



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE

Matunga, Mumbai-400 019

[Autonomous Institute affiliated to University of Mumbai]

EXAMINATION	Mid Semester Test October 2022	DATE OF EXAM	October 3, 2022
SEMESTER & PROGRAM	VII B.Tech	TIME	11:00AM-12:30PM
TIME ALLOWED	1.5Hr	MARKS	40
COURSE NAME – (CODE)	Entrepreneurship Development (R4PE4601S)		

Instructions

- Figures to the right indicate full marks.
- Attempt **all** questions.
- Assume additional data if required

Q.01	a) Discuss Principles of Effectuation in detail. b) Discuss customer adoption pattern in detail considering suitable example.	5	CO1
		5	CO2
Q.02	a) Explain different types of Market with appropriate examples. b) What do you understand by Entrepreneurial styles? Discuss different entrepreneurial styles with suitable examples.	5	CO2
		5	CO1
Q.03	Let L&T is the contractor for the Government project of constructing a 10Km long bridge; there is a reward for early completion of the project from stipulated time period of 1.5 years, also a penalty for delay in completion. The speed of average work carried out by L&T is 200m per day during normal traffic conditions and 500m per day during minimum traffic conditions. Now unfortunately we have a lockdown situation happened in middle of the completion of the project. Consider the above situation and implement the design thinking process.	10	CO2
Q.04	a) How do you explain the importance of segmentation, marketing and positioning for your product as an entrepreneur. b) Describe the process and outcome of the idea generation.	5	CO2
		5	CO1



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE

Matunga, Mumbai-400 019
Autonomous Institute affiliated to University of Mumbai

EXAMINATION	Mid Semester Examination (MST) Oct 2022	DATE OF EXAM	4 th Oct 2022
SEMESTER & PROGRAM	Sem-VII, B Tech (Information Technology)	TIME	2:00 pm to 3:30 pm
TIME ALLOWED	1.5 HRS.	MARKS	40
COURSE NAME - (CODE)	Data Mining and Data Warehousing (R4CO4001T)		

Instructions

1. All questions carry equal marks.
3. Figures to the right indicate full marks.
4. Make assumptions wherever necessary.

Q.1 a Find the outliers and extreme values, if any, for the following data set, and draw the box plot.. Mark any outliers with an asterisk and any extreme values with an open dot. 03 CO1

24 ,27, 19, 21, 48, 34, 22

- b How can we perform Data Characterization using OLAP operations? Explain with example 03 CO1
- b List the Parametric and Non Parametric methods used as Data Reduction Strategy 02 CO1
- c How can concept hierarchy be generated from Binning? 02 CO1

Q.2 a Suppose that a data warehouse consists of the four dimensions *date*, *spectator*, *location*, and *game*, and the two measures *count* and *charge*, where *charge* is the fare that a spectator pays when watching a game on a given date. Spectators may be students, adults, or seniors, with each category having its own charge rate. 04 CO1 CO3

(a) Draw a star schema diagram for the data warehouse.

- b For the case study in Q.2 a, show the bit map index used for the cube. Make assumptions if necessary. How will this index be useful for aggregation operations? 03 CO1
- c For the following data cube , illustrate the 0-d, 1-d and 2-d cuboids. 03 CO1

A	B	C	M
1	1	1	50
2	1	2	60
3	3	1	100
4	5	3	80
5	5	2	120

Q.3 a A database has four transactions.

05 CO2

Cust_ID	TID	Items bought
01	T100	A,B,C,D
02	T200	C,B,E,F,D
01	T300	E,B,D,F
03	T400	D,B,C

What is the maximum size of frequent itemsets that can be extracted (assuming minsup > 0)?

BEFD 2

Write an expression for the maximum number of size-3 itemsets that can be derived from this data set.

BCD 3

Find an itemset (of size 2 or larger) that has the largest support.

BED 2

b How is attribute oriented induction performed? Write the steps

05 CO1

Q.4 a An educational psychologist wants to use association analysis to analyze test results. The test consists of 100 questions with four possible answers each.

