channel access control?
Option A:	ARC
Option B:	CSMA/CD and CSMA/CA
Option C:	Sliding window protocol
Option D:	HDLC

<b>Q2.</b>	<b>Solve any Four out of Six</b>	<b>5 marks each</b>
A	Explain different types of CSMA protocols	
B	Explain IPV 4 header format with appropriate diagram	
C	Explain different internetworking devices	
D	Explain token bucket algorithm	
E	Explain optic fiber with appropriate diagram	
F	How does congestion control work in TCP?	

<b>Q3.</b>	
<b>A</b>	<b>Solve any Two</b> <span style="float: right;"><b>5 marks each</b></span>
i.	Explain DNS with appropriate diagram
ii.	Explain 3 - way connection establishment handshake in TCP
iii.	Given the network address 132.21.0.0, find the class, the block, and the range of the addresses, how many host addresses can be created?
<b>B</b>	<b>Solve any One</b> <span style="float: right;"><b>10 marks each</b></span>
i.	Explain Distance vector routing. What is its limitations and how are they overcome?
ii.	Explain pure Aloha and slotted Aloha with appropriate diagram

<b>Q4.</b>	
<b>A</b>	<b>Solve any Two</b> <span style="float: right;"><b>5 marks each</b></span>
i.	Explain Dijkstra's algorithm with appropriate diagram
ii.	Explain SMTP
iii.	Consider a message represented by the polynomial $M(x) = x^5 + x^4 + x$ . Consider a generating polynomial $G(x) = x^3 + x^2 + 1$ . Generate a 3 bit CRC code and show what will be the transmitted frame?
<b>B</b>	<b>Solve any One</b> <span style="float: right;"><b>10 marks each</b></span>
i.	An ISP is granted a block of addresses starting with 190.100.0.0/16 (65536) addresses the ISP needs to distribute these addresses to three groups of customers as follows: 1) The first group has 64 customers, each needs 256 addresses 2) The second group has 128 customers; each need 128 bits. 3) The third group has 128 customers; each need 64 addresses.
ii.	Explain HTTP & Telnet protocols in detail.

6. Please Upload complete scanned answer copy in a single PDF file. \*

Files submitted:

7. Have you uploaded correct scanned copy of the answer sheets. \*

Mark only one oval.

YES

---

This content is neither created nor endorsed by Google.

# Google Forms

# UOM Exam Second half 2021\_ Question paper\_R2019/Comp/CSC504 - Data Warehousing and Mining/Sem-V

Dear Student,

Please note before you attempt this section of examination:

1. Q1, Q2, Q3 and Q4 carry 20 marks each.
2. This paper contains 20 Marks MCQ and 60 marks subjective section for 150 minutes duration.
3. It is mandatory for all the students to upload their answer papers in a single PDF format only.
4. You have to write Date of Examination, Seat number, Program, Scheme and semester, Subject name, Signature on EVERY PAGE.
5. Remain in the meet with your camera on and you in clear view throughout the duration of the exam.

---

\* Required

1. Email \*

---

2. Student Name (As per exam form filled) \*

---

3. Seat No \*

Refer Hall ticket

---

4. Class \*

*Mark only one oval.*

TE3

TE4

TE9

Solve Questions as per the instructions given separately.

- Please upload a single PDF for Q1 to Q4
- For MCQs Question write Question number & correct option with complete text in option.
- Q2 to Q4 are subjective questions - Solve Questions as per the instructions and marks allotted.

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	Assume that the database D is given by the table below. Follow single link technique to find clusters in D. Use Euclidean distance measure. (A : 2,1) (B: 3,2) (C:3,2) (D:4,4) (E:5,5) Final Clusters are
Option A:	k1(A,B,C), k2(D,E)
Option B:	k1(A,B), k2(C,D,E)
Option C:	k1(A), k2(B,C,D,E)
Option D:	k1(B,C), k2(A,D,E)
2.	_____ involves capturing data through the transaction logs or Database triggers or capture in Source application.
Option A:	Immediate data extraction
Option B:	deferred data extraction
Option C:	quick data extraction
Option D:	delay data extraction
3.	Data mining is?
Option A:	time variant non-volatile collection of data
Option B:	The actual discovery phase of a knowledge
Option C:	The stage of selecting the right data
Option D:	The pattern matching algorithm
4.	Web structure mining comprises of a process of _____
Option A:	Understanding trends and associations
Option B:	Extracting useful information from server
Option C:	Discovering patterns and structures
Option D:	Using graphs and network mining theory
5.	Confidence C is defined as _____. ( consider Rule L -> M)
Option A:	Support_count (L)/Support_count (M)
Option B:	Support_count (L and M occurs together)/Support_count (L)
Option C:	Support_count (M)/Support_count (L)
Option D:	Support_count ((L and M occurs together)/Support_count (M)
6.	Maximal frequent itemset defines _____
Option A:	If none of the immediate supersets is frequent
Option B:	If all of the immediate supersets is frequent
Option C:	If all of the immediate subsets is frequent
Option D:	If none of the immediate subsets is frequent

7.	Which of the following is not a valid Visualization technique?
Option A:	Scatter plot
Option B:	Box plot
Option C:	Histogram
Option D:	Decision tree
8.	Which of the following statements are not true?
Option A:	Clustering analysis is an unsupervised learning since it does not require labeled training data.
Option B:	When clustering, we want to put two dissimilar data objects into the same cluster.
Option C:	Clustering is a technique of analyzing and dividing data.
Option D:	Data are broken into hierarchy clusters in Hierarchical clustering algorithms
9.	_____ is the process of transforming numerical variables into categorical counterparts
Option A:	clustering
Option B:	Binning
Option C:	classification
Option D:	Association mining
10.	A database has 5 transactions T1-T5. Let min sup = 40% and min conf =65% AND T1={A,B,D,E} T2={A,C,A,B} T3={C,B,A,C} T4={B,A,D} T5={C,D} . The following are some of the strong association rules
Option A:	$(A,B) \rightarrow C$ , $(A,C) \rightarrow B$
Option B:	$C \rightarrow (A,B)$ , $(A,C) \rightarrow B$
Option C:	$B \rightarrow (A,C)$ , $(A,C) \rightarrow B$
Option D:	$A \rightarrow (B,C)$ , $B \rightarrow (A,C)$

