

K. J. Somaiya Institute of Engineering and Information Technology
Sion, Mumbai - 400022
NAAC Accredited Institute with 'A' Grade
NBA Accredited 3 Programs (Computer Engineering, Electronics & Telecommunication Engineering and Electronics Engineering) Permanently Affiliated to University of Mumbai

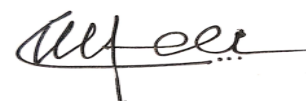
EXAMINATION TIME TABLE (June 2021)

T.E.(COMPUTER)(Sem VI) (REV. -2016) (Choice Based)

Days and Dates	Time	Course Code	Paper
Wednesday, June 02, 2021	11.30 a.m. to 1.30 p.m.	CSC601	Software Engineering
Friday, June 04, 2021	11.30 a.m. to 1.30 p.m.	CSC602	System Programming & Compiler Construction
Monday, June 07, 2021	11.30 a.m. to 1.30 p.m.	CSC603	Data Warehousing & Mining
Wednesday, June 09, 2021	11.30 a.m. to 1.30 p.m.	CSC604	Cryptography & System Security
Friday, June 11, 2021	11.30 a.m. to 1.30 p.m.	CSDLO6021	Department Level Optional Course -II:-Machine Learning
Friday, June 11, 2021	11.30 a.m. to 1.30 p.m.	CSDLO6022	Advance Database System
Friday, June 11, 2021	11.30 a.m. to 1.30 p.m.	CSDLO6023	Enterprise Resource Planning
Friday, June 11, 2021	11.30 a.m. to 1.30 p.m.	CSDLO6024	Advance Computer Network

Change if any, in the time table shall be communicated on the college web site.

Mumbai
Wednesday, May 12, 2021



Principal

University of Mumbai
Examination 2021 under cluster __ (Lead College: _____)

Examinations Commencing from 1st June 2021 to 14th June 2021

Program: **Computer Engineering**

Curriculum Scheme: Rev 2016

Examination: TE Semester VI

Course Code: **CSC601** and Course Name: **Software Engineering**

Time: 2 hours

Max. Marks: 80

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	COCOMO-II model is an example of :
Option A:	Risk Management
Option B:	Estimation Models
Option C:	Requirement Analysis
Option D:	software testing
2.	Empirical Estimations model are constructed on:
Option A:	Expert judgment based on past projects
Option B:	Regression models derived from historical project data
Option C:	Expected value estimation
Option D:	Trial and error parameter values
3.	Which of the following does not fall under project scheduling?
Option A:	Effort validation
Option B:	Market assessment
Option C:	Compartmentalization
Option D:	Time allocation
4.	Which of the following software process models couples the iterative nature of prototyping with the controlled and systematic factors of the linear sequential model?
Option A:	The Spiral Model.
Option B:	The Waterfall Model.
Option C:	The Incremental Model.
Option D:	The Revolutionary Model
5.	A Person is anyone within the company that has business interest in the product to be built and might be rewarded for the outcome or criticized if the attempt fails.
Option A:	Developer
Option B:	Stakeholder
Option C:	Coder
Option D:	Proprietor

6.	A technique for handling the introduction of products with an emphasis on chronic transparency and not overburdening the development team is -----
Option A:	Kanban
Option B:	Scrum
Option C:	Agile
Option D:	Development
7.	Which of the following is a useful measure for measuring the quality of a system?
Option A:	integrity, sales, usability, maintainability
Option B:	Stakeholders ,integrity, usability, sales
Option C:	correctness, usability, maintainability, integrity
Option D:	Correctness ,size ,usability ,maintainability
8.	The 3 P's in Project management are:
Option A:	Process, Performance and Product
Option B:	Process, Product and People
Option C:	Product, Performance and People
Option D:	People, Process and Performance
9.	In LOC Estimation techniques Problem decompositions are based on:
Option A:	project schedule
Option B:	process activities
Option C:	product specification
Option D:	software function
10.	SRS is said to be consistent if and only if
Option A:	its structure and style are such that any changes to the requirements can be made easily while retaining the style and structure
Option B:	every requirement stated therein is verifiable
Option C:	every requirement stated therein is one that the software shall meet
Option D:	no subset of individual requirements described in it conflict with each other
11.	What questions do black-box tests answer?
Option A:	Are all independent paths within a module exercised?
Option B:	Is the system particularly sensitive to certain input values?
Option C:	Does the internal structure to ensure their validity are exercised?
Option D:	Do all loops at their boundaries and within their operational bounds are exercised?
12.	In the Change control process, the change report is evaluated finally by whom?
Option A:	Software Developer
Option B:	Project Manager
Option C:	Software Configuration Manager
Option D:	Change Control authority

13.	Which design concept defines a direct outgrowth of modularity and the concepts of abstraction and information hiding?
Option A:	Refinement
Option B:	Architectural Patterns
Option C:	Functional Independence
Option D:	Refactoring
14.	The reverse engineering is concerned with
Option A:	Any adaptation of the system
Option B:	Any reconstruction of the system
Option C:	Any maintenance of the system
Option D:	Documentation change of the software
15.	Estimate the risk exposure, if the risk probability is given as 70%, 15 components need to be developed from scratch and the average component is 100 LOC with software engineering cost for each LOC is Rs.12.
Option A:	Rs.10,500
Option B:	Rs.18,000
Option C:	Rs.8,400
Option D:	Rs.12, 600
16.	Which one among the following provides the upper bound on the number of test cases that will be required to guarantee that every statement in the program has been executed at least once
Option A:	Cyclomatic Complexity
Option B:	Flowchart and flow graph
Option C:	Boundary value analysis
Option D:	Independent Program Paths
17.	Which of the following errors should not be tested when error handling is evaluated?
Option A:	Error description is impossible to understand
Option B:	Error noted does not correspond to error encountered
Option C:	Error condition causes system intervention
Option D:	Error description provide enough information to assist in the location of the cause of the error
18.	Which of the following is not a SQA plan for a project?
Option A:	evaluations to be performed
Option B:	duration of technical work
Option C:	audits and reviews to be performed
Option D:	procedures for error reporting and tracking
19.	Which of the following is not the golden rule for user interface design?
Option A:	Place the user in control

Option B:	Reduce the user's memory load
Option C:	Make the interface consistent
Option D:	Risk identification
20.	Independence of a module is measured using the following 2 qualitative criteria :
Option A:	Module and modularity
Option B:	Cyclomatic complexity and modularity
Option C:	Cohesion and coupling
Option D:	Abstraction and function point

Q2.	Solve any Two Questions out of Three	10 marks each
A	Differentiate between Spiral and Agile process models. Explain which process model is appropriate for developing any Mobile application.	
B	Explain the SCM Process. Differentiate between Quality Assurance and Quality control	
C	Describe the various testing strategies for a conventional system. Also discuss the different testing methods applicable for Web application.	

Q3.	Solve any Two Questions out of Three	10 marks each
A	Is Risk can be quantified? Justify your answer. How to practice risk management? Explain in detail.	
B	Explain COCOMO II Model with a suitable example. A project size of 200 KLOC is to be developed. Software development team has average experience on similar types of projects. The project schedule is not very tight. Calculate the Effort, development time, average staff size, and productivity for the project.	
C	Describe verification and validation with example. What comes first? Justify	

University of Mumbai
Examination 2021 under cluster __ (Lead College: _____)

Examinations Commencing from 1st June 2021 to 14th June 2021

Program: **Computer Engineering**

Curriculum Scheme: Rev 2016

Examination: TE Semester VI

Course Code: **CSC601** and Course Name: **Software Engineering**

Time: 2 hours

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	B
Q6	A
Q7	C
Q8.	B
Q9.	D
Q10.	D
Q11.	B
Q12.	D
Q13.	C
Q14.	B
Q15.	D
Q16.	A
Q17.	A
Q18.	B
Q19.	D
Q20.	C

Option 2 - Subjective Type

Q.2

(A) Agile Vs Spiral Process Model (5 Marks)

Agile model

The main principle of the Agile model is to achieve agility by removing unnecessary activities that waste time and effort.

The Agile model focuses on the delivery of an increment to the customer after each Time-box, so customer interaction is more frequent.

Agile model is suitable for large projects that are easy to divide into small parts that can be easily developed incrementally over each iteration.

Agile model does not rely on documentation.

Spiral model

The main principle of the Spiral model is risk handling.

Spiral model mainly deals with various kinds of unanticipated risks but customer interaction is less.

The Spiral model is suitable for those projects that are prone to various kinds of risks that are difficult to anticipate at the beginning of the project.

Proper documentation is required for Spiral model.

The Agile model is best suitable for mobile applications because it follows a combination of iterative and incremental approach which helps the project to adapt to changes rapidly.

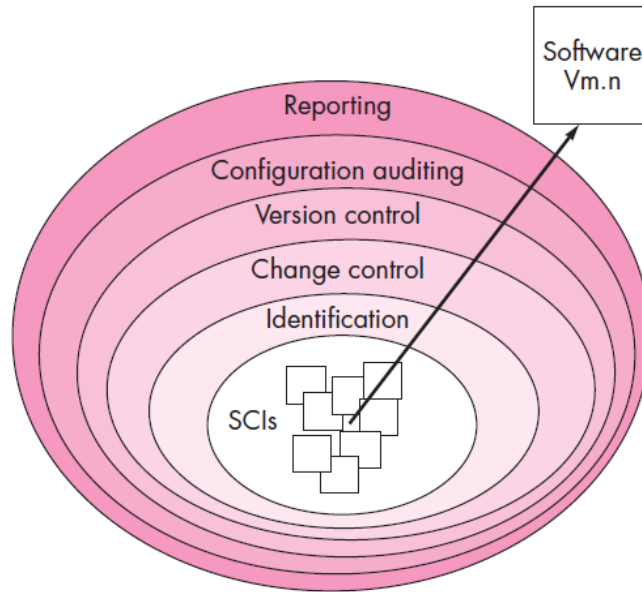
Need explanation for justification as well (5 Marks)

(B) SCM Process (6 Marks)

Answer Key: Explanation of the following is required.

It uses the tools which keep that the necessary change has been implemented adequately to the appropriate component. The SCM process defines a number of tasks:

- Identification of objects in the software configuration
- Version Control
- Change Control
- Configuration Audit
- Status Reporting



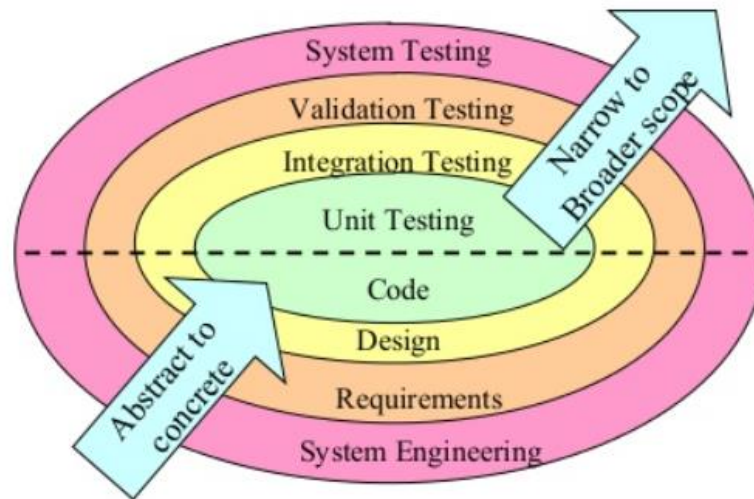
Differentiate between Quality Assurance and Quality control (4 Marks)

Quality Assurance (QA)	Quality Control (QC)
It focuses on providing assurance that quality requested will be achieved.	It focuses on fulfilling the quality requested.
It is the technique of managing quality	It is the technique to verify quality
It does not include the execution of the program.	It always includes the execution of the program.
It is process oriented.	It is product oriented.
It is responsible for a full software development life cycle. Example: Verification	It is responsible for software testing life cycle. Example: Validation

(C) Testing Strategies (4 Marks)

Answer Key: Explanation of the following is required

- Unit / Module Testing
- Integration Testing
- Validation testing
- System Testing



Testing methods applicable for web app Testing. (6 Marks)

Answer Key : Explanation for the following is required

- Functional Testing
- Usability Testing
- Interface Testing
- Database Testing
- Compatibility Testing
- Performance Testing
- Security
- Crowd Testing

Q3

Answer A : Risk Can Be Quantified and Justification (4 Marks)

Answer Key : Risk is the probability of failing to achieve particular costs, performance, and schedule objectives, and the consequences of failing to meet those objectives.