05 CO2

CO3

How would you convert this data into a form suitable for association analysis? In particular, what type of attributes would you have and how many of them are there?

b Consider the following table

05 CO2
CO3

Instance	A1	A2	A3	Target class
1	T	T	2.0	+
2	T	T	6.0	+
3	T	F	5.0	-
4	F	F	4.0	+
5	F	T	7.0	-
6	F	T	5.0	-
7	F	F	8.0	-
8	T	F	7.0	+
9	F	T	5.0	-

A1
4 5
T F

- i) What is the entropy of this collection of training examples ?
- ii) What are the information gains of a_1 and a_2 relative to these training examples?

What is the best split among (a_1, a_2) according to the information gain?

$$-\frac{4}{9} \log_2 \frac{4}{9} - \frac{5}{9} \log_2 \frac{5}{9}$$



$$-1.1699$$

$$0.991076$$



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE
[Central Technological Institute, Maharashtra State]
Matunga, Mumbai-400 019

SEMESTER EXAMINATION
SEMESTER & PROGRAM

Mid Semester Examination-Oct 2022
Sem-VII Final Year. B.Tech. (IT)

DATE :
TIME

06/10/2022
2.00 PM to
3.30 PM

TIME ALLOWED

1.30 Hr

MARKS

40

COURSE (Course Code)

Cyber Security (R4IT4002T)

Q1. a Solve RSA

5M CO1

P=193,q=17 If Plaintext ='XY' what will be the ciphertext ?Show process of encryption and decryption of input.XY are last two digits of your roll no.

Q1. b Illustrate the process of hash functions.

5M CO1

Q2. a Illustrate OWASP top 3 attacks .

5M CO2

b Demonstrate version Rollback attack in SSL.

5M CO2

Q3. a In the exercise involving corrupted cipher text when using block modes where you changed the single bit of the 29th byte in the encrypted file (encrypted using AES-128), answer the following.

i. What is the number of the cipher block (assume blocks start from 1) corrupted?

ii. For different modes ECB, CBC, CFB, and OFB, specify which plaintext blocks and how many bytes within that block were corrupted?

Q3. b Show the process of AES single Round.

6M CO1

Q4. a Justify the process of TCP SYN flood attack.

4M CO2

Q4 b Justify the requirements of Hash function.

4M CO1

Q4 c Differentiate between active attack and passive attack

2M CO1



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE

Matunga, Mumbai-400 019
Autonomous Institute affiliated to University of Mumbai

EXAMINATION	Mid Semester Examination (MST) October 2022	DATE OF EXAM	08 th October 2022
SEMESTER & PROGRAM	Sem-VII, B Tech Computer Engineering and IT	TIME	11:00 am to 12:30 pm
TIME ALLOWED	1.5 HRS.	MARKS	40
COURSE NAME - (CODE)	PE-II : Natural Language Processing (R4CO4204T)		

- Instructions 1. All questions carry equal marks.
 2. Figures to the right indicate full marks.

- Q.1** a. Briefly, explain the morphology operations: compounding, derivation, inflection. Given the root **snow** give an example each of the result of applying the three morphology operations. 04M CO2
- b. Write regular expressions for the following languages: You may use python notation or the minimal "algebraic" notations but make sure to say which one you are using. 03M CO1
1. the set of all alphabetic strings.
 2. the set of all lowercase alphabetic strings ending in a b. ab
- c. Distinguish between homonymy and polysemy and give an original example of each indicating the meaning in each case. 03M CO1
- Q.2** a. What are the ambiguities observed at different levels of classical NLP model. Justify with examples for each 06M CO1
- b. What is Morphological Parsing? Explain steps of morphological parser. 04M CO2
- Q.3** Consider the following corpus with following 4 sentences in it. 10M CO2
1. <s> ¹ ² ³ ⁴ ⁵ ⁶ ⁷ ⁸ ⁹ ¹⁰ ¹¹ three friends meeta priya and geeta are reading book </s>
 2. <s> ¹² ¹³ ¹⁴ meeta is reading malgudi days </s>
 3. <s> ¹⁵ ¹⁶ priya is reading a detective book </s>
 4. <s> ¹⁷ ¹⁸ ¹⁹ geeta is reading a book by rk narayan </s>
- a. Calculate a Bi-gram language model for given corpus.
- b. Consider the bi-gram model from Q3a and calculate probability for sentence as:

$$P(<\text{s}> \text{meeta} \text{ is reading a book} </\text{s}>)$$

- c. Given the same language model as Q3a. and Calculate the perplexity of sentence