Question 2	Solve any Two out of Three <span style="float: right;">10 marks each</span>																																																		
A	<p>Apply statistical based algorithm to obtain the actual probabilities of each event to classify the new tuple. Consider new tuple as <math>t = (Adam, M, 1.95m)</math>.</p> <p>Use the following data :</p> <table border="1" data-bbox="501 544 1281 1697" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>Person ID</th> <th>Name</th> <th>Gender</th> <th>Height</th> <th>Class</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Kristina</td> <td>Female</td> <td>1.6 m</td> <td>Short</td> </tr> <tr> <td>2</td> <td>Jim</td> <td>Male</td> <td>2 m</td> <td>Tall</td> </tr> <tr> <td>3</td> <td>Maggie</td> <td>Female</td> <td>1.9 m</td> <td>Medium</td> </tr> <tr> <td>4</td> <td>Martha</td> <td>Female</td> <td>2.1 m</td> <td>Tall</td> </tr> <tr> <td>5</td> <td>Stephanie</td> <td>Female</td> <td>1.7 m</td> <td>Short</td> </tr> <tr> <td>6</td> <td>Bob</td> <td>Male</td> <td>1.85 m</td> <td>Medium</td> </tr> <tr> <td>7</td> <td>Catherine</td> <td>Female</td> <td>1.6 m</td> <td>Short</td> </tr> <tr> <td>8</td> <td>Dave</td> <td>Male</td> <td>1.7 m</td> <td>Short</td> </tr> <tr> <td>9</td> <td>Wilson</td> <td>Male</td> <td>2.2 m</td> <td>Tall</td> </tr> </tbody> </table>	Person ID	Name	Gender	Height	Class	1	Kristina	Female	1.6 m	Short	2	Jim	Male	2 m	Tall	3	Maggie	Female	1.9 m	Medium	4	Martha	Female	2.1 m	Tall	5	Stephanie	Female	1.7 m	Short	6	Bob	Male	1.85 m	Medium	7	Catherine	Female	1.6 m	Short	8	Dave	Male	1.7 m	Short	9	Wilson	Male	2.2 m	Tall
Person ID	Name	Gender	Height	Class																																															
1	Kristina	Female	1.6 m	Short																																															
2	Jim	Male	2 m	Tall																																															
3	Maggie	Female	1.9 m	Medium																																															
4	Martha	Female	2.1 m	Tall																																															
5	Stephanie	Female	1.7 m	Short																																															
6	Bob	Male	1.85 m	Medium																																															
7	Catherine	Female	1.6 m	Short																																															
8	Dave	Male	1.7 m	Short																																															
9	Wilson	Male	2.2 m	Tall																																															
B	<p>Consider a datawarehouse for order analysis application where there are four dimensions 1. Customer 2. Order date 3 Sales person and 4. Product. Consider two measures: cost and quantity sold. Design star and Snowflake schema.</p>																																																		
C	<p>Apply Apriori algorithm to find all frequent item sets from the following table :</p> <p style="margin-left: 40px;">Minimum Support count = 2</p> <p style="margin-left: 40px;">Minimum Confidence = 70%</p>																																																		

		TID	Item	
		100	1,2,5	
		200	2,4	
		300	2,3	
		400	1,2,4	
		500	1,3	
		600	1,2,3,5	

<b>Question 3</b>	<b>Solve any Two out of Three</b>	<b>10 marks each</b>																								
A	<p>Calculate 2 clusters using k-means cluster algorithm. For finding the distance use Euclidian distance</p> <table border="1"> <thead> <tr> <th>Subject</th> <th>A</th> <th>B</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>1.0</td> <td>1.0</td> </tr> <tr> <td>2</td> <td>1.5</td> <td>2.0</td> </tr> <tr> <td>3</td> <td>3.0</td> <td>4.0</td> </tr> <tr> <td>4</td> <td>5.0</td> <td>7.0</td> </tr> <tr> <td>5</td> <td>3.5</td> <td>5.0</td> </tr> <tr> <td>6</td> <td>4.5</td> <td>5.0</td> </tr> <tr> <td>7</td> <td>3.5</td> <td>4.5</td> </tr> </tbody> </table>	Subject	A	B	1	1.0	1.0	2	1.5	2.0	3	3.0	4.0	4	5.0	7.0	5	3.5	5.0	6	4.5	5.0	7	3.5	4.5	
Subject	A	B																								
1	1.0	1.0																								
2	1.5	2.0																								
3	3.0	4.0																								
4	5.0	7.0																								
5	3.5	5.0																								
6	4.5	5.0																								
7	3.5	4.5																								
B	Describe the various multilevel association rule mining. In detail provide an example of multilevel mining with reduced support																									
C	Explain the structure of web log with the example of each field. Consider any fragment from server log.																									



<b>Question 4</b>	<b>Solve any Four Questions out of Six</b>	<b>5 marks each</b>
<b>A</b>	<b>Compare OLTP vs OLAP systems.</b>	
<b>B</b>	<b>Explain ROLAP and MOLAP models.</b>	
<b>C</b>	<b>Explain how the accuracy of a classifier can be measured. How Bagging strategy helps improving the classifier accuracy?</b>	
<b>D</b>	<b>Describe the steps involved in the FP growth algorithm. Compare and contrast its steps with Apriori algorithm</b>	
<b>E</b>	<b>Elaborate on the taxonomy of web mining</b>	
<b>F</b>	<b>In real-world data, tuples with missing values for some attributes are a common occurrence. Describe various methods for handling this problem</b>	

6. Please Upload complete scanned answer copy in a single PDF file. \*

Files submitted:

7. Have you uploaded correct scanned copy of the answer sheets. \*

*Mark only one oval.*

YES

---

This content is neither created nor endorsed by Google.

**Google Forms**

# UOM Exam Second half 2021\_ Question paper\_R2019/Comp/CSDLO 5012 - Internet Programming/Sem-V

Dear Student,

Please note before you attempt this section of examination:

1. Q1, Q2, Q3 and Q4 carry 20 marks each.
2. This paper contains 20 Marks MCQ and 60 marks subjective section for 150 minutes duration.
3. It is mandatory for all the students to upload their answer papers in a single PDF format only.
4. You have to write Date of Examination, Seat number, Program, Scheme and semester, Subject name, Signature on EVERY PAGE.
5. Remain in the meet with your camera on and you in clear view throughout the duration of the exam.

---

\* Required

1. Email \*

---

2. Student Name (As per exam form filled) \*

---

3. Seat No \*

Refer Hall ticket

---

4. Class \*

*Mark only one oval.*

TE3

TE4

TE9

5. Roll Number \*

---

Solve Questions as per the instructions given separately.

- Please upload a single PDF for Q1 to Q4
- For MCQs Question write Question number & correct option with complete text in option.
- Q2 to Q4 are subjective questions - Solve Questions as per the instructions and marks allotted.

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	If we want to define style for an unique element , then which CSS selector will we use
Option A:	Text
Option B:	Name
Option C:	Class
Option D:	Id
2.	Which one of the following method is used to retrieve the number of rows added due to an INSERT query
Option A:	affectedrows()
Option B:	changed_rows()
Option C:	affected_rows()
Option D:	rows_affected()
3.	Which of the following is correct to set "Black as Background of page?
Option A:	bgcolor = "#000000"
Option B:	<body background = "#000000">
Option C:	<body background color = "#000000">
Option D:	All of above
4.	What is the output of the following JavaScript code? <pre>&lt;script type="text/javascript"&gt; x = 4 + "4"; document.write(x); &lt;/script&gt;</pre>
Option A:	44
Option B:	8
Option C:	4
Option D:	Error
5.	What is the correct signature of _jspService() method of HttpJspPage class?
Option A:	void _jspService(HttpServletRequest request, HttpServletResponse response)
Option B:	void _jspService(HttpServletRequest request, HttpServletResponse response) throws ServletException, IOException
Option C:	void _jspService()
Option D:	void _jspService() throws ServletException, IOException
6.	JSP life cycle includes the following phases. Arrange them as they occur 1. Instantiation 2. Request processing 3. Initialization