Risk Exposure = Prob(Loss) x Size(Loss)

But precise quantification is difficult for software project

Answer Key: Explanation of Risk Management (6 Marks)

To practice risk management:

- Identify your risks
- Determine the odds of each risk manifesting a problem
- Estimate your exposure in the risks occur (time, money, effort spent)
- Determine which risks to manage
- Take action on risks you have control over
- Plan contingency for those beyond immediate action

Answer B

COCOMO II Model – Explanation (4 marks) and Problem Solving (6 Marks)

- Stands for COConstructive COst MOdel
- Became one of the well-known and widely used estimation models in the industry
- It has evolved into a more comprehensive estimation model called COCOMO II
- COCOMO II is actually a hierarchy of three estimation models
- As with all estimation models, it requires sizing information. Its LOC based model.
- Definition of Application Composition Model, Early design Stage model and Post-architecture stage model.

Solution: The semidetached mode is the most appropriate mode, keeping in view the size, schedule and experience of development time.

$$\text{Hence } E = 3.0(200)1.12 = 1133.12 \text{ PM}$$

$$D = 2.5(1133.12)0.35 = 29.3 \text{ PM}$$

$$\text{Average Staff Size (SS)} = \frac{E}{D} \text{ Persons}$$

$$= \frac{1133.12}{29.3} = 38.67 \text{ Persons}$$

$$\text{Productivity} = \frac{\text{KLOC}}{E} = \frac{200}{1133.12} = 0.1765 \text{ KLOC/PM}$$

$$P = 176 \text{ LOC/PM}$$

Answer C

Answer Key: Explanation of Verification and Validation (4 Marks)

Verification in Software Testing is a process of checking documents, design, code, and program in order to check if the software has been built according to the requirements or not. The main goal of the verification process is to ensure quality of software application, design, architecture etc. The verification process involves activities like reviews, walk-throughs and inspection.

Validation in Software Testing is a dynamic mechanism of testing and validating if the software product actually meets the exact needs of the customer or not. The process helps to ensure that the software fulfils the desired use in an appropriate environment. The validation process involves activities like unit testing, integration testing, system testing and user acceptance testing.

KEY DIFFERENCE (4 Marks)

- Verification process includes checking of documents, design, code and program whereas Validation process includes testing and validation of the actual product.
- Verification does not involve code execution while Validation involves code execution.
- Verification uses methods like reviews, walkthroughs, inspections and desk-checking whereas Validation uses methods like black box testing, white box testing and non-functional testing.
- Verification checks whether the software confirms a specification whereas Validation checks whether the software meets the requirements and expectations.
- Verification finds the bugs early in the development cycle whereas Validation finds the bugs that verification can not catch.
- Verification process targets on software architecture, design, database, etc. while Validation process targets the actual software product.
- Verification is done by the QA team while Validation is done by the involvement of the testing team with the QA team.

Verification process comes before validation whereas Validation process comes after verification.

Example of verification and validation: In Software Engineering, consider the following specification (2 Marks)

A clickable button with name Submet

- Verification would check the design doc and correct the spelling mistake.
- Otherwise, the development team will create a button like
- So new specification is



A clickable button with name Submit

- Once the code is ready, Validation is done. A Validation test found –
- Owing to Validation testing, the development team will make the submit button clickable



University of Mumbai
Examination June 2021

Examinations Commencing from 1st June 2021

Program: Computer Engineering

Curriculum Scheme: Rev2016

Examination: TE Semester VI

Course Code: CSC602 and Course Name: System Programming & Compiler Construction

Time: 2 hour

Max. Marks: 80

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	Which language processor bridges an execution gap but not translator?
Option A:	Pre-processor
Option B:	Assembler
Option C:	Compiler
Option D:	Loader
2.	What are the fields present in MOT used in two pass assembler design? 1. Mnemonic opcode 2. Binary opcode 3. Instruction length 4. Instruction format
Option A:	1 & 2 only
Option B:	2 & 3 only
Option C:	1,2,3 & 4
Option D:	3 & 4 only
3.	Which of the following types of statements are present in assembly program?
Option A:	Imperative statements and assembler directives
Option B:	imperative and declarative statement
Option C:	imperative and declarative statement as well as assembler directive
Option D:	declarative statements and assembler directive
4.	In a two-pass assembler, which of the following process is/are done under first pass? 1. Adding literals to literal table 2. Address resolution of local symbol
Option A:	1 only
Option B:	2 only
Option C:	1 & 2 both
Option D:	Neither 1 nor 2
5.	Consider following code. What will be the entry come under MNT ? The current values for MDTC and MNTC are MDTC=1 and MNTC=1 MACRO INCR4 &AR1, &AR2, &AR3 A 1, &AR1

	A 2, &AR2 A 3, &AR3 MEND
Option A:	Index-4, Macro name- INCR4, MDT index- 1
Option B:	Index-1, Macro name- MACRO, MDT index- 1
Option C:	Index-1, Macro name- INCR4, MDT index- 1
Option D:	Index-4, Macro name- MACRO, MDT index- 4
6.	What is the process of replacing macro name by the statements and instructions included in macro definition is called?
Option A:	Expanding Macro
Option B:	Inserting a Macro
Option C:	Initializing a Macro
Option D:	Installing a Macro
7.	Which of the following is not a data structure used during Macro Processor design?
Option A:	Symbol table
Option B:	MNT
Option C:	MDT
Option D:	ALA
8.	In case of Direct Linking Loader, which are the fields are present in Relocation and Linkage Directory (RLD) card?
Option A:	source card reference number, ESD ID, Length, Flag, Relative Address
Option B:	source card reference number, ESD ID, Flag
Option C:	source card reference number, ESD ID, Relative Address
Option D:	source card reference number, Relative Address, Length
9.	In case of absolute loading scheme, which function is performed by loader?
Option A:	Loading and Allocation
Option B:	Loading
Option C:	Relocation
Option D:	Allocation
10.	In which of the following scheme the loading and linking of external references is done at execution time?
Option A:	Absolute Loading
Option B:	Dynamic Linking
Option C:	General Loading
Option D:	Compile and go loading
11.	Consider following Expression. $-(p * q) + (r + s) - (p + q + r + s)$ How many numbers of temporary variables are required to construct 3 address code?
Option A:	8
Option B:	7
Option C:	9
Option D:	6

12.	In which Code Optimization technique, variables are replaced with constants that have been assigned to them?
Option A:	loop optimization
Option B:	constant folding
Option C:	local optimization
Option D:	Constant propagation
13.	Which technique is applicable to optimize the given code? <pre>a=10; for (j=0 ; j< a*2; j++) { x= j+2; }</pre>
Option A:	Code Motion
Option B:	Copy Propagation
Option C:	Induction Variable Reduction
Option D:	Common Sub-expression Elimination
14.	Which of the following cannot be used as intermediate code form?
Option A:	Post fix notation
Option B:	Three address code
Option C:	Abstract Syntax tree
Option D:	Token
15.	What of the following graph represents flow of control among the set of basic blocks?
Option A:	Hamiltonian graph
Option B:	Control graph
Option C:	Flow graph
Option D:	DAG
16.	What will be the FOLLOW(A) for following grammar? $S \rightarrow AaAb$ $S \rightarrow BaBb$ $A \rightarrow \epsilon$ $B \rightarrow \epsilon$
Option A:	Only a
Option B:	a, b
Option C:	Only b
Option D:	Only ϵ
17.	Which of the following grammar is appropriate for operator precedence grammar?
Option A:	$S \rightarrow EF$
Option B:	$S \rightarrow E * F \mid \epsilon$
Option C:	$S \rightarrow E + F$
Option D:	$S \rightarrow +EF$
18.	Which of the following statement are correct for Syntax Directed Definition? i. The terminals do not have inherited attributes. ii. The non-terminal can have both inherited and synthesized attributes.

	iii. Each grammar symbol is associated with a set of attributes.
Option A:	i only
Option B:	i, ii and iii
Option C:	ii and iii
Option D:	iii only
19.	Which of the following approach is used to evaluate the attributes in L-attributed SDTs?
Option A:	DFS with left-to-right Parsing
Option B:	BFS with left-to-right Parsing
Option C:	DFS with right-to-left Parsing
Option D:	BFS with right-to-left Parsing
20.	Which sentence/s is correct with respect to lexical analyzer? 1. Recognizing the tokens 2. To organize the variables in a lexical order 3. Building a literal and identifier table
Option A:	1 only
Option B:	2 & 3 only
Option C:	1, 2 & 3
Option D:	1 & 3 only

Q2.	Solve any Two 10 marks each
A	Generate SLR parsing table for the following grammar. $S \rightarrow DD$ $D \rightarrow dD \mid e$
B	Explain databases used in Single pass assembler design with suitable example.
C	What is Macro call, Macro expansion, Macro definition? How is macro different from subroutine?
Q3.	
A	Solve any Two 5 marks each
i.	Explain the process of elimination of left recursion with example.
ii.	Compare application software and system software.
iii.	Generate Three address code. For(i=0;i<10;i++) { If (i<5) a=b+c*3; else x=y+z; }
B	Solve any One 10 marks each
i.	What is fundamental process of a loader? Explain dynamic loading in detail.
ii.	Explain loop optimization techniques with example.

University of Mumbai

Examination June 2021

Examinations Commencing from 1st June 2021

Program: Computer Engineering

Curriculum Scheme: Rev2016

Examination: TE Semester VI

Course Code: CSC602 and Course Name: System Programming & Compiler Construction

Time: 2 hour

Max. Marks: 80

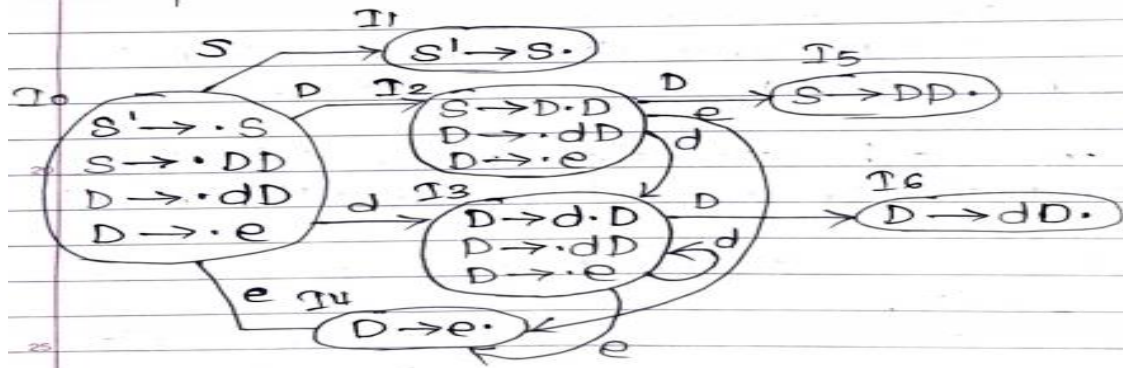
Question Number	Correct Option (Enter either 'A' or 'B' or 'C' or 'D')
Q1.	A
Q2.	C
Q3.	C
Q4	A
Q5	C
Q6	A
Q7	A
Q8.	A
Q9.	B
Q10.	B
Q11.	B
Q12.	D
Q13.	A
Q14.	D
Q15.	C
Q16.	B
Q17.	C
Q18.	B
Q19.	A
Q20.	D

Q2.A. SLR

Step 1: Augment the grammar

- $S' \rightarrow S$
- $S \rightarrow DD$ — ①
- $D \rightarrow dD$ — ②
- $D \rightarrow e$ — ③

Step 2: Construct GOTO graph



Step 3: Design LR(0) parser.