< s > meeta is reading a book < /s >

- d. Given the same language model as Q3a. and calculate probability for the following sentence with Add-1 smoothing:

P(< s > akash is reading a story book < /s >)

Q.4 a. Consider following sentences :

05M CO2

- earnings growth took a back seat
- a small building in the back
- a clear majority of senators back the bill
- Dave began to back toward the door
- enable the country to buy back debt
- I was twenty-one back then

What is the Part of Speech for the word "**back**" in every sentence?
Also mention the PoS tag for them in Penn Tree Bank PoS tags.

- b. Explain Interpolated Kneser-Ney Smoothing technique? Why is it beneficial better for language modeling tasks?

05M CO2



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE

Matunga, Mumbai-400 019

[Autonomous Institute affiliated to University of Mumbai]

EXAMINATION	End Semester Examination November/December 2022 [odd term 2022-23]	DATE OF EXAM	December 1, 2022
SEMESTER & PROGRAM	VII sem Open Elective	TIME	11.00 AM to 02.00 PM
TIME ALLOWED	3 hrs	MARKS	100
COURSE NAME – (CODE)	Entrepreneurship Development (R4PE4601S)		

Instructions

1. Figures to the right indicate full marks.
 2. Attempt **all** questions.
 3. Assume additional data if required

- | | | | | |
|------------|----------|---|-----------|------------|
| Q.1 | a | Illustrate and explain all the elements of a Lean Canvas | 10 | CO2 |
| | b | Elaborate the concept of Idea generation through various stages involved and Business outcome | 10 | CO3 |
| Q.2 | a | What is MVP? Define the 5 steps involved to define your MVP with example | 10 | CO4 |
| | b | Compare and contrast the business Model Canvas to the Lean Model Canvas. | 05 | CO2 |
| | c | Discuss personas in detail | 05 | CO2 |
| Q.3 | a | What is the meaning of Marketing Channels, and Why is It Important? | 05 | CO1 |
| | b | Discuss different qualities of entrepreneurs with appropriate examples considering they are action oriented, people oriented or with deep skill | 05 | CO1 |
| | c | What is value proposition and discuss in detail value proposition canvas | 10 | CO4 |

- Q.4** a What is leveraging social media? 05 CO4
- b Discuss the three major influences on pricing decision. Distinguish between short run and long run pricing decision 10 CO3
- c How you will describe Entrepreneurship? relate entrepreneur with their flow in life using suitable examples 05 CO1
- Q.5** a Explain Sales Planning and its importance 04 CO4
- b Discuss strategies to come up with creating solutions for the identified problem 08 CO3
- c What is competition? why it is important? Define in terms of how it is relevant in favour of customers 04 CO2
- d Differentiate between debt and equity financing 04 CO3



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE

Matunga, Mumbai-400 019

Autonomous Institute affiliated to University of Mumbai

EXAMINATION	End Semester Examination, Even Sem December 2022	DATE OF EXAM	03/12/2022
SEMESTER & PROGRAM	Sem-VII, B Tech Computer Engg and IT	TIME	11:00 am – 2:00 pm
TIME ALLOWED	3 HRS.	MARKS	100
COURSE NAME – (CODE)	PE-II : Natural Language Processing (R4CO4204T / R4IT4204T)		

Instruction

1. All questions are compulsory and carry equal marks.
 2. Figures in last column show COs mapped and 2nd last shows full marks (M).
 3. Mention valid assumptions in answers if required.

Q.1	<p>a. Distinguish between semantics, pragmatics and discourse with examples.</p> <p>b. What are morphemes? Brief about inflection, derivation and compounding