	4. Classloading 5. Compilation of JSP Page 6. Destroy 7. Translation of JSP Page								
Option A:	6-5-41-3-2-7								
Option B:	7-5-41-3-2-6								
Option C:	7-5-43-1-2-6								
Option D:	1-5-47-3-2-6								
7.	Which of the following is the correct data flow sequence of flux concept?								
Option A:	Dispatcher->Action->Store->View								
Option B:	Action->Dispatcher->View->Store								
Option C:	Action->Dispatcher->Store->View								
Option D:	Action->Store->Dispatcher->View								
8.	Match the correct pair								
	<table border="1"> <tr> <td>1. XSLT</td> <td>a. It is a language for form attingXML docum ents.</td> </tr> <tr> <td>2. XPath</td> <td>b. It is a language for transforming XML docum ents into various other types of docum ents</td> </tr> <tr> <td>3. XQuery</td> <td>c. It is a language for navigating in XML docum ents.</td> </tr> <tr> <td>4. XSL-FO</td> <td>d. It is a language for queryingXML docum ents.</td> </tr> </table>	1. XSLT	a. It is a language for form attingXML docum ents.	2. XPath	b. It is a language for transforming XML docum ents into various other types of docum ents	3. XQuery	c. It is a language for navigating in XML docum ents.	4. XSL-FO	d. It is a language for queryingXML docum ents.
1. XSLT	a. It is a language for form attingXML docum ents.								
2. XPath	b. It is a language for transforming XML docum ents into various other types of docum ents								
3. XQuery	c. It is a language for navigating in XML docum ents.								
4. XSL-FO	d. It is a language for queryingXML docum ents.								
Option A:	1-d, 2-c, 3-b, 4-a								
Option B:	1-c, 2-d, 3-b, 4-a								
Option C:	1-b, 2-d, 3-a, 4-c								
Option D:	1-b, 2-c, 3-d, 4-a								
9.	What will be the output of the following PHP program? <pre>&lt;?php \$month = array("Jan", "Feb", "Mar",array("Apr","May")); echo (count(\$month, 1)); ?&gt;</pre>								
Option A:	5								
Option B:	4								
Option C:	3								
Option D:	6								
10.	The XSLT Processor accepts which files for processing								
Option A:	DHTML document, XSL Document								
Option B:	HTML document, DOM document								
Option C:	XML document, XSL Document								
Option D:	HTML document, Javascript document								

<b>Q2</b>	<b>Solve any Four out of Six 5 marks each</b>
A	List and explain various CSS Transition properties
B	What is the use of XMLHttpRequest object? Explain methods that are used to send request to server using AJAX.
C	Write HTML5 code to drag a specific image and drop it on a specific location
D	Explain with the help of an example how exceptions are handled in JavaScript
E	Differentiate between HTML and XML
F	Explain the following elements of XSL with examples i. <xsl:value-of> ii. <xsl:for-each>

<b>Q3</b>	<b>Solve any Two Questions out of Three 10 marks each</b>
1	What is validation. Create a form that has following fields, "Username", "Password" and "Confirm Password". Using JavaScript validate each field as follows: a) Username should of minimum 10 characters b) Password should contain 1 upper case letter, 1 lower case letter, 1 digit and 1 special character and the length of the password should be minimum 8. c) Confirm password should match password entered.
2	What are Servlets? Demonstrate the Servlet architecture and explain its working principle.
3	Create a well-formed XML document containing details of a car like: id, company_name, model, engine and mileage. Format this information in the tabular manner using XSLT

<b>Q4.</b>	
A	<b>Solve any Two 5 marks each</b>
i.	What is the purpose of implicit arrays \$_POST and \$_GET in PHP? Consider that a webpage displays a form containing two text boxes named n1 and n2, where user enters numeric data. Write a PHP script which collects this form data and finds greatest of two numbers and displays the same.
ii.	What is session tracking? Show how session tracking is achieved using cookies
iii.	Draw the diagram of AJAX application model and traditional application web model and compare them
B	<b>Solve any One 10 marks each</b>
i.	Create an HTML form that accepts Emp_Id, First_Name, Last_Name and Gender from user. Write a PHP code to store this information into Employee table using MySQL database
ii.	Write HTML5 code for embedding audio and video elements in the web page.

6. Please Upload complete scanned answer copy in a single PDF file. \*

Files submitted:

7. Have you uploaded correct scanned copy of the answer sheets. \*

*Mark only one oval.*

YES

---

This content is neither created nor endorsed by Google.

Google Forms

# UOM Exam Second half 2021\_ Question paper\_R2019/Comp/CSC501 - THEORETICAL COMPUTER SCIENCE/Sem-V

Dear Student,

Please note before you attempt this section of examination:

1. Q1, Q2, Q3 and Q4 carry 20 marks each.
2. This paper contains 20 Marks MCQ and 60 marks subjective section for 150 minutes duration.
3. It is mandatory for all the students to upload their answer papers in a single PDF format only.
4. You have to write Date of Examination, Seat number, Program, Scheme and semester, Subject name, Signature on EVERY PAGE.
5. Remain in the meet with your camera on and you in clear view throughout the duration of the exam.

---

\* Required

1. Email \*

---

2. Student Name (As per exam form filled) \*

---

3. Seat No \*

Refer Hall ticket

---



4. Class \*

*Mark only one oval.*

TE3

TE4

TE9

5. Roll No. \*

---

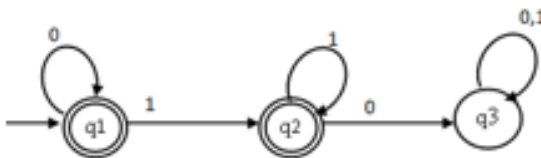
Solve Questions as per the instructions given separately.

- Please upload a single PDF for Q1 to Q4
- For MCQs Question write Question number & correct option with complete text in option.
- Q2 to Q4 are subjective questions - Solve Questions as per the instructions and marks allotted.

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	The no. of minimum state for divisibility by four tester FSM for binary numbers
Option A:	Three
Option B:	Four
Option C:	Five
Option D:	Six
2.	In Moore Machine if the length of I/p sequence is n, then length of o/p sequence is
Option A:	$n^2$
Option B:	$2n$
Option C:	$n+1$
Option D:	$N$
3.	If P, Q, R be the three regular expressions then the equation $R = Q + RP$ has a unique solution given by
Option A:	$R = QP^*$
Option B:	$R = P^*R$
Option C:	$R = RP$
Option D:	$R = Q^*P$
4.	Consider the following statements. I. If $L \cup M$ is regular, then both L and M must be regular. II. The class of regular languages is closed under infinite union.
Option A:	I only
Option B:	II only
Option C:	Both I and II
Option D:	Neither I nor II
5.	The production of the form $A \rightarrow BC$ or $A \rightarrow a$ , where A, B & C are non-terminals and a is terminal. This type of grammar is,
Option A:	Unit production
Option B:	Chomsky Normal Form
Option C:	Null production
Option D:	Greibach Normal Form
6.	_____ is Type 2 grammar according to Chomsky Hierarchy.
Option A:	Regular Grammar
Option B:	Context Sensitive Grammar
Option C:	Context Free Grammar
Option D:	Unrestricted Grammar

7.	A PDA machine configuration $(q, a, X)$ can be correctly represented as ,
Option A:	(Current state, unprocessed input, stack content)
Option B:	(Current state, stack content, unprocessed input ,)
Option C:	(unprocessed input ,current state, ,stack content)
Option D:	(stack content, current state, unprocessed input)
8.	----- can be used to simulate any Turing machine
Option A:	Finite state machine
Option B:	Push down automata
Option C:	Universal Turing machine
Option D:	Counter machines
9.	Recursively enumerable language can be recognized by a,
Option A:	Finite state machine
Option B:	Deterministic push down automata
Option C:	Turing machine
Option D:	Linear bounded automata
10.	Consider the following statement I. Each Turing Acceptable language is need not to be Turing Decidable II. Every Turing Decidable language is Turing acceptable
Option A:	only I is true
Option B:	only II is true
Option C:	Both are False
Option D:	Both are True