	Action			GOTO	
	d	e	\$	S	D
0	S ₃	S ₄		1	2
1			Accept		
2	S ₃	S ₄			5
3	S ₃	S ₄			6
4	r ₃	r ₃	r ₃		
5	r ₁	r ₁	r ₁		
6	r ₂	r ₂	r ₂		

Step 4: Algorithm

repeat forever

{

let 'x' be stack top number

let 'a' be input symbol

if $M[x, a] = \text{Accept}$ then Accept & Break

else if $M[x, a] = S_i$ then shift 'a' and Push 'i'

else if $M[x, a] = r_j$ then reduce using r_j & perform
GOTO

else ERROR();

}

follow (S') = { \$ }

follow (S) = { \$ }

follow (D) = { e, \$ }

SLR parsing table

	Actions			GOTO	
	d	e	\$	S	D
0	S ₃	S ₄		1	2
1			Accept		
2	S ₃	S ₄			5
3	S ₃	S ₄			6
4		r ₃	r ₃		
5			r ₁		
6			r ₂		

Q. 2 B

The following databases must be explained along with their formats and one example entry in it

1. Symbol Table (03 M)
2. Forward Reference Table (03 M)
3. Cross Reference Table (01 M)
4. Segment Register Table (02 M)
5. Mnemonic Operation Table (01 M)

Q. 2 C

- Macro definition format (02 M)
- How macro name can be used to call macro (01 M)
- How macro can be expanded (02 M)
- Difference between Macro and Subroutine (04 M)

Q. 3 –A-i

Rule for eliminating Left Recursion : (03 M)

$$A \rightarrow A\alpha | \beta$$

then

$$A \rightarrow \beta A'$$
$$A' \rightarrow \alpha A' | \epsilon$$

Example (02 M)

Q. 3 –A-ii

Compare System Program and Application program (05 M)

Q3-A-iii.

Three address code

For(i=0;i<10;i++)

{

If (i<5)

a=b+c*3;

else

x=y+z;

}

101	i=0
102	if(i<10) then goto 104
103	goto 115
104	if(i<5) then goto 106
105	goto 110
106	t2=c*3
107	t3=b+t2
108	a=t3
109	goto 112
110	t4=y+z
111	x=t4
112	t5=i+1
113	i=t5
114	goto 102
115	

Q3-B-i Functions of loader (05 M)

Dynamic loading (05 M)

Q. 3 B ii

The following loop optimization techniques to be discussed along with suitable example

- Code motion (03 M)
- Induction-variable elimination (04 M)
- Reduction in strength (03 M)

University of Mumbai
Examination June 2021

Examinations Commencing from 1st June 2021

Program: Computer Engineering

Curriculum Scheme: Rev2016

Examination: TE Semester VI

Course Code: CSC603 and Course Name: Data Warehousing and Mining

Time: 2 hour

Max. Marks: 80

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	The purpose of the operational system is used to _____
Option A:	Run the business in real time and is based on historical data
Option B:	Takes strategic decisions for business
Option C:	Support decision making and is based on historical data
Option D:	Run the business in real time and is based on current data
2.	Which of following describes a data warehouse well?
Option A:	Can be updated by end users.
Option B:	Contains numerous naming conventions and formats.
Option C:	Organized around important subject areas.
Option D:	Contains only current data
3.	Expected amount of information (in bits) needed to assign a class to a randomly drawn object is _____
Option A:	Gain ratio
Option B:	Gini Index
Option C:	Entropy
Option D:	Information Gain
4.	Which of the following achieves data reduction by detecting redundant attributes
Option A:	Data cube aggregation
Option B:	Dimension reduction
Option C:	Data compression
Option D:	Numerosity reduction
5.	The fraudulent usage of credit card- can be detected using data mining task should be used
Option A:	Prediction
Option B:	Outlier analysis
Option C:	Association analysis
Option D:	Correlation
6.	Given the record of users and movies viewed. Using Jaccard similarity measures, find similarity between {A-B,A-C,B-C }

		Users	Movie 1	Movie 2	Movie 3	Movie 4	movie 5	
		A	1	0	1	0	1	
		B	0	0	1	0	1	
		C	0	1	0	0	1	
Option A:	{0.67,0.25,0.33}							
Option B:	{0.67,0.33,0.25}							
Option C:	{0.5,0.33,0.67}							
Option D:	{0.5,0.25,0.67}							
7.	Five-number summary of a distribution (Minimum, Q1, Median, Q3, Maximum) is displayed by-----							
Option A:	Histogram							
Option B:	quantile plot							
Option C:	Scatterplot							
Option D:	Box plot							
8.	If a set is a frequent set and no superset of this set is a frequent set, then it is called _____.							
Option A:	maximal frequent set							
Option B:	border set							
Option C:	lattice							
Option D:	infrequent sets							
9.	_____ is a mining task that examines the web and hyperlinks structure that connect web pages.							
Option A:	Web content mining							
Option B:	Web structure mining							
Option C:	Web usage mining							
Option D:	Web link mining							
10.	What does Web content mining involve?							
Option A:	analyzing the universal resource locator in Web pages							
Option B:	analyzing the unstructured content of Web pages							
Option C:	analyzing the pattern of visits to a Web site							
Option D:	analyzing the PageRank and other metadata of a Web page							
11.	A sub-database which consists of set of prefix paths in the FP-tree co-occurring with the suffix pattern is called as							
Option A:	Suffix path							
Option B:	FP-tree							
Option C:	Prefix path							
Option D:	Condition pattern base							
12.	In star schema, there is one fact table as F1 is connected with four-dimension							

	tables D1, D2, D3, D4 then fact table will have how many foreign keys?
Option A:	2
Option B:	4
Option C:	3
Option D:	5
13.	If Mean salary is 54,000 Rs. and standard deviation is 16,000 Rs. then find z score value of 73,600 Rs. salary
Option A:	1.225
Option B:	0.351
Option C:	1.671
Option D:	1.862
14.	The generalization of cross-tab which is represented visually is _____ which is also called as data cube.
Option A:	Two-dimensional cube
Option B:	Multidimensional cube
Option C:	N-dimensional cube
Option D:	Cuboid
15.	In KDD and Data mining, noise is referred to as
Option A:	Complex data
Option B:	Meta data
Option C:	Error
Option D:	Repeated data
16.	Find the IQR of the data set {3, 7, 8, 5, 12, 14, 21, 13, 18}.
Option A:	6
Option B:	12
Option C:	16
Option D:	10
17.	Which of the following is not a method to estimate a classifier's accuracy
Option A:	Holdout method
Option B:	Random Sampling
Option C:	Information Gain
Option D:	Bootstrap
18.	For questions given below consider the data Transactions : T1 {F, A, D, B} T2 {D, A, C, E, B} T3 {C, A, B, E} T4 {B, A, D} With minimum support is 60% and the minimum confidence is 80%. Which of the following is not valid association rule?
Option A:	A -> B
Option B:	B -> A
Option C:	D -> A
Option D:	A -> D

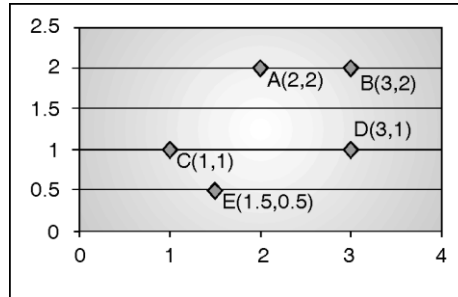
19.	To calculate distance between two isotheticrectangles, _____ is efficient approach and produces cluster of high quality
Option A:	CLARA
Option B:	PAM
Option C:	Spatial mining
Option D:	IR Approximation
20.	Geographers typically model the world with objects located at different places on surface of the earth. Through _____ model, the real word entities are represented by lines, points and polygons
Option A:	Vector data model
Option B:	Raster data model
Option C:	Network data model
Option D:	Topology data model

Q2	Solve any Four out of Six 5 marks each																				
A	<i>Consider Metadata as an equivalent of Amazon book store, where each data element is book. What this meta data will contain. Explain.</i>																				
B	<i>Suppose a group of sales price records has been sorted as follows: 6, 9, 12, 13, 15, 25, 50, 70, 72, 92, 204, 232. Partition them into three bins by equal-frequency (Equi-depth) partitioning method. Perform data smoothing by bin mean.</i>																				
C	<i>Suppose that the data for analysis includes the attribute age. The age values for the data tuples are (in increasing order): 13, 15, 16, 16, 19, 20, 23, 29, 35, 41, 44, 53, 62, 69, 72 Use min-max normalization to transform the value 45 for age onto the range [0:0, 1:0].</i>																				
D	<i>Use K-means algorithm to create 3 - clusters for given set of values: {2, 3, 6, 8, 9, 12, 15, 18, 22}</i>																				
E	<i>Transaction database is given Below. Min Support = 2. Draw FP-Tree.</i> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>TID</th> <th>List of item_Ids</th> </tr> </thead> <tbody> <tr> <td>T100</td> <td>I1, I2, I5</td> </tr> <tr> <td>T200</td> <td>I2, I4</td> </tr> <tr> <td>T300</td> <td>I2, I3</td> </tr> <tr> <td>T400</td> <td>I1, I2, I4</td> </tr> <tr> <td>T500</td> <td>I1, I3</td> </tr> <tr> <td>T600</td> <td>I2, I3</td> </tr> <tr> <td>T700</td> <td>I1, I3</td> </tr> <tr> <td>T800</td> <td>I1, I2, I3, I5</td> </tr> <tr> <td>T900</td> <td>I1, I2, I3</td> </tr> </tbody> </table>	TID	List of item_Ids	T100	I1, I2, I5	T200	I2, I4	T300	I2, I3	T400	I1, I2, I4	T500	I1, I3	T600	I2, I3	T700	I1, I3	T800	I1, I2, I3, I5	T900	I1, I2, I3
TID	List of item_Ids																				
T100	I1, I2, I5																				
T200	I2, I4																				
T300	I2, I3																				
T400	I1, I2, I4																				
T500	I1, I3																				
T600	I2, I3																				
T700	I1, I3																				
T800	I1, I2, I3, I5																				
T900	I1, I2, I3																				
F	<i>Write short note on Spatial Clustering Techniques : CLARANS .</i>																				
Q3	Solve any Two Questions out of Three 10 marks each																				
A	<i>For a Supermarket Chain consider the following dimensions, namely Product, store, time , promotion. The schema contains a central fact tables sales facts with three measures unit_sales, dollars_sales and dollar_cost.</i>																				

Design star schema and calculate the maximum number of base fact table records for the values given below :
 Time period : 5 years
 Store : 300 stores reporting daily sales
 Product : 40,000 products in each store (about 4000 sell in each store daily)
 Promotion : a sold item may be in only one promotion in a store on a given day

B

Use the data given below. Create adjacency matrix. Use complete link algorithm to cluster given data set. Draw dendrogram.



C

Using the following training data set. Create classification model using decision-tree and draw final Tree.

Tid	Income	Age	Own House
1.	Very High	Young	Yes
2.	High	Medium	Yes
3.	Low	Young	Rented
4.	High	Medium	Yes
5.	Very high	Medium	Yes
6.	Medium	Young	Yes
7.	High	Old	Yes
8.	Medium	Medium	Rented
9.	Low	Medium	Rented
10.	Low	Old	Rented
11.	High	Young	Yes
12.	medium	Old	Rented

University of Mumbai

Examination June 2021

Examinations Commencing from 1st June 2021

Program: Computer Engineering

Curriculum Scheme: Rev2016

Examination: TE Semester VI

Course Code: CSC603 and Course Name: Data Warehousing and Mining

Time: 2 hour

Max. Marks: 80

=====

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

Q.2.