<b>Q2</b>	
<b>A</b>	<b>Solve any Two 5 marks each</b>
i.	Write Short note on :Universal Turing machine
ii.	Convert following context free grammar to equivalent CNF $S \rightarrow ABA$ $A \rightarrow aA   \epsilon$ $B \rightarrow bB   \epsilon$
iii.	Prove that $L = \{ WW^R \mid W \in (0, 1)^* \}$ is not regular where $w^R$ is reverse of $w$ using pumping lemma.
<b>B</b>	<b>Solve any One 10 marks each</b>
i.	Design Turing machine for the language $L = \{ a^n b^{n+1} \mid n > 0 \}$
ii.	Design a DFA for language of string with even number of 0's and odd number of 1's over $\Sigma = \{ 0, 1 \}$

<b>Q3</b>	
<b>A</b>	<b>Solve any Two</b> <span style="float: right;"><b>5 marks each</b></span>
i.	Find the solution for the following instances of PCP. $x = \{1, 10111, 10\}$ $y = \{111, 10, 0\}$ have a solution? Justify your answer
ii.	Use Arden's Theorem to find the regular expression represented by the following DFA: 
iii.	Design PDA for language $\{(a^n b^m a^n \mid m, n \geq 1)\}$
<b>B</b>	<b>Solve any One</b> <span style="float: right;"><b>10 marks each</b></span>
i.	Convert the following RE to $\epsilon$ -NFA & then convert it to DFA: $R = ((0+1)^* 10 + (00)^* (11)^*)^*$
ii.	Design Moore machine to convert each occurrence of "1000" to "1001".

<b>Q4</b>	
<b>A</b>	<b>Solve any Two</b> <span style="float: right;"><b>5 marks each</b></span>
i.	Write Short Note on Chomsky Hierarchy
ii.	Differentiate between DFA and NFA
iii.	Design PDA for $L = \{ x \mid n_a(x) = n_b(x) \}$
<b>B</b>	<b>Solve any One</b> <span style="float: right;"><b>10 marks each</b></span>
i.	Consider the following Grammar $S \rightarrow SAS \mid b$ $A \rightarrow ba \mid b$ For the string 'bbabbbbab' find the following LMD, RMD, parse tree and check whether the grammar is ambiguous or not.
ii.	Design Turing machine to perform addition of two unary numbers. i.e. $m+n$ I/P = $B0^m \# 0^n B$ O/P = $B0^p B$ where $p = m+n$

6. Please Upload complete scanned answer copy in a single PDF file. \*

Files submitted:

7. Have you uploaded correct scanned copy of the answer sheets. \*

*Mark only one oval.*

YES

---

This content is neither created nor endorsed by Google.

**Google Forms**

# UOM Exam Second half 2021\_ Question paper\_R2019/Comp/CSDLO5011 - Probabilistic Graphical Models/Sem-V

Dear Student,

Please note before you attempt this section of examination:

1. Q1, Q2, Q3 and Q4 carry 20 marks each.
2. This paper contains 20 Marks MCQ and 60 marks subjective section for 150 minutes duration.
3. It is mandatory for all the students to upload their answer papers in a single PDF format only.
4. You have to write Date of Examination, Seat number, Program, Scheme and semester, Subject name, Signature on EVERY PAGE.
5. Remain in the meet with your camera on and you in clear view throughout the duration of the exam.

---

\* Required

1. Email \*

---

2. Student Name (As per exam form filled) \*

---

3. Seat No \*

Refer Hall ticket

---

4. Class \*

*Mark only one oval.*

TE3

TE4


TE9

5. Roll Number: \*

Solve Questions as per the instructions given separately.

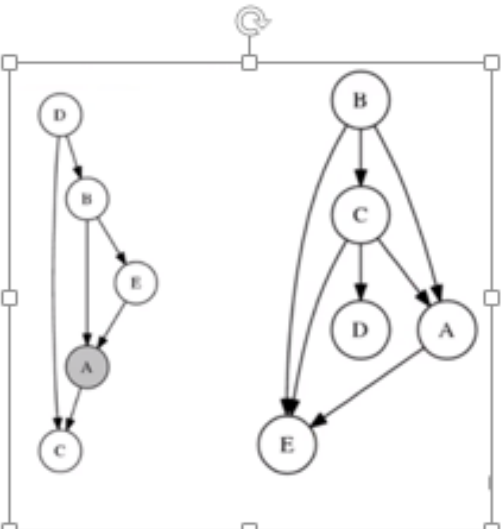
- Please upload a single PDF for Q1 to Q4
- For MCQs Question write Question number & correct option with complete text in option.
- Q2 to Q4 are subjective questions - Solve Questions as per the instructions and marks allotted.

Page 1/5

<b>Q.1</b>	<b>Choose the correct option for following questions. All the Questions are compulsory and carry equal marks</b>
1.	Which algorithm is used for solving temporal probabilistic reasoning?
Option A:	Hill Climbing Algorithm
Option B:	Hidden Markov Model
Option C:	Depth-first search
Option D:	Breadth-first search
2.	Functional dependencies that holds at a particular point in time are known as
Option A:	Interval functional dependency
Option B:	Temporal functional dependency
Option C:	Snapshot functional dependency
Option D:	Pointed functional dependency
3.	<p>What is the Markov Blanket of variable <math>S_3</math>.</p>  <pre> graph TD     D1((D1)) --&gt; S1[S1]     D1((D1)) --&gt; S2[S2]     D1((D1)) --&gt; S3[S3]     D2((D2)) --&gt; S2[S2]     D2((D2)) --&gt; S3[S3]     S1[S1] --&gt; S3[S3]     S2[S2] --&gt; S3[S3]     D1((D1)) --&gt; D2((D2))     </pre>

Option A:	D1
Option B:	D2
Option C:	D1 and D2
Option D:	S2
4.	Find the incorrect statement. Bayesian (BN) versus Markov Network (MN)
Option A:	In BN, we use conditional probability as factors. In MN also, we use conditional probability.
Option B:	In MN, we want to capture the affinity by a real number. In BN the factors are probability between 0 and 1.
Option C:	MN is restricted to discrete state space while BN can be both discrete and continuous.
Option D:	Unlike BN which have directed edges and clear directions of causality, MN have undirected edges and only encode associations.
5.	Which of these is not a goal of learning
Option A:	Density estimation
Option B:	Specific Prediction tasks
Option C:	Knowledge Discovery
Option D:	Introducing bias
6.	For n random binary variables the joint distribution requires _____ parameters.
Option A:	$2n - 1$
Option B:	$n - 1$
Option C:	$2^n - 1$
Option D:	$2^n$
7.	If the various probabilities are given as: $P(B1)=P(B2)=P(B3)=P(B4)=1/4$ and $P(D/B1)=0.05$ , $P(D/B2)=0.4$ , $P(D/B3)=0.1$ , $P(D/B4)=0.1$ . Find $P(B2/D)$ .
Option A:	$13/80$
Option B:	$13/8$
Option C:	$8/13$
Option D:	0



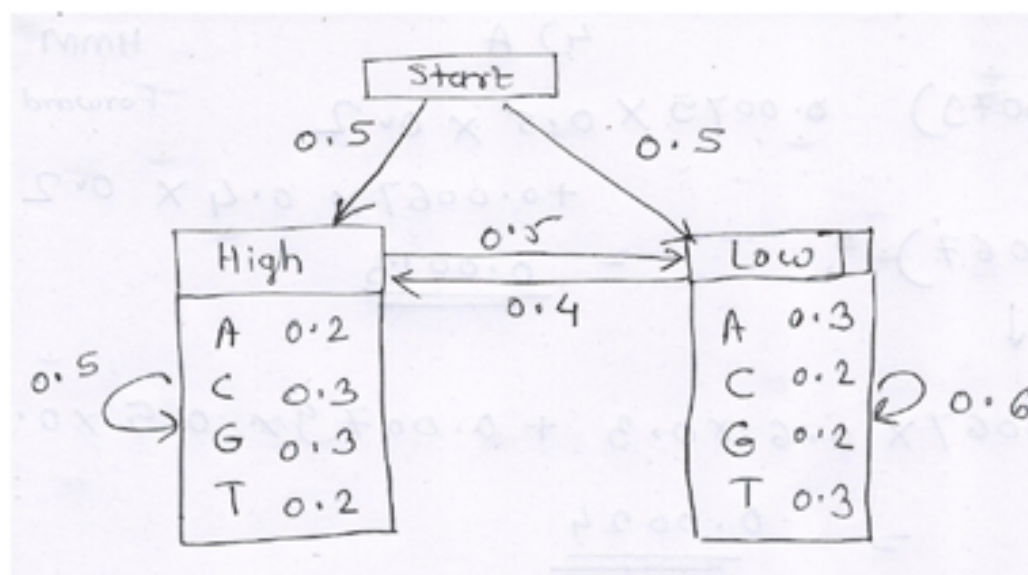
8.	<p>In Fig 1 is C independent of E if A is observed? In Fig 2 is D independent of A?</p> 
Option A:	Fig 1 & 2 – No
Option B:	Fig 1 - Yes, Fig 2 – No
Option C:	Fig 1 - No, Fig 2 – Yes
Option D:	Fig 1 & 2 – Yes
9.	Which of the options is not a type of local probabilistic model?
Option A:	Deterministic CPDs
Option B:	Context-Specific CPDs
Option C:	Tabular CPDs
Option D:	Generalized quadratic Models
10.	A decision-making situation D is defined by which of the following elements:
Option A:	A set of possible outcomes and a set of possible actions
Option B:	A probabilistic outcome model
Option C:	An utility function
Option D:	All of the above

<b>Q.2</b>	<b>Solve any Four out of Six [5 marks each]</b>
A	Differentiate between marginal and joint distributions with an example.
B	Explain Causal and Evidential Reasoning patterns with example.
C	Explain factor graph in HMM with the help of an example.
D	What is a temporal setting with respect to FGM? Explain with the help of an example.
E	Explain speech recognition as an application HMM.
F	Explain Expected Log Likelihood metric.

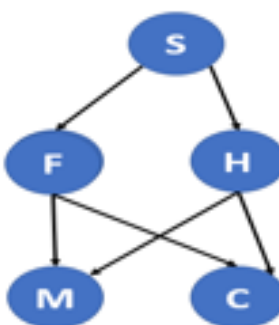
**Q.3 Solve any Two Questions out of Three [10 marks each]**

A  $S = G G C A C T$

Decode the above sequence using Hidden Markov Model.



B Differentiate between conditional dependence and conditional independence. List out 5 independencies in the graph given below:



C	Given the following factor table between variables A, B and C, Find the joint distribution function																													
	<table border="1"> <thead> <tr> <th>A</th> <th>B</th> <th>Phi</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>0</td> <td>20</td> </tr> <tr> <td>0</td> <td>1</td> <td>5</td> </tr> <tr> <td>1</td> <td>0</td> <td>3</td> </tr> <tr> <td>1</td> <td>1</td> <td>12</td> </tr> </tbody> </table> <table border="1"> <thead> <tr> <th>B</th> <th>C</th> <th>Phi</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>0</td> <td>40</td> </tr> <tr> <td>0</td> <td>1</td> <td>15</td> </tr> <tr> <td>1</td> <td>0</td> <td>13</td> </tr> <tr> <td>1</td> <td>1</td> <td>50</td> </tr> </tbody> </table>	A	B	Phi	0	0	20	0	1	5	1	0	3	1	1	12	B	C	Phi	0	0	40	0	1	15	1	0	13	1	1
A	B	Phi																												
0	0	20																												
0	1	5																												
1	0	3																												
1	1	12																												
B	C	Phi																												
0	0	40																												
0	1	15																												
1	0	13																												
1	1	50																												

<b>Q. 4</b>	<b>Solve any Four out of Six [5 marks each]</b>
A	Write a short note on Conditioning a Markov Network Model.
B	Explain confounding variables with the help an example?
C	Explain structured decision problems with the help of decision tree.
D	Explain the concept of LogLinear Parameterization with the help of an example.
E	Explain any one application of Bayesian Network with respect to PGM.
F	Explain variable elimination with the help of an example.

6. Please Upload complete scanned answer copy in a single PDF file. \*

Files submitted:

7. Have you uploaded correct scanned copy of the answer sheets. \*

*Mark only one oval.*

YES

---

This content is neither created nor endorsed by Google.

**Google Forms**

# UOM Exam Second half 2021\_ Question paper\_R2019/Comp/CSC502 - Software Engineering/Sem-V

Dear Student,

Please note before you attempt this section of examination:

1. Q1, Q2, Q3 and Q4 carry 20 marks each.
2. This paper contains 20 Marks MCQ and 60 marks subjective section for 150 minutes duration.
3. It is mandatory for all the students to upload their answer papers in a single PDF format only.
4. You have to write Date of Examination, Seat number, Program, Scheme and semester, Subject name, Signature on EVERY PAGE.
5. Remain in the meet with your camera on and you in clear view throughout the duration of the exam.

---

**\* Required**

1. Email \*

---

2. Student Name (As per exam form filled) \*

---

3. Seat No \*

Refer Hall ticket

---

4. Class \*

*Mark only one oval.*

TE3

TE4

TE9

5. Roll No. \*

---

Solve Questions as per the instructions given separately.

- Please upload a single PDF for Q1 to Q4
- For MCQs Question write Question number & correct option with complete text in option.
- Q2 to Q4 are subjective questions - Solve Questions as per the instructions and marks allotted.

<b>Q1.</b>	<b>Choose the correct option for following questions. All the Questions are compulsory and carry equal marks</b>
1.	The incremental model of software development is
Option A:	A reasonable approach when requirements are well defined.
Option B:	A good approach when a working core product is required quickly.
Option C:	The best approach to use for projects with large development team.
Option D:	A revolutionary model that is not used for commercial products.
2.	Which of the following testing is the part of non-functional testing?
Option A:	Unit testing
Option B:	Performance testing
Option C:	System testing
Option D:	Integration testing
3.	Software Testing with real data in real environment is known as
Option A:	alpha testing
Option B:	beta testing
Option C:	regression testing
Option D:	retesting
4.	Which of the following risk is the failure of a purchased component to perform as expected?
Option A:	Product risk
Option B:	Project risk
Option C:	Business risk
Option D:	Programming risk
5.	LOC based estimation techniques require problem decomposition based on
Option A:	Information domain values
Option B:	Project schedule
Option C:	Size of Software functions
Option D:	Process activities
6.	In scrum the team activity is monitored and coordinated on ..... basis.
Option A:	Hourly
Option B:	Daily
Option C:	Monthly
Option D:	Weekly
7.	Which of the following framework activities are found in the Extreme Programming?
Option A:	Planning, Analysis, Design, Coding
Option B:	Analysis, Design, Coding, Testing
Option C:	Planning, Design, Coding, Testing
Option D:	Design, Analysis, Coding, Testing

8.	Cohesion is a qualitative indication to the degree to which a module
Option A:	Can be written more compactly.
Option B:	Focuses on just one thing.
Option C:	Is able to complete its function in a timely manner.
Option D:	Is connected to other modules and the outside world.
9.	Which of the following is the process of assembling program components, data, and libraries, and then compiling and linking these to create an executable system?
Option A:	System building
Option B:	Release management
Option C:	Change management
Option D:	Version management
10.	Which of the following is not a SQA plan for a project?
Option A:	Evaluations to be performed
Option B:	Amount of technical work
Option C:	Audits and reviews to be performed
Option D:	Documents to be produced by the SQA group

<b>Q2</b>	<b>Solve any Four out of Six.</b>	<b>5 markseach.</b>
A	Explain spiral process model.	
B	What are the advantages of Agile process.	
C	Differentiate between White Box testing and Black Box testing.	
D	Explain – Reverse Engineering .	
E	Explain scenario based model?	
F	Describe Boundaryvalue testing with an suitable example.	