(B)

1. Sort the given data

6, 9, 12, 13, 15, 25, 50, 70, 72, 92, 204, 232

2. Partition the data into equal frequency bin of size =4

Bin1: 6, 9, 12, 13

Bin 2: 15, 25, 50, 70

Bin 3: 72, 92, 204, 232

3. Calculate the arithmetic mean of each bin

Bin 1: 10

Bin 2: 40

Bin 3: 150

4. Replace each value in the bin with its respective arithmetic mean

Bin1: 10,10,10,10

Bin 2: 40,40,40,40

Bin 3: 150,150,150,150

(C)

Min-max normalization performs a linear transformation on the original data.

Let A be the attribute age

Suppose that min_A and max_A are the minimum and maximum values of an attribute, *age* so using formula,

$$v'_i = \frac{v_i - \min_A}{\max_A - \min_A(\text{new_max}_A - \text{new_min}_A) + \text{new_min}_A} .$$

$\min_A = 13$ and $\max_A = 72$

$\text{new_min}_A = 0$ and $\text{new_max}_A = 1.0$

$$v_i = 45$$

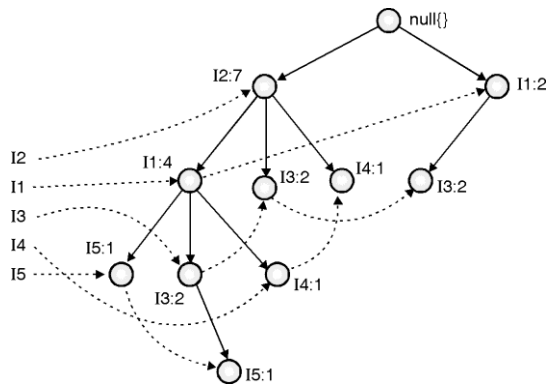
$$v'_i = \frac{(45 - 13)}{(72 - 13) \times (1.0 - 0) + 0}$$

$$v'_i = 0.542$$

(D)

Cluster 1 = {6,8,9} , Cluster 2 = { 2,3 } , Cluster 3 = { 12,15,18,22 }

(E)

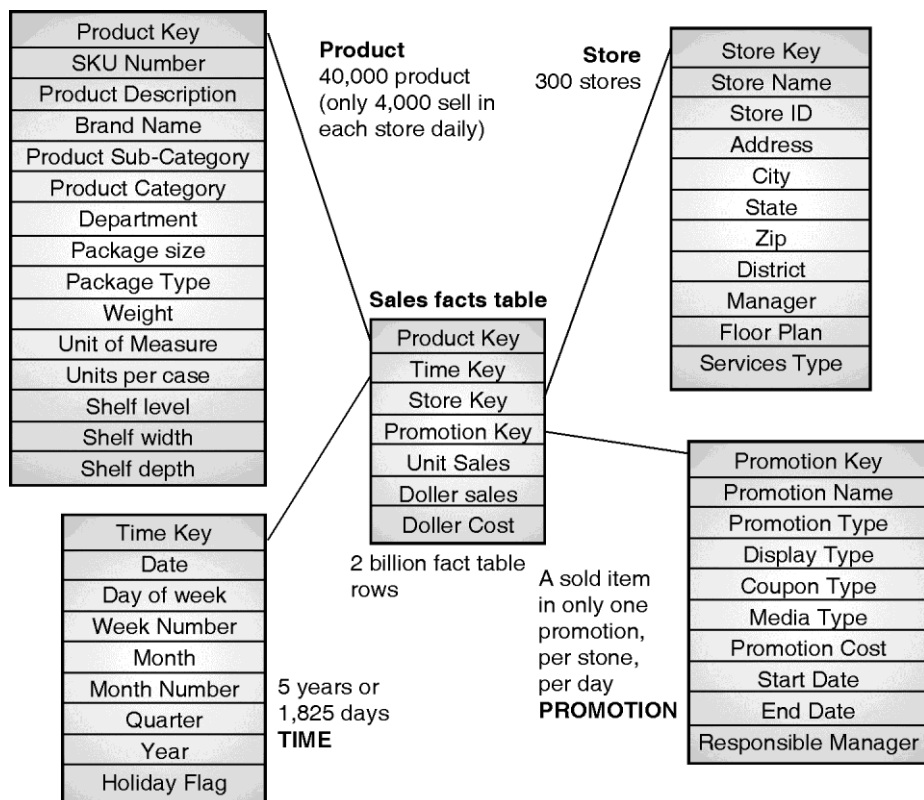


Item ID	Support Count
I2	7
I1	6
I3	6
I4	2
I5	2

Q.3

(A)

(a) Star schema :



(b) Time period = 5 years × 365 days = 1825

There are 300 stores,

Each stores daily sale = 4000

Promotion = 1

Maximum number of fact table records: 1825 × 300 × 4000 × 1 = 2 billion

(B)

Adjacency Matrix

	A	B	C	D	E
A	0				
B	1	0			
C	1.41	2.24	0		
D	1.41	1	2	0	
E	1.58	2.12	0.71	1.58	0

Step 1 : Closest clusters are merged where the distance is the smallest measured by looking at the maximum distance between any two point.

Since C, E is minimum we can combine clusters C, E.

	A	B	(C, E)	D
A	0			
B	1	0		
(C,E)	1.58	2.24	0	
D	1.41	1	2	0

Step 2 : Now A and B is having minimum closest measure value therefore we merge these two clusters.

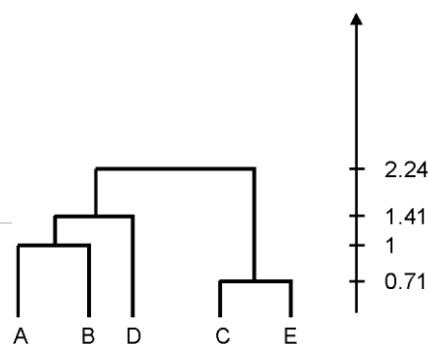
	(A, B)	(C,E)	D
(A,B)	0		
(C, E)	2.24	0	
D	1.41	2	0

Step 3 : Cluster (A,B) and D can be merged together as they are having minimum distance value.

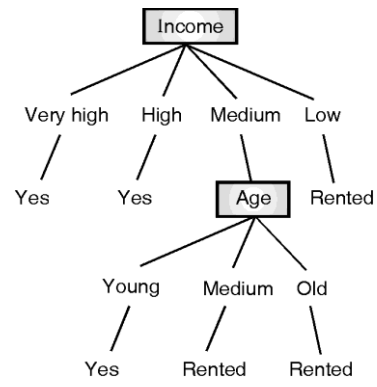
	(A, B,D)	(C,E)
(A,B,D)	0	
(C, E)	2.24	0

Step 4 : In the last step there are only two clusters to be combined they are, (A,B,D) and (C,E).

Final dendrogram



(C)



University of Mumbai
Examination June 2021

Examinations Commencing from 1st June 2021

Program: **Computer Engineering**

Curriculum Scheme: Rev2016

Examination: TE Semester VI

Course Code: **CSC604** and Course Name: **Cryptography and System Security**

Time: 2 hour

Max. Marks: 80

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	_____ defines a security service as a service that is provided by a protocol layer of communicating open systems and that ensures adequate security of the systems or of data transfers.
Option A:	X.800
Option B:	X.809
Option C:	X.832
Option D:	X.802
2.	_____ are fundamental to a number of public-key algorithms, including and the digital signature algorithm (DSA).
Option A:	Discrete logarithms
Option B:	Chinese remainder theorem
Option C:	Fermat's theorem
Option D:	Miller and Rabin algorithm
3.	Plain text message is: "meet me after the toga party" with a rail fence of depth 2. Compute cipher text.
Option A:	MEMATRHTGPRYETEFETEOAAT
Option B:	MEMATRHTGPRYETEFETFOAAT
Option C:	MEMATRHTHPRYETEFETEOAAT
Option D:	MEMATRHTGPRYETEFFTEOAOT
4.	In _____ mode, the same plaintext value will always result in the same cipher text value.
Option A:	Cipher Block Chaining
Option B:	Cipher Feedback
Option C:	Electronic code book
Option D:	Output Feedback
5.	DES encrypting the plaintext as block of _____ bits.
Option A:	64
Option B:	56
Option C:	128
Option D:	32
6.	_____ is a symmetric block cipher that is intended to replace DES as the approved standard for a wide range of applications.
Option A:	AES

Option B:	RSA
Option C:	MD5
Option D:	RC5
7.	The number of rounds in RC5 can range from 0 to _____
Option A:	127
Option B:	63
Option C:	31
Option D:	255
8.	How many rounds does the AES-192 perform?
Option A:	10
Option B:	14
Option C:	16
Option D:	12
9.	For the Knapsack: {1 6 8 15 24}, Find the cipher text value for the plain text 10011.
Option A:	40
Option B:	15
Option C:	14
Option D:	39
10.	Which of the following is not possible through hash value?
Option A:	Password check
Option B:	Data integrity check
Option C:	Data retrieval
Option D:	Digital signature
11.	Which of the following is not an element/field of the X.509 certificates?
Option A:	Issuer Name
Option B:	Serial Modifier
Option C:	Issue unique identifier
Option D:	Signature
12.	_____ is responsible for distributing keys to pairs of users (hosts, processes, applications) as needed
Option A:	Key distribution center
Option B:	Key analysis center
Option C:	UKey storing center
Option D:	HKey storing center
13.	A digital certificate system is _____.
Option A:	uses third-party CAs to validate a user's identity
Option B:	uses digital signatures to validate a user's identity
Option C:	uses tokens to validate a user's identity
Option D:	are used primarily by individuals for personal correspondence
14.	Hashed message is signed by a sender using
Option A:	His public key
Option B:	His private key

Option C:	Receivers public key
Option D:	Receivers private key
15.	The man-in-the-middle attack can endanger the security of the Diffie-Hellman method if two parties are not
Option A:	Authenticated
Option B:	Joined
Option C:	Submit
Option D:	Separate
16.	Which of the following does authorization aim to accomplish?.
Option A:	Restrict what operations/data the user can access
Option B:	Determine if the user is an attacker
Option C:	Flag the user if he/she misbehaves
Option D:	Determine who the user is
17.	_____ operates in the transport mode or the tunnel mode.
Option A:	IPSec
Option B:	SSL
Option C:	PGP
Option D:	BGP
18.	When a hash function is used to provide message authentication, the hash function value is referred to as
Option A:	Message Field
Option B:	Message Digest
Option C:	Message Score
Option D:	Message Leap
19.	Which of the following tool would NOT be useful in figuring out what spyware or viruses could be installed on a client's computer?
Option A:	Wireshark
Option B:	Malware Bytes
Option C:	HighjackThis
Option D:	HitmanPro
20.	What is honey pot attack?
Option A:	dummy device put into the network to attract attackers
Option B:	single line threat
Option C:	Ip spoofing bypass
Option D:	recognition attack

Q2	Solve any Two	10 marks each
A	Explain Security Services and Mechanisms in detail. Explain the relationship between them.	
B	What is meant by the Diffie-Hellman key exchange algorithm? Explain with example.	
C	Describe HMAC algorithm. Comment on the security of HMAC.	
Q3	Solve any Two	10 marks each
A	Describe signing and verification in Digital Signature Algorithm.	

B	Explain any 2 ways to classify Intrusion Detection Systems.
C	Explain Man-in-the-Middle and Flooding attacks concept in detail.