<b>Q3</b>	<b>Solve any Two Questions out of Three.</b>	<b>10 marks each.</b>
A	Explain risk and its types? Explain the steps involved in setting up or generating RMMIM plan.	
B	Differentiate between FP based & LOC based cost estimation techniques.	
C	Develop the SRS for University Management System.	

<b>Q4.</b>		
<b>A</b>	<b>Solve any Two.</b>	<b>5 marks each.</b>
i.	Explain Integration testing .	
ii.	Distinguish between cohesion and coupling.	
iii.	What is cost estimation? Explain LOC method.	
<b>B</b>	<b>Solve any One.</b>	<b>10 marks each.</b>
i.	What is FTR? Explain the review guidelines considered during FTR.	
ii.	Explain the process of SCM.	

6. Please Upload complete scanned answer copy in a single PDF file. \*

Files submitted:



7. Have you uploaded correct scanned copy of the answer sheets. \*

*Mark only one oval.*

YES

---

This content is neither created nor endorsed by Google.

Google Forms

**University of Mumbai**  
**Examination Second Half 2021**

**Examinations Commencing from 22<sup>nd</sup> November 2021 to 5<sup>th</sup> January 2022**

Program: COMPUTER ENGINEERING

Curriculum Scheme: Rev2019

Examination: TE Semester V

Course Code: CSC501 and Course Name: THEORETICAL COMPUTER SCIENCE

Time: 2 hour 30 minutes

Max. Marks: 80

---

---

Question Number	Correct Option (Enter either 'A' or 'B' or 'C' or 'D')
Q1.	B
Q2.	C
Q3.	A
Q4	D
Q5	B
Q6	C
Q7	A
Q8.	C
Q9.	C
Q10.	D

Descriptive Q2 , Q3 and Q4

Ans : Solution / Expected points  
/Distribution of Marks.

**University of Mumbai**

**Examination Second Half 2021 under cluster \_\_ (Lead College: \_\_\_\_\_)**

**Examinations Commencing from 22<sup>nd</sup> November 2021 to 5<sup>th</sup> January 2022**

**Program: Computer Engineering**

**Curriculum Scheme: Rev2019**

**Examination: TE Semester V**

**Course Code: CSC502 and Course Name: Software Engineering**

**Time: 2 hour 30 minutes**

**Max. Marks: 80**

---

---

<b>Question Number</b>	<b>Correct Option (Enter either 'A' or 'B' or 'C' or 'D')</b>
Q1.	<b>B</b>
Q2.	<b>B</b>
Q3.	<b>B</b>
Q4	<b>A</b>
Q5	<b>C</b>
Q6	<b>B</b>
Q7	<b>C</b>
Q8.	<b>B</b>
Q9.	<b>A</b>
Q10.	<b>B</b>

**University of Mumbai**  
**Examination Second Half 2021**

Examinations Commencing from 22<sup>nd</sup> November 2021 to 5<sup>th</sup> January 2022

Program: Computer Engineering

Curriculum Scheme: Rev2019

Examination: TE Semester V

Course Code: CSDLO5011 and Course Name: Probabilistic Graphical Models

Time: 2 hour 30 minutes

Max. Marks: 80

=====

Question Number	Correct Option (Enter either 'A' or 'B' or 'C' or 'D')
Q1.	B
Q2.	B
Q3.	B
Q4	A
Q5	D
Q6	C
Q7	C
Q8.	A
Q9.	D
Q10.	D

**University of Mumbai**  
**Examination Second Half 2021**

Examinations Commencing from 22<sup>nd</sup> November 2021 to 5<sup>th</sup> January 2022

Program: **Computer Engineering**

Curriculum Scheme: Rev2019

Examination: TE Semester V

Course Code: CSDLO 5012 and Course Name: Internet Programming

Time: 2 hour 30 minutes

Max. Marks: 80

---

---

Question Number	Correct Option (Enter either 'A' or 'B' or 'C' or 'D')
Q1.	D
Q2.	C
Q3.	A
Q4	A
Q5	B
Q6	B
Q7	C
Q8.	D
Q9.	D
Q10.	C

**Q2.**

Ans A. Each property -- 1 mark

Ans B. Use XMLHttpRequest object -- 2mark

List all methods – 1 marks

Explain methods – 2 marks

Ans C. Correct code for drag using API –2 marks

Defining destination in code – 1 marks

Correct code for drop using API –2 marks

Ans D. listing all exception handling statements-1 mark

Explain each method - 1 mark each

( 4 methods – each method 1 mark)

Ans E. Any 5 points – Each correct point – 1mark

Ans F. <xsl : value- of> - 2 mark ,1 mark example

<xsl : for -each> - 2 mark, 1 mark example

**Q4.**

Ans A i. \$\_GET – 1 mark

\$\_POST – 1 mark

Program – 3 marks

A ii. Session tracking concept – 2 marks

Session tracking using cookies-3 marks

A iii. AJAX application model – 1 mark

Traditional application model – 1 mark

Any 3 valid points – 3 marks

Ans B i – HTML form – 5 marks

PHP Code – 5 marks

Ans Bii. Audio clip with all possibility – 5 marks

Video clip with all possibility – 5 marks

**Q3.**

**Ans 1:** Validation explanation – 2 marks

HTML form – 2 marks

Username validation – 2 marks

Password validation – 2 marks

Confirm Password – 2 marks

**Ans 2:** Servlets explanation – 2 marks

Listing of Components & explanation – 3 marks

Architecture diagram & explanation – 3 marks

**Ans 3:**

Valid XML Document – 5 marks

Correct XSLT - 5 marks

**University of Mumbai**

**Examination Second Half 2021 under cluster \_\_ (Lead College: \_\_\_\_\_)**

**Examinations Commencing from 22<sup>nd</sup> November 2021 to 5<sup>th</sup> January 2022**

Program: **Computer Engineering** \_\_\_\_\_

Curriculum Scheme: Rev2019

Examination: TE Semester: V

Course Code: **CSC504** and Course Name: **Data Warehousing and Mining**

Time: 2 hour 30 minutes

Max. Marks: 80

=====

<b>Question Number</b>	<b>Correct Option (Enter either 'A' or 'B' or 'C' or 'D')</b>
Q1.	A
Q2.	A
Q3.	B
Q4	D
Q5	C
Q6	A
Q7	D
Q8.	B
Q9.	B
Q10.	B

Question 2	Solve any Two out of Three	10 marks each
A	$P(\text{short}   t) = (P(t \text{short}) * p(\text{short})) / P(t) = (0 * 4 / 9) / 0.11 = 0$ $P(\text{medium} t) = (0 * 2/9) / 0.11 = 0$ $P(\text{Tall} t) = (0.11 * 3/9) / 0.11 = 0.33$ <p>New tuple is a Tall as it has the highest probability.</p>	
B	<p>Fig Star Schema</p>	
C	<p>the steps for the apriori algorithm:</p> <p>Step-1: Determine the support of itemsets in the transactional database, and select the minimum support and confidence.</p> <p>Step-2: Take all supports in the transaction with higher support value than the minimum or selected support value.</p>	



Step-3: Find all the rules of these subsets that have higher confidence value than the threshold or minimum confidence.

Step-4: Sort the rules as the decreasing order of lift.

For the following dataset that has various transactions, we need to find the frequent itemsets and generate the association rules using the Apriori algorithm:

Step-1: Calculating C1 and L1:

- In the first step, we will create a table that contains support count (The frequency of each itemset individually in the dataset) of each itemset in the given dataset. This table is called the Candidate set or C1.

Item Set	Support Count
1	4
2	5
3	3
4	2
5	2

Now, we will take out all the itemsets that have the greater support count than the Minimum Support (2). It will give us the table for the frequent itemset L1.

Since all the itemsets have greater or equal support count than the minimum support,

Item Set	Support Count
1	4
2	5
3	3
4	2
5	2

Step-2: Candidate Generation C2, and L2:

1. In this step, we will generate C2 with the help of L1. In C2, we will

create the pair of the itemsets of L1 in the form of subsets.

2. After creating the subsets, we will again find the support count from the main transaction table of datasets, i.e., how many times these pairs have occurred together in the given dataset. So, we will get the below table for C2:

Item Set	Support Count
1,2	3
1,3	2
1,4	1
1,5	2
2,3	2
2,4	2
2,5	2
3,4	-
3,5	1
4,5	-

Step-3: Candidate generation C3, and L3:

Again, we need to compare the C2 Support count with the minimum support count, and after comparing, the itemset with less support count will be eliminated from the table C2. Similarly perform further combinations

Item Set	Support Count
1,2,3	1
1,2,4	1
1,2,5	2
2,3,4	-
2,3,5	1
3,4,5	-
4,5,1	-

	4,5,2	-
<p>Step-4: Finding the association rules for the subsets:</p> <p>To generate the association rules, first, we will create a new table with the possible rules from the occurred combination {A, B.C}. For all the rules, we will calculate the Confidence using formula <math>\frac{\text{sup}(A \wedge B)}{\text{sup}(A)}</math>. After calculating the confidence value for all rules, we will exclude the rules that have less confidence than the minimum threshold(70%).</p>		

Question 3	Solve any Two out of Three	10 marks each
A	Final Centroid $C_1 = (1.83, 2.33)$ , $C_2 = (4.12, 5.37)$ final clusters (1,2) and (3,4,5,6,7)	
B		
C	<p>The Web log contains the following information :</p> <ol style="list-style-type: none"> <li>1. The user's IP address,</li> <li>2. The user's authentication name,</li> <li>3. The date-time stamp of the access,</li> <li>4. The HTTP request,</li> <li>5. The response status,</li> <li>6. The size of the requested resource, and optionally a lso,</li> <li>7. The referrer URL (the page the user "came from"),</li> <li>8. The user's browser identification.</li> </ol>	

**Question 4**

**Solve any Four Questions out of Five**

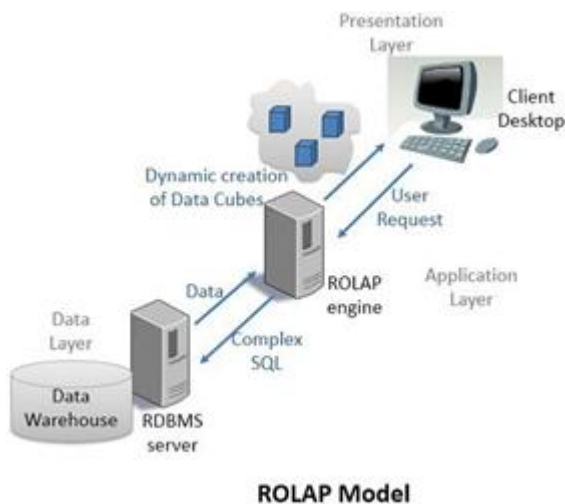
**5 marks each**

A

	<b>OLTP System Online Transaction Processing (Operational System)</b>	<b>OLAP System Online Analytical Processing (Data Warehouse)</b>
Source of data	Operational data; OLTPs are the original source of the data.	Consolidation data; OLAP data comes from the various OLTP Databases
Purpose of data	To control and run fundamental business tasks	To help with planning, problem solving, and decision support
What the data	Reveals a snapshot of ongoing business processes	Multi-dimensional views of various kinds of business activities
Inserts and Updates	Short and fast inserts and updates initiated by end users	Periodic long-running batch jobs refresh the data
Queries	Relatively standardized and simple queries Returning relatively few records	Often complex queries involving aggregations
Processing Speed	Typically very fast	Depends on the amount of data involved; batch data refreshes and complex queries may take many hours; query speed can be improved by creating indexes
Space Requirements	Can be relatively small if historical data is archived	Larger due to the existence of aggregation structures and history data; requires more indexes than OLTP
Database Design	Highly normalized with many tables	Typically de-normalized with fewer tables; use of star and/or snowflake schemas
Backup and Recovery	Backup religiously; operational data is critical to run the business, data loss is likely to entail significant monetary loss and legal liability	Instead of regular backups, some environments may consider simply reloading the OLTP data as a recovery method

B

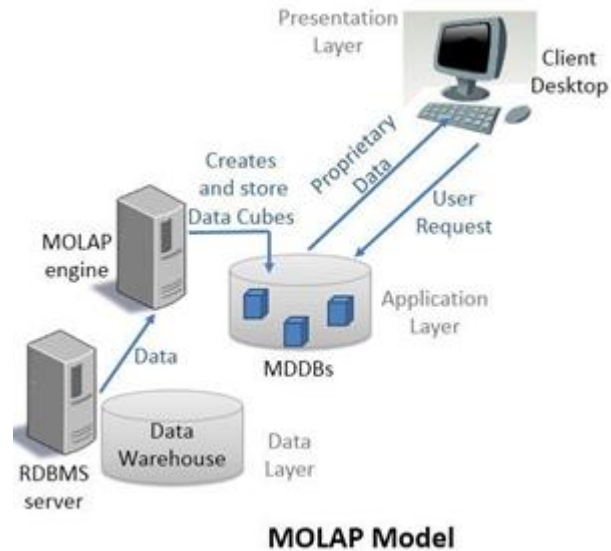
ROLAP is Relational Online Analytical Processing model, where the data is stored as in relational database i.e. rows and columns in the data warehouse. In the ROLAP model data is present in the front of the user in the multidimensional form. To display the data, in a multidimensional view, a semantic layer of metadata is created that maps dimension to the relational tables. Metadata also supports aggregation of the data.



MOLAP is a Multidimensional Online Analytical Processing model. The data used for analysis is stored in specialized multidimensional databases (MDDBs). The multidimensional database management systems are

proprietary software systems.

These multidimensional databases are formed from the large multidimensional array. The cells or data cubes of this multidimensional databases carry precalculated and prefabricated data. Proprietary software systems create this precalculated and fabricated data, while the data is loaded to MDDBs from the main databases.



### Structure of Web Log

The Web log contains the following information :

1. The user's IP address,
2. The user's authentication name,
3. The date-time stamp of the access,
4. The HTTP request,
5. The response status,
6. The size of the requested resource, and optionally a lso,
7. The referrer URL (the page the user "came from"),
8. The user's browser identification.