University of Mumbai

Examination June 2021

Examinations Commencing from 1st June 2021

Program: Computer Engineering

Curriculum Scheme: Rev2016

Examination: TE Semester VI

Course Code: CSC604 and Course Name: Cryptography and System Security

Time: 2 hour

Max. Marks: 80

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

University of Mumbai
Examination June 2021

Examinations Commencing from 1st June 2021

Program: Computer Engineering

Curriculum Scheme: Rev2016

Examination: TE Semester VI

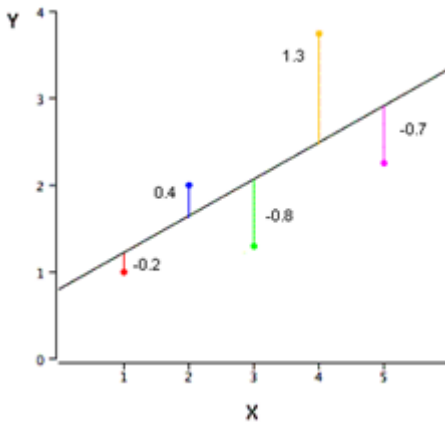
Course Code: CSDLO6021 and Course Name: Machine Learning

Time: 2 hour

Max. Marks: 80

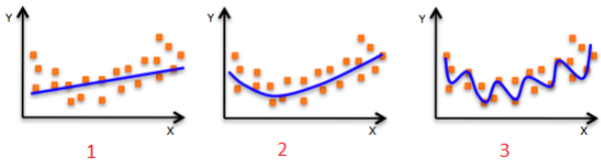
Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	Which of the following are examples of unsupervised learning? i. Modeling a spam filter from a set of labeled emails as spam and not spam ii. Given a set of news articles found on the web, group them into articles under different categories iii. Given a database of customer data, automatically discover market segments and group customers into different market segments iv. Given a database of patients diagnosed as either having diabetes or not, learn to classify new patients as having diabetes or not
Option A:	Both i and iv
Option B:	Both i and iii
Option C:	Both ii and iii
Option D:	Both iii and iv
2.	Which of the following options are true about Machine Learning? 1. Machine learning is automatic learning based on experience 2. Machine learning is programmed so that it learns, and past experience is not required. 3. It can learn and improve from the past experience without being explicitly programmed. 4. Machines can learn from past experience, but it must be explicitly programmed.
Option A:	1 and 2
Option B:	2 and 4
Option C:	1 and 4
Option D:	3 and 4
3.	Which of the following is an example of reinforcement learning?
Option A:	Stock price prediction
Option B:	Sentiment analysis

Option C:	Customer segmentation
Option D:	Robot in a maze
4.	In Downhill Simplex method, if $f(x)$ at the reflected point is greater than $f(x)$ at worst point (N) then the new point is obtained by
Option A:	Contraction
Option B:	Multiple Reflection
Option C:	Expansion
Option D:	Multiple contraction
5.	In classical Newton's Method, having Hessian Matrix H, Gradient G, X_{k+1} is computed using
Option A:	$X_{k+1} = X_k + H_k^{-1} * G_k$
Option B:	$X_{k+1} = X_k - H_k^{-1} * G_k$
Option C:	$X_{k+1} = X_k - H_k * G_k$
Option D:	$X_{k+1} = X_k + H_k * G_k$
6.	Which of the following is not true about the derivative free techniques?
Option A:	They use evolutionary concepts.
Option B:	The objective function has to be differentiable
Option C:	These methods use an empirical approach for analysis.
Option D:	Random search and Downhill Simplex are examples of Derivative free techniques.
7.	Given $X=[1 \ 2 \ 3 \ 4]$ $W=[1 \ 1 \ -1 \ -1]$ compute $f(\text{net})$ given $\lambda = 0.5$ using i. Bipolar continuous ii. Unipolar continuous activation function
Option A:	i. 0.7615 ii. 0.880
Option B:	i. 0.880 ii. 0.7615
Option C:	i. -0.7615 ii. 0.1192
Option D:	i. 0.119 ii. -0.7615
8.	Hebbian learning is an example of _____ and perceptron learning is an example of _____
Option A:	Feedforward supervised learning, supervised binary response
Option B:	Feedforward unsupervised learning, supervised binary response
Option C:	Feedback supervised learning, unsupervised binary response
Option D:	Feedback unsupervised learning, supervised multivariate response
9.	_____ is a type of learning rule which works with a layer of neurons.
Option A:	Perceptron
Option B:	Hebbian
Option C:	Widrow Hoff
Option D:	Winner takes all

10.	Which of these statements are false with respect to the metrics in linear regression? a. For a strong linear regression R^2 value should be high b. Multiple R value of 1 represents perfect positive relationship c. Karl Pearson value of -1 indicates total negative linear correlation d. High value of Sum of Squared Errors(SSE) indicates perfect fit									
Option A:	Both A and B are false									
Option B:	Both A and C are false									
Option C:	Both B and C are false									
Option D:	Only D is false									
11.	The graph below represents a regression line predicting Y from X. The values on the graph shows the residuals for each predicted value. Use this information to compute the Sum of squared errors (SSE) 									
Option A:	4.02									
Option B:	3.02									
Option C:	1.01									
Option D:	0									
12.	<table border="1" data-bbox="375 1624 1428 1736"> <thead> <tr> <th></th> <th>Actual True</th> <th>Actual False</th> </tr> </thead> <tbody> <tr> <th>Predicted True</th> <td>156</td> <td>20</td> </tr> <tr> <th>Predicted False</th> <td>14</td> <td>50</td> </tr> </tbody> </table> <p>Compute the specificity and the precision?</p>		Actual True	Actual False	Predicted True	156	20	Predicted False	14	50
	Actual True	Actual False								
Predicted True	156	20								
Predicted False	14	50								
Option A:	Specificity = 88.6% Precision = 71.4%									
Option B:	Specificity = 71.4 % Precision = 88.6%									
Option C:	Specificity = 28.5% Precision = 11.36%									
Option D:	Specificity = 71.4% Precision = 11.36%									
13.	Which is not true statement about Kernel Trick									

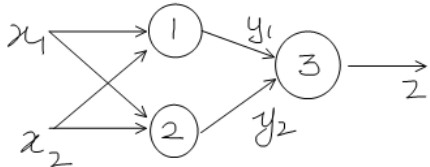
Option A:	A Kernel Trick is a method where a Non Linear data is projected onto a higher dimension space so as to make it easier to classify the data where it could be linearly divided by a plane.
Option B:	A Kernel Trick is a method of transforming the original (non-linear) input data into a higher dimensional space (as a linear representation of data).
Option C:	The Kernel Trick allows us to take linear Support Vector Machines and extend their functionality to classify non-linear data sets.
Option D:	A Kernel Trick is a method which can easily separates the data points in a lower dimensionality space
14.	The difference between naïve Bayesian classifier and Bayesian belief networks is
Option A:	The joint conditional probability distributions are considered in Bayesian Belief networks
Option B:	The joint conditional probability distribution is not considered in Bayesian Belief networks
Option C:	Class conditional independence is always considered in Bayesian Belief networks
Option D:	Class conditional independence is sometimes considered in Bayesian Belief Networks
15.	<p>Today's weather Tomorrow's weather Initial Probability values</p> <p>Sunny 0.25 Rainy 0.75 Foggy 0.30</p> <pre> graph TD Sunny((Sunny)) -- 0.8 --> Sunny Sunny -- 0.2 --> Rainy((Rainy)) Rainy -- 0.6 --> Rainy Rainy -- 0.05 --> Sunny Rainy -- 0.3 --> Foggy((Foggy)) Foggy -- 0.5 --> Foggy Foggy -- 0.15 --> Sunny Foggy -- 0.2 --> Rainy </pre> <p>Given that today is sunny what is the probability that tomorrow is sunny and the day after is rainy</p>
Option A:	0.01
Option B:	0.004

Option C:	0.04
Option D:	0.32
16.	<p>What is true about Markov Property</p> <p>I. Markov Property is very useful for explaining events, and it cannot be the true model of the underlying situation in most cases.</p> <p>II. The state of the system at time $t+1$ depends only on the state of the system at time t</p> <p>III. The advantages of Markov property are complexity and forecasting accuracy.</p> <p>IV. Markov property is used to forecast the value of a variable whose predicted value is influenced only by its current state</p>
Option A:	i and ii
Option B:	ii and iii
Option C:	ii and iv
Option D:	iii and iv
17.	<p>A square matrix is _____ if all eigen values are _____</p> <p>Positive definite, Positive</p> <p>Negative definite, Negative</p> <p>Positive definite, Negative</p> <p>Negative definite, positive</p>
Option A:	Both ii and i are correct
Option B:	Both iii and iv are correct
Option C:	All four options are wrong
Option D:	Either iii or iv is right
18.	<p>Identify the correct options regarding Principal Component Analysis</p> <p>(a) Principal component analysis (PCA) can be used with variables of any mathematical types: quantitative, qualitative, or a mixture of these types</p> <p>(b) The major principal component axis has dimensions having the maximum variance.</p> <p>(c) The major principal component axis has dimensions having the minimum variance</p> <p>(d) The most information is retained among the top few principal axes.</p>
Option A:	Both a and b

Option B:	Both b and d
Option C:	Both a and d
Option D:	Both c and d
19.	Compute the eigen values for matrix $A = \begin{bmatrix} 7 & 3 & 3 & -1 \end{bmatrix}$
Option A:	$\lambda_1 = 8; \lambda_2 = -2$
Option B:	$\lambda_1 = -8; \lambda_2 = 2$
Option C:	$\lambda_1 = 4; \lambda_2 = -4$
Option D:	$\lambda_1 = -4; \lambda_2 = 4$
20.	 <p>In the graphs 1, 2 and 3 which is best fitted and which is overfitted?</p>
Option A:	2 is best-fitted and 1 is over-fitted
Option B:	1 is best-fitted and 2 is over-fitted
Option C:	2 is best-fitted and 3 is over-fitted
Option D:	1 is best-fitted and 3 is over-fitted

Q2 (20 Marks Each)																																					
A	Solve any Two 5 marks each																																				
i.	Why is the Support Vector Machine(SVM) called the maximum margin classifier? Explain mathematically the formulation of margin.																																				
ii.	What is a saddle point? Minimize $f(x) = x_1^2 + x_2^2 + 2x_1x_2$, with starting initial point X_0 is $[0.5, -0.1]$ (Perform 2 iteration only) using the steepest descent method																																				
iii.	What are the steps in designing a Machine Learning Application																																				
B	Solve any One 10 marks each																																				
i.	<p><i>Two questions of 10 marks each have to be asked</i></p> <p>For the following data, to construct the decision tree calculate Gini indexes and determine which attribute is the root attribute. (4)</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>Sr. No</th> <th>Income</th> <th>Defaulting Level</th> <th>Credit Score</th> <th>Location</th> <th>Give Loan?</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>low</td> <td>high</td> <td>high</td> <td>bad</td> <td>no</td> </tr> <tr> <td>2</td> <td>low</td> <td>high</td> <td>high</td> <td>good</td> <td>no</td> </tr> <tr> <td>3</td> <td>high</td> <td>high</td> <td>high</td> <td>bad</td> <td>yes</td> </tr> <tr> <td>4</td> <td>medium</td> <td>medium</td> <td>high</td> <td>bad</td> <td>yes</td> </tr> <tr> <td>5</td> <td>medium</td> <td>low</td> <td>low</td> <td>bad</td> <td>no</td> </tr> </tbody> </table>	Sr. No	Income	Defaulting Level	Credit Score	Location	Give Loan?	1	low	high	high	bad	no	2	low	high	high	good	no	3	high	high	high	bad	yes	4	medium	medium	high	bad	yes	5	medium	low	low	bad	no
Sr. No	Income	Defaulting Level	Credit Score	Location	Give Loan?																																
1	low	high	high	bad	no																																
2	low	high	high	good	no																																
3	high	high	high	bad	yes																																
4	medium	medium	high	bad	yes																																
5	medium	low	low	bad	no																																