C

Web log fields

- IP
    - o 152.152.98.11
    - o IP address - can be converted to host name, such as xyz.example.com
  - Name
    - o The name of the remote user (usually omitted and replaced by a dash “-”)
  - Login
    - o Login of the remote user (also usually omitted and replaced by a dash “-”)
  - Date/Time/TZ
    - 16/Nov/2005:16:32:50 -0500
- Etc

**Fragment Example :**

```
111.111.111.111
-
-
[08/Oct/2007:11:17:55 -0400]
“GET / HTTP/1.1”
200
10801
“http://www.google.com/search?q=log+analyzer&ie=utf-8&oe=utf-8&aq=t&rls=org.mozilla:en-US:official&client=firefox-a”
“Mozilla/5.0 (Windows; U; Windows NT 5.2; en-US; rv:1.8.1.7) Gecko/20070914 Firefox/2.0.0.7”
```

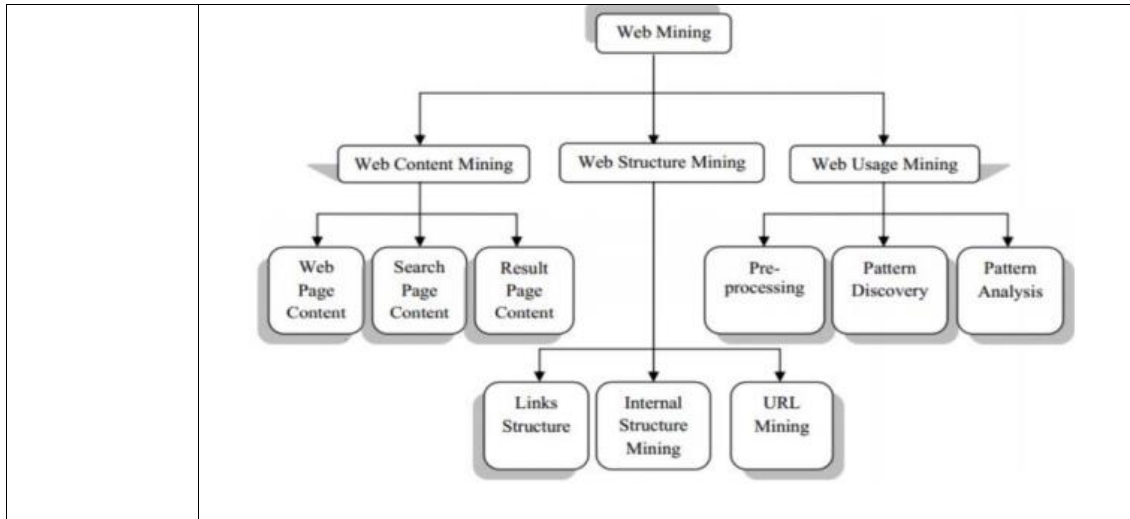
Give explanation of each field

D

Describe the steps involved in the FP growth algorithm . Scan DB once, find frequent 1-itemset (single item pattern)

- Sort frequent items in frequency descending order, f-list.
- Scan DB again, construct FP-tree.

	<ul style="list-style-type: none"> <li>● Construct the conditional FP tree in the sequence of reverse order of F - List - generate frequent item set.</li> </ul> <p>Compare and contrast its steps with Apriori algorithm</p> <table border="0" style="width: 100%;"> <thead> <tr> <th style="text-align: left; width: 50%;"><b>Apriori</b></th> <th style="text-align: left; width: 50%;"><b>FP Growth</b></th> </tr> </thead> <tbody> <tr> <td>Apriori generates the frequent patterns by making the itemsets using pairing such as single item set, double itemset, triple itemset.</td> <td>FP Growth generates an FP-Tree for making frequent patterns.</td> </tr> <tr> <td>Apriori uses candidate generation where frequent subsets are extended one item at a time.</td> <td>FP-growth generates conditional FP-Tree for every item in the data.</td> </tr> <tr> <td>Since apriori scans the database in each of its steps it becomes time-consuming for data where the number of items is larger.</td> <td>FP-tree requires only one scan of the database in its beginning steps so it consumes less time.</td> </tr> <tr> <td>A converted version of the database is saved in the memory</td> <td>Set of conditional FP-tree for every item is saved in the memory</td> </tr> <tr> <td>It uses breadth-first search</td> <td>It uses a depth-first search.</td> </tr> </tbody> </table>	<b>Apriori</b>	<b>FP Growth</b>	Apriori generates the frequent patterns by making the itemsets using pairing such as single item set, double itemset, triple itemset.	FP Growth generates an FP-Tree for making frequent patterns.	Apriori uses candidate generation where frequent subsets are extended one item at a time.	FP-growth generates conditional FP-Tree for every item in the data.	Since apriori scans the database in each of its steps it becomes time-consuming for data where the number of items is larger.	FP-tree requires only one scan of the database in its beginning steps so it consumes less time.	A converted version of the database is saved in the memory	Set of conditional FP-tree for every item is saved in the memory	It uses breadth-first search	It uses a depth-first search.
<b>Apriori</b>	<b>FP Growth</b>												
Apriori generates the frequent patterns by making the itemsets using pairing such as single item set, double itemset, triple itemset.	FP Growth generates an FP-Tree for making frequent patterns.												
Apriori uses candidate generation where frequent subsets are extended one item at a time.	FP-growth generates conditional FP-Tree for every item in the data.												
Since apriori scans the database in each of its steps it becomes time-consuming for data where the number of items is larger.	FP-tree requires only one scan of the database in its beginning steps so it consumes less time.												
A converted version of the database is saved in the memory	Set of conditional FP-tree for every item is saved in the memory												
It uses breadth-first search	It uses a depth-first search.												
E	taxonomy of web mining												





**University of Mumbai**  
**Examination Second Half 2021**

Examinations Commencing from 22<sup>nd</sup> November 2021 to 5<sup>th</sup> January 2022

Program: Computer Engineering

Curriculum Scheme: Rev2019

Examination: TE Semester V

Course Code: CSC503 and Course Name: Computer Network

Time: 2 hour 30 minutes

Max. Marks: 80

---

Question Number	Correct Option (Enter either 'A' or 'B' or 'C' or 'D')
Q1.	C
Q2.	B
Q3.	A
Q4	D
Q5	A
Q6	D
Q7	A
Q8.	B
Q9.	D
Q10.	B

Solution:

Q3.

A) iii) Class : Class B

Block : 132.21

Range : 132.21.0.0 to 132.21.255.255

Host address :  $2^{16} - 2 = 65536 - 2 = 65534$

Q4.

A) iii) CRC : 100

Transmitted frame : 110010100

Q5.

B) i) 190.100.0.0/16

1) 64 customers with 256 address each

190.100.0.0/24 to 190.100.63.0/24

2) 128 customers with 128 address each

190.100.64.0/25 to 192.100.127.128/24

3) 34 customers need 128 address each

190.100.128.0/26 to 190.100.159.192/26

**University of Mumbai**  
**Examination Second Half 2021**

Examinations Commencing from 22<sup>nd</sup> November 2021 to 5<sup>th</sup> January 2022

Program: Computer

Curriculum Scheme: Rev2019

Examination: TE Semester V

Course Code: CSDLO5013 and Course Name: Advance Database Management System

Time: 2 hour 30 minutes

Max. Marks: 80

---

---

<b>Question Number</b>	<b>Correct Option (Enter either 'A' or 'B' or 'C' or 'D')</b>
Q1.	D
Q2.	D
Q3.	B
Q4	C
Q5	A
Q6	A
Q7	C
Q8.	B
Q9.	D
Q10.	B