	<table border="1"> <tr><td>6</td><td>medium</td><td>low</td><td>low</td><td>good</td><td>yes</td></tr> <tr><td>7</td><td>high</td><td>low</td><td>low</td><td>good</td><td>yes</td></tr> <tr><td>8</td><td>low</td><td>medium</td><td>high</td><td>bad</td><td>no</td></tr> <tr><td>9</td><td>low</td><td>low</td><td>low</td><td>bad</td><td>no</td></tr> <tr><td>10</td><td>medium</td><td>medium</td><td>low</td><td>bad</td><td>no</td></tr> <tr><td>11</td><td>low</td><td>medium</td><td>low</td><td>good</td><td>yes</td></tr> <tr><td>12</td><td>high</td><td>medium</td><td>high</td><td>good</td><td>yes</td></tr> <tr><td>13</td><td>high</td><td>high</td><td>low</td><td>bad</td><td>no</td></tr> <tr><td>14</td><td>medium</td><td>medium</td><td>high</td><td>good</td><td>yes</td></tr> </table>	6	medium	low	low	good	yes	7	high	low	low	good	yes	8	low	medium	high	bad	no	9	low	low	low	bad	no	10	medium	medium	low	bad	no	11	low	medium	low	good	yes	12	high	medium	high	good	yes	13	high	high	low	bad	no	14	medium	medium	high	good	yes
6	medium	low	low	good	yes																																																		
7	high	low	low	good	yes																																																		
8	low	medium	high	bad	no																																																		
9	low	low	low	bad	no																																																		
10	medium	medium	low	bad	no																																																		
11	low	medium	low	good	yes																																																		
12	high	medium	high	good	yes																																																		
13	high	high	low	bad	no																																																		
14	medium	medium	high	good	yes																																																		
ii.	<p>List down the steps of PCA Using PCA compute the transformed matrix of A Where A is</p> $\begin{bmatrix} 1 & 2 \\ 0.5 & 1.5 \\ 0 & 0.5 \\ -0.5 & 0.25 \end{bmatrix}$																																																						

Q3. (20 Marks Each)	
A	Solve any Two 5 marks each
i.	Define logit function. Explain the importance of logit function in logistic regression with appropriate example
ii.	<p>Given</p>  <p style="text-align: center;">$X = [3, 5] \quad W = \begin{bmatrix} 1 & 2 \\ 4 & -2 \end{bmatrix} \quad Y = [1, -5] \quad c = 1$</p> <p>Compute output Z using binary bipolar activation function. Also compute the new weights $y_1, y_2, w_{11}, w_{12}, w_{21}, w_{22}$</p>
iii.	Define covariance ? For the given dataset, compute the covariance matrix

	X_1 X_2 2.5 2.4 0.5 0.7 2.2 2.9 1.9 2.2 3.1 3.0 2.3 2.7 2.0 1.6 1.0 1.1 1.5 1.6 1.2 0.9
B	Solve any One 10 marks each
ii.	Explain Linear Separability problem? (2) Solve a linearly separable problem (AND Gate) Solve a linearly non separable problem (XOR gate) both using McCulloch Pitt Model ?
ii.	What is the role of radial basis function in separating nonlinear patterns? Explain with XOR Example.

University of Mumbai
Examination June 2021

Examinations Commencing from 1st June 2021

Program: Computer Engineering

Curriculum Scheme: Rev2016

Examination: TE Semester VI

Course Code: CSDLO6021 and Course Name: Machine Learning

Time: 2 hour

Max.Marks: 80

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

Q19.	A
Q20.	C

Q2. (20 Marks Each)	
A	Solve any Two 5 marks each
i.	<p>Why SVM is called as maximum margin classifier? Explain mathematically the formulation of margin. Marking scheme: <i>SVM classifier as maximum margin classifier (2Marks)</i> <i>Mathematical Formulation of margin (3Marks)</i></p>
ii.	<p>What is a saddle point? Minimize $f(x)=x_1^2+x_2^2+2x_1x_2$, with starting initial point X_0 is $[0.5,-0.1]$ (Perform 2 iteration only)using the steepest descent method Marking scheme: <i>What is a saddle point? (2Mark)</i> <i>$X_1 = [0.1 -0.5]$ $X_2 = [-0.3 -0.9]$ (3 Marks)</i></p>
iii.	<p>What are the steps in designing a Machine Learning Application. Marking scheme: 5 marks <i>Explanation of following steps are required</i> <i>Following are the steps in designing a learning system</i></p> <ul style="list-style-type: none"> ● <i>Choosing the training experience</i> ● <i>Choosing the target function</i> ● <i>Choosing the representation of the target function</i> ● <i>Choosing the function approximation algorithm</i> ● <i>Final Design</i>
B	Solve any One 10 marks each
i.	<p>Marking scheme: <i>Calculating gini index for Income - 0.39 (2M)</i> <i>Calculating gini index for Defaulting Level- 0.44 (2M)</i> <i>Calculating gini index for Credit Score-0.46 (2M)</i> <i>Calculating gini index for Location - 0.34 (2M)</i> <i>Determining the root attribute-2M</i></p>

Step 1: Calculating gini index for Income

Gini (Income = low)

$$= 1 - \left[\left(\frac{1}{5} \right)^2 + \left(\frac{4}{5} \right)^2 \right] = 1 - [0.4 + 0.64] = 1 - 1.04 = -0.04$$

Gini (Income = medium)

$$= 1 - \left[\left(\frac{3}{5} \right)^2 + \left(\frac{2}{5} \right)^2 \right] = 1 - [0.36 + 0.16] = 1 - 0.52 = 0.48$$

Gini (Income = high)

$$= 1 - \left[\left(\frac{3}{4} \right)^2 + \left(\frac{1}{4} \right)^2 \right] = 1 - [0.5625 + 0.0625] = 1 - 0.625 = 0.375$$

$$\begin{aligned} \text{Gini (Income)} &= \frac{5}{14} \times 0.04 + \frac{5}{14} \times 0.48 + \frac{4}{14} \times 0.375 \\ &= 0.14 + 0.17 + 0.11 \\ &= \boxed{0.39} \end{aligned}$$

Step 2: Calculating gini index for Defaulting level

Gini (D.L. = low)

$$= 1 - \left[\left(\frac{2}{4} \right)^2 + \left(\frac{2}{4} \right)^2 \right] = 1 - [0.25 + 0.25] = 0.5$$

Gini (D.L. = medium)

$$= 1 - \left[\left(\frac{4}{6} \right)^2 + \left(\frac{2}{6} \right)^2 \right] = 1 - [0.45 + 0.11] = 0.44$$

Gini (D.L. = high)

$$= 1 - \left[\left(\frac{1}{4} \right)^2 + \left(\frac{3}{4} \right)^2 \right] = 0.375$$

Gini (Defaulting level)

$$\begin{aligned} &= \frac{0.4}{14} \times 0.5 + \frac{6}{14} \times 0.44 + \frac{4}{14} \times 0.375 \\ &= 0.14 + 0.19 + 0.11 = \boxed{0.44} \end{aligned}$$

Gini

Step 3: Calculating gini index for credit score

Gini (Score = low)

$$= 1 - \left[\left(\frac{2}{7} \right)^2 + \left(\frac{5}{7} \right)^2 \right]$$

$$= 1 - [0.08 + 0.50] = 1 - 0.58 = 0.42$$

Gini (Score = high)

$$= 1 - \left[\left(\frac{4}{7} \right)^2 + \left(\frac{3}{7} \right)^2 \right]$$

$$= 1 - [0.32 + 0.18] = 0.5$$

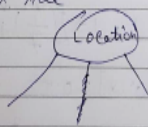
$$\text{Gini (Score)} = \frac{7}{14} \times 0.42 + \frac{7}{14} \times 0.5 = \boxed{0.46}$$

Step 4: Calculating gini index for Location

Gini (Location = good)

$$= 1 - \left[\left(\frac{5}{6} \right)^2 + \left(\frac{1}{6} \right)^2 \right]$$

$$= 1 - [0.67 + 0.03] = 0.3$$

$$\begin{aligned}
 & \text{Gini}(\text{Location} = \text{Bad}) \\
 &= 1 - \left[\left(\frac{6}{8}\right)^2 + \left(\frac{2}{8}\right)^2 \right] \\
 &= 1 - [0.56 + 0.0625] = 0.375 \\
 \\
 & \text{Gini}(\text{Location}) \\
 &= \frac{6}{14} \times 0.3 + \frac{8}{14} \times 0.375 \\
 &= 0.129 + 0.213 = \boxed{0.34} \\
 \\
 & \text{Gini}(\text{Income}) = 0.39 \\
 & \text{Gini}(\text{Defaulting level}) = 0.44 \\
 & \text{Gini}(\text{Credit score}) = 0.46 \\
 & \text{Gini}(\text{Location}) = 0.34 \\
 \\
 & \therefore \text{Location will be the root of the} \\
 & \text{decision tree}
 \end{aligned}$$


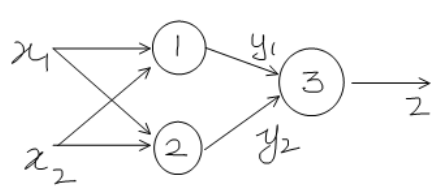
ii

1	2
0.5	1.5
0	0.5
-0.5	0.25

Marks:-

Steps of PCA algorithm (3Marks)

1. Normalise the data by Mean normalization method (say Z is normalized data)
2. Calculate covariance matrix $C = \frac{1}{N-1} Z'Z$
3. Compute Eigenvalues and normalize them. Sort eigenvalues in descending order.
4. Find eigenvectors for each eigenvalue.
5. choose the k eigenvectors that correspond to the k largest eigenvalues where k is the number of dimensions of the new feature subspace ($k \leq d$) (d : Number of attributes)
6. Put them in matrix V
7. Compute Principal Components by $P = Z V^T$

	<p>Marks distribution:- Mean – 1 mark [0.25,1.0] Covariance matrix – 1 mark 0.41 0.52 0.52 0.68 Eigen values and eigen vectors -5 marks 1.0 [-0.613 -0.789] 0.011 [-0.789 0.6135] Arranging in proper order – 1 mark Transformed matrix – 2 marks -1.2 -0.017 -0.49 0.07 0.59 -0.14 1.10 0.09</p>
<p>Q3. (20 Marks Each)</p>	
<p>A</p>	<p>Solve any Two 5 marks each</p>
<p>i.</p>	<p>Define logit function. Explain the importance of logit function in logistic regression with appropriate example. Marking scheme: Logit function explanation – 2 marks Example – 2 marks Relevance of logit function in logistic regression – 1 mark</p>
<p>ii.</p>	<p>Given</p>  <p style="text-align: center;"> $X = [3, 5]$ $W = \begin{bmatrix} 1 & 2 \\ 4 & -2 \end{bmatrix}$ $Y = [1, -5]$ $c = 1$ </p> <p>Compute output Z using binary bipolar activation function. Also compute the new weights $y_1, y_2, W_{11}, W_{12}, W_{21}, W_{22}$</p>

$$W_{k+1} = W_k + c f(\text{net}) x_k$$

$$\begin{aligned} \text{net}_1 &= x_1 w_{11} + x_2 w_{12} \\ &= 3 \times 1 + 5 \times 2 = 13 \end{aligned}$$

$$f(\text{net}_1) = 1$$

$$\begin{aligned} \text{net}_2 &= x_2 w_{22} + x_1 w_{21} \\ &= 5(-2) + 3 \times 4 \\ &= -10 + 12 = 2 \end{aligned}$$

$$f(\text{net}_2) = 1$$

$$\begin{aligned} \text{net}_3 &= x_3 y_1 + x_4 y_2 \\ &= 1 \times 1 + 1 \times (-5) = 1 - 5 = -4 \end{aligned}$$

$$-5 + c(1)(-1) \\ = -5 - 1 = \underline{-6}$$

$$W_{11}^{(1)} = W_{11}^{(0)} + c x_1 f(\text{net}_1) \\ = 1 + 1 * 3 * 1 = 4$$

$$W_{12}^{(1)} = W_{12}^{(0)} + c x_2 f(\text{net}_1) \\ = 2 + 1 * 5 * 1$$

$$= 2 + 5 = \underline{7}$$

$$W_{21}^{(1)} = W_{21}^{(0)} + c x_1 f(\text{net}_2)$$

$$4 + 1 * 3 * 1 = \underline{7}$$

$$W_{22}^{(1)} = W_{22}^{(0)} + c x_2 f(\text{net}_2)$$

$$= -2 + 1 * 5 * 1 = \underline{3}$$

$$f(\text{net}_3) = -1 = z$$




$$\underline{y_1^1} = y_1^0 + c y_1 z$$

$$= 1 + 1 * 1 * (-1)$$

$$= 1 - 1 = \underline{0}$$

$$W^{k+1} = W^k + c \Delta W \\ \Delta W = \text{opt}$$

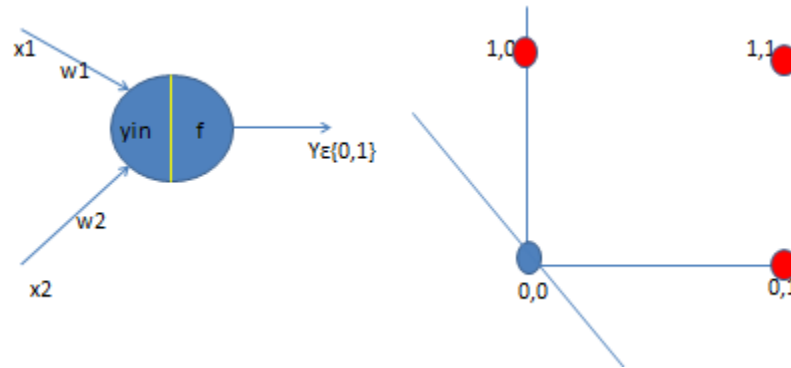
$$\underline{y_2^1} = y_2^0 + c y_2 z$$

iii.	<p>Define covariance ? For the given dataset, compute the covariance matrix.</p> <p>Marking scheme: <i>Define covariance? (1Marks)</i> <i>How is it different from correlation? (1Marks)</i> <i>Co variance matrix – 3marks</i></p> <p style="text-align: center;">PCA Process – STEP 1</p> <table style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>X_1</th> <th>X_2</th> <th></th> <th>X'_1</th> <th>X'_2</th> </tr> </thead> <tbody> <tr> <td>2.5</td> <td>2.4</td> <td rowspan="10" style="text-align: center; vertical-align: middle;">  </td> <td>0.69</td> <td>0.49</td> </tr> <tr> <td>0.5</td> <td>0.7</td> <td>-1.31</td> <td>-1.21</td> </tr> <tr> <td>2.2</td> <td>2.9</td> <td>0.39</td> <td>0.99</td> </tr> <tr> <td>1.9</td> <td>2.2</td> <td>0.09</td> <td>0.29</td> </tr> <tr> <td>3.1</td> <td>3.0</td> <td>$\Rightarrow \bar{X}_1 = 1.81$</td> <td>$\Rightarrow$</td> <td>1.29</td> <td>1.09</td> </tr> <tr> <td>2.3</td> <td>2.7</td> <td>$\bar{X}_2 = 1.91$</td> <td>\Rightarrow</td> <td>0.49</td> <td>0.79</td> </tr> <tr> <td>2.0</td> <td>1.6</td> <td></td> <td></td> <td>0.19</td> <td>-0.31</td> </tr> <tr> <td>1.0</td> <td>1.1</td> <td></td> <td></td> <td>-0.81</td> <td>-0.81</td> </tr> <tr> <td>1.5</td> <td>1.6</td> <td></td> <td></td> <td>-0.31</td> <td>-0.31</td> </tr> <tr> <td>1.2</td> <td>0.9</td> <td></td> <td></td> <td>-0.71</td> <td>-1.01</td> </tr> </tbody> </table> <p style="text-align: center; margin-top: 20px;"> $\text{cov} = \begin{bmatrix} 0.616555556 & 0.615444444 \\ 0.615444444 & 0.716555556 \end{bmatrix}$ </p>	X_1	X_2		X'_1	X'_2	2.5	2.4		0.69	0.49	0.5	0.7	-1.31	-1.21	2.2	2.9	0.39	0.99	1.9	2.2	0.09	0.29	3.1	3.0	$\Rightarrow \bar{X}_1 = 1.81$	\Rightarrow	1.29	1.09	2.3	2.7	$\bar{X}_2 = 1.91$	\Rightarrow	0.49	0.79	2.0	1.6			0.19	-0.31	1.0	1.1			-0.81	-0.81	1.5	1.6			-0.31	-0.31	1.2	0.9			-0.71	-1.01
X_1	X_2		X'_1	X'_2																																																							
2.5	2.4		0.69	0.49																																																							
0.5	0.7		-1.31	-1.21																																																							
2.2	2.9		0.39	0.99																																																							
1.9	2.2		0.09	0.29																																																							
3.1	3.0		$\Rightarrow \bar{X}_1 = 1.81$	\Rightarrow	1.29	1.09																																																					
2.3	2.7		$\bar{X}_2 = 1.91$	\Rightarrow	0.49	0.79																																																					
2.0	1.6				0.19	-0.31																																																					
1.0	1.1				-0.81	-0.81																																																					
1.5	1.6				-0.31	-0.31																																																					
1.2	0.9				-0.71	-1.01																																																					

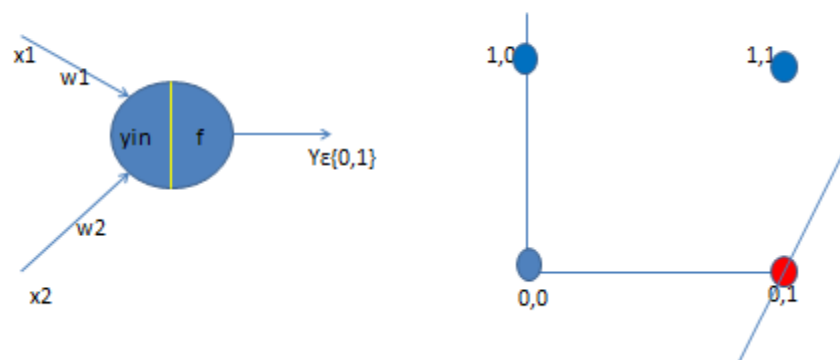
B	Solve any One	10 marks each
----------	----------------------	----------------------

i.	<div style="border: 1px solid black; padding: 10px;"> <p>Explain Linear Separability problem? Solve XOR linear separability problem using McCulloch Pitt Model ?</p> <p>Marking scheme: <i>Definition of Linear separability problem – 2 marks</i> <i>EXOR Solution of Linear separability problem – 8 marks</i></p> </div>
----	--

Gates	output	Inputs	Weights	yin	Threshold
NOR	y	x1 x2	W1 w2		
	1	0 0	-1 -1	0	≥ 0
	0	0 1	-1 -1	-1	< 0
	0	1 0	-1 -1	-1	< 0
	0	1 1	-1 -1	-2	< 0



Gates	output	Inputs	Weights	yin	Threshold
ANDNOT	y	x1 x2	W1 w2		
	0	0 0	1 -1	0	< 1
	0	0 1	1 -1	-1	< 1
	1	1 0	1 -1	1	≥ 1
	0	1 1	1 -1	0	< 1



EXOR

x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0

Solve it using the McCulloch Pitts model

$$y = x_1 \bar{x}_2 + \bar{x}_1 x_2$$

$$y = z_1 + z_2$$

$$z_1 = x_1 \bar{x}_2, z_2 = \bar{x}_1 x_2, y = z_1 + z_2,$$

- ii. What is the role of radial basis function in separating nonlinear patterns? Explain with XOR Example.

Marking scheme:

RBF architecture – 2 marks

Covers theorem – 2 marks

EXOR solution – 6 marks

University of Mumbai

Examination June 2021

Examinations Commencing from 1st June 2021

Program: Computer Engineering

Curriculum Scheme: Rev2016

Examination: TE Semester VI

Course Code: CSDLO6022 and Course Name: Adv. Database System

Time: 2 hours

Max. Marks: 80

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	The type of file which provides fast access to records under certain search conditions. This organization is usually called a
Option A:	Heap File
Option B:	Hash File
Option C:	Sequential File
Option D:	Direct File
2.	In Multilevel Indexing the index which leaves some space in each of its blocks for inserting new entries is called
Option A:	Dynamic Multilevel Index
Option B:	Dense Index
Option C:	Primary Index
Option D:	Clustering Index
3.	When the hash field value of a record that is being inserted hashes to an address that already contains a different record. It is called as
Option A:	Indexing
Option B:	Hashing
Option C:	Collision
Option D:	Chaining
4.	Semi -join is generally used for unnesting -----subqueries.
Option A:	Not In
Option B:	All
Option C:	Not Exists
Option D:	Exists
5.	The processor that has task of running the query code, whether in compiled or interpreted mode, to produce query result.
Option A:	Runtime Database Processor
Option B:	Query Graphic Processor
Option C:	Parser Runtime Processor
Option D:	Query Optimizer Processor
6.	The algorithms that are suitable for sorting large files of records stored on disk that do not fit entirely in main memory

Option A:	Internal Sorting
Option B:	Secondary sorting
Option C:	Parser Sorting
Option D:	external Sorting
7.	Which type of expression is represented by Query Graph?
Option A:	Tuple Relational Calculus
Option B:	Simple Expressions
Option C:	Relational Algebra
Option D:	Relational Calculus
8.	The real use of the Two-phase commit protocol is
Option A:	Deadlock will not occur.
Option B:	Concurrency control can be avoided.
Option C:	Atomicity, i.e, all-or-nothing commits at all sites.
Option D:	Both Availability and Robustness.
9.	In log based recovery, log is a sequence of
Option A:	Filter
Option B:	Records
Option C:	Blocks
Option D:	Numbers
10.	In Distributed Database if transaction can read, but cannot update that data item. It is called as
Option A:	Read Lock
Option B:	Write Lock
Option C:	Upgradation Lock
Option D:	Down gradation Lock
11.	The Probability that the system can continue its normal execution according to the specification at a given point in time in spite of failures.
Option A:	Availability
Option B:	Scalability
Option C:	Reliability
Option D:	Check pointing
12.	'enum' keyword used to..
Option A:	Define data type range
Option B:	Define a class range
Option C:	Define relationship range
Option D:	Define a range for an attribute
13.	Which is not a consistency level of Document Database?
Option A:	Bounded-staleness
Option B:	Elastic
Option C:	Session
Option D:	Strong
14.	The Most Well-Known object oriented Databases

Option A:	SimpleDB
Option B:	eXist
Option C:	BaseX
Option D:	Objectstore
15.	The characteristic of Multimedia system
Option A:	High storage
Option B:	Both High storage and High data rates
Option C:	High data rates
Option D:	Low Storage
16.	Which data is primarily managed by vertical application in Mobile database?
Option A:	Individual
Option B:	Shared
Option C:	Public
Option D:	Private
17.	XML stands for
Option A:	Extensible Markup Language
Option B:	Extended Markup Language
Option C:	Extensive Markup Language
Option D:	Exhausted Markup Language
18.	Which is not a valid access control mechanism?
Option A:	Discretionary Access Control
Option B:	Mandatory Access Control
Option C:	Role Based Access Control
Option D:	Subjective Access Control
19.	In statistical database, a set of tuples of a relation (table) that satisfy some selection condition called as
Option A:	Interinstance
Option B:	Population
Option C:	Infer
Option D:	Integrity
20.	Time based SQL injection attack is called..
Option A:	Initial Exploitation
Option B:	Inline Comments
Option C:	Quick detection
Option D:	Blind SQL Injection

Q2	Solve any Four out of Six	5 marks each
A	Compare B-Tree and B+ Tree with respect to their structure, advantages and disadvantages.	
B	Explain in detail Communication of Two Phase Commit Protocol.	
C	Explain Two Phase Locking Protocol in Distributed Database.	
D	How to manage continuous data in Spatial data models?	
E	Explain how authorization will be a Database Security issue? Give alternate solution in authorization.	
F	Explain FLWR expression in XML with an example.	

Q3	Solve any Two Questions out of Three	10 marks each
A	<p>Consider the following recursive DTD.</p> <pre><!DOCTYPE parts [<!ELEMENT parts (part+)> <!ELEMENT part (name, subpartinfo*)> <!ELEMENT subpartinfo (part, quantity)> <!ELEMENT name (#PCDATA)> <!ELEMENT quantity (#PCDATA)>]></pre> <p>a. Give a small example of data corresponding to the above DTD. b. Show how to map this DTD to a relational schema. You can assume that part names are unique, that is, wherever a part appears, its subpart structure will be the same.</p>	
B	Discuss how Multimedia Databases used in Mobile Databases? Explain what dirty data in terms of multimedia databases is. Consider any one type of multimedia databases.	
C	Discuss in detail Distributed Transaction Management with an example.	

University of Mumbai

Examination June 2021

Examinations Commencing from 1st June 2021

Program: Computer Engineering

Curriculum Scheme: Rev2016

Examination: TE Semester VI

Course Code: CSDLO6022 and Course Name: Adv. Database System

Time: 2 hours

Max. Marks: 80

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

Q.3. A

```
<!DOCTYPE parts [  
<!ELEMENT parts (part+)>  
<!ELEMENT part (name, subpartinfo*)>  
<!ELEMENT subpartinfo (part, quantity)>  
<!ELEMENT name ( #PCDATA )>  
<!ELEMENT quantity ( #PCDATA )>  

```

a. Give a small example of data corresponding to the above DTD.

b. Show how to map this DTD to a relational schema. You can assume that part names are unique, that is, wherever a part appears, its subpart structure will be the same.

Ans .

```
<parts>  
  <part>  
    <name> bicycle </name>  
    <subpartinfo>  
      <part>  

```

```
<subpartinfo>
  <part>
    <name> gear </name>
  </part>
  <qty> 3 </qty>
</subpartinfo>
<subpartinfo>
  <part>
    <name> frame </name>
  </part>
  <qty> 1 </qty>
</subpartinfo>
</part>
</parts>
```

ANS-B

Show how to map this DTD to a relational schema.

part (partid,name)

subpartinfo (partid, subpartid, qty)

Attributes partid and subpartid of subpartinfo are foreign keys to part.

University of Mumbai

Examination June 2021

Examinations Commencing from 1st June 2021

Program: Computer Engineering

Curriculum Scheme: Rev2016

Examination: TE Semester VI

Course Code: **CSDL06023** and Course Name: **Enterprise Resource Planning**

Time: 2 hour

Max. Marks: 80

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	What is at the heart of any ERP systems?
Option A:	Information
Option B:	Employees
Option C:	Customers
Option D:	Database
2.	One common mistake that companies resort when selecting an ERP system is
Option A:	Non biased selection
Option B:	Over-emphasis on system cost
Option C:	Complete set of requirements
Option D:	Not relying on vendor demos
3.	The first phase of ERP implementation life cycle is ,
Option A:	Maintenance
Option B:	Training
Option C:	Package selection
Option D:	Project planning
4.	Which of the following is the first step of accounting flow process
Option A:	Posting
Option B:	Trial balance
Option C:	Transaction
Option D:	Voucher entry
5.	What is OLAP
Option A:	Online analytical processing
Option B:	Online analytical programming
Option C:	Online analysis
Option D:	Offline analytical processing
6.	_____ module enables the enterprise to marry technology with business process to produce integrated solution
Option A:	CRM
Option B:	Manufacturing
Option C:	Procurement

Option D:	Supply chain
7.	Which is the module of Extended ERP?
Option A:	Accounting
Option B:	Procurement
Option C:	HR
Option D:	BI
8.	Which of the following are not fundamental technologies in ERP
Option A:	Data Mining
Option B:	BI
Option C:	CRM
Option D:	AI
9.	“Make to Stock” is based on -----
Option A:	Manufacturing Capacity
Option B:	Raw material availability
Option C:	Forecasted Demand
Option D:	Current Demand
10.	What does KPI refers to
Option A:	Key process indicator
Option B:	Key performance indicator
Option C:	key product indicator
Option D:	key product information
11.	The most important step of ERP implementation is _____ phase
Option A:	Testing
Option B:	Gap Analysis
Option C:	Training
Option D:	Installing
12.	Which of the following is not myth about ERP?
Option A:	ERP means more work and procedures
Option B:	ERP makes many employees redundant
Option C:	ERP is sole responsibility of management
Option D:	ERP integrates and automates organization process
13.	Objective of SCM is to -----
Option A:	To improve the efficiency across whole supply chain
Option B:	To improve the responsiveness across whole supply chain
Option C:	To deliver improved value to the customers
Option D:	To improve the efficiency across whole supply chain and responsiveness across whole supply chain to deliver improved value to the customers
14.	What is the meaning of parallel adaption in ERP implementation
Option A:	Changing the system incrementally
Option B:	Migrate to new system at once
Option C:	Keep the ERP system and legacy system operate parallely

Option D:	Limiting the access to ERP system to desktop working environment
15.	Name the parts 1,2,3,4,5 in the diagram
Option A:	1.position marker , 2.Format , 3.Timing Patterns , 4.version , 5.Alignment Marker
Option B:	1.Timing Patterns, 2.Format , 3.position marker , 4.Alignment Marker , 5.Version
Option C:	1.position marker, 2.Version , 3.Timing Patterns , 4.Alignment Marker , 5.Format
Option D:	1.position marker, 2.Format , 3.Timing Patterns , 4.Alignment Marker , 5.Version
16.	What are several different types of software, which sit in the middle of and provide connectivity between two or more software applications?
Option A:	Middleware
Option B:	Enterprise application Integration
Option C:	Automated business processes
Option D:	e-business infrastructure
17.	What is OLAP
Option A:	Online analytical processing
Option B:	Online analytical programming
Option C:	Online analysis
Option D:	Offline analytical processing
18.	Which of the following statement is true,
Option A:	DSS supplements MIS
Option B:	MIS replaces DSS
Option C:	DSS and MIS are distinct
Option D:	DSS replaces MIS
19.	Big bang implementation approach should be used for.....
Option A:	Small, simple organization
Option B:	Centralized organization
Option C:	Big organization
Option D:	Any type of organization
20.	The sequence of the stages of PLC is....
Option A:	Introduction-growth-maturity-decline
Option B:	Introduction-growth-decline-maturity
Option C:	Introduction-maturity-growth-decline
Option D:	Introduction-maturity-decline-growth

Q2 (20 Marks)	Solve any Four out of Six	5 marks each
A	Enumerate challenges in ERP Implementation?	
B	List different ERP related technologies and brief any one?	
C	What are different stages of ERP implementation life cycle?	
D	Write different types of EAI, brief any one.	
E	What is difference between data warehouse and data mart?	
F	What are OLAP and OLTP?	

Q3. (20 Marks)	Solve any Two Questions out of Three	10 marks each
A	What is extended ERP?	
B	What are different ERP data security technologies?	
C	Write short note on , "Oracle ERP Implementation at Maruti Suzuki"	

University of Mumbai

Examination June 2021

Examinations Commencing from 1st June 2021

Program: Computer Engineering

Curriculum Scheme: Rev2016

Examination: TE Semester VI

Course Code: **CSDLO6023** and Course Name: **Enterprise Resource Planning**

Time: 2 hour

Max. Marks: 80

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

University of Mumbai
Examination June 2021

Examinations Commencing from 1st June 2021

Program: Computer Engineering

Curriculum Scheme: Rev2016

Examination: TE Semester VI

Course Code: CSDLO6024 and Course Name: Advanced Computer Network

Time: 2 hour

Max. Marks: 80

=====

=====

=====

Q1.	Choose the correct option for following questions. All the Questions are compulsory and carry equal marks
1.	How many layers SONET has?
Option A:	7
Option B:	4
Option C:	5
Option D:	2
2.	_____ in SONET corresponds to the physical layer of the OSI model.
Option A:	Photonic Layer
Option B:	Section Layer
Option C:	Line Layer
Option D:	Path Layer
3.	_____ is a repeater, that takes an optical signal and regenerates it.
Option A:	Multiplexer
Option B:	Amplifier
Option C:	Regenerator
Option D:	Router
4.	ATM stands for _____ in telecommunication standards.
Option A:	Asynchronous Transmission Mode
Option B:	Asynchronous Transfer Multiplexing
Option C:	Anonymous Transfer Mode
Option D:	Asynchronous Transfer Mode
5.	An ATM cell has header of _____ bytes and the payload are of _____ bytes
Option A:	5,48
Option B:	8,32
Option C:	1,8
Option D:	8,16

6.	AAL1 supports applications that transfer information at
Option A:	high bit rates
Option B:	low bit rates
Option C:	constant bit rates
Option D:	variable bit rates
7.	Which is simplest ATM switch?
Option A:	knockout
Option B:	crossbar
Option C:	banyan
Option D:	Batcher-Banyan
8.	IPv6 uses bit addresses
Option A:	64
Option B:	128
Option C:	32
Option D:	256
9.	The header length of an IPv6 datagram is
Option A:	10 bytes
Option B:	20 bytes
Option C:	60 bytes
Option D:	40 bytes
10.	X.25 is protocol suite for packet-switched data communication in
Option A:	WAN
Option B:	LAN
Option C:	MAN
Option D:	Universal
11.	How many layers X.25 has?
Option A:	Seven
Option B:	Three
Option C:	Four
Option D:	Eight
12.	Open Shortest Path First (OSPF) is also termed as
Option A:	Border gateway protocol
Option B:	Routing information protocol
Option C:	Error-correction protocol
Option D:	Link state protocol
13.	An intra-domain routing protocol RIP is based on routing.
Option A:	path vector
Option B:	link state
Option C:	distance vector
Option D:	distance code
14.	BGP stands for

Option A:	Border Gateway Protocol
Option B:	Border Gigabyte Protocol
Option C:	Broadcast Gateway Protocol
Option D:	Broadband Gateway Protocol
15.	In Distance Vector Routing, each node shares its routing table with its
Option A:	First node neighbors
Option B:	Immediate Neighbors
Option C:	Next lane Neighbors
Option D:	Distant Neighbors
16.	What is minimum size of RTP header?
Option A:	16 bytes
Option B:	8 bytes
Option C:	12 bytes
Option D:	32 bytes
17.	The variation in delay for packets belonging to same flow is called as
Option A:	Jitter
Option B:	Bandwidth
Option C:	Reliability
Option D:	Attenuation
18.	What will happen about delay when load on network reaches network capacity?
Option A:	Increases sharply
Option B:	Decreases sharply
Option C:	Remains constant
Option D:	Unpredictable
19.	MIB stands for in SNMP
Option A:	Management Information Broadcast
Option B:	Management Inspection Base
Option C:	Management Information Block
Option D:	Management Information Base
20.	SNMP process is run by manager which is host.
Option A:	server
Option B:	data
Option C:	client
Option D:	variable

Q2.	Solve any Four out of Six	5 marks each
A	<i>Explain advantages of SONET in details.</i>	
B	<i>Discuss throughput and jitter parameter in short.</i>	
C	<i>Explain concept of Resource Reservation Protocol (RSVP) in short.</i>	
D	<i>What are advantages and disadvantages of Asynchronous Transfer Mode (ATM)?</i>	

E	<i>What is the function of x 25 explain in detail?</i>
F	<i>Explain concept of SMI in short.</i>

Q3.	Solve any Two Questions out of Three	10 marks each
A	<i>Explain ATM Adaptation layers in detail.</i>	
B	<i>Compare IPV4 and IPV6.</i>	
C	<i>Explain H.323 protocol in detail.</i>	

University of Mumbai

Examination June 2021

Examinations Commencing from 1st June 2021

Program: Computer Engineering

Curriculum Scheme: Rev2016

Examination: TE Semester VI

Course Code: CSDLO6024 and Course Name: Advance Computer Network

Time: 2 hour

Max. Marks: 80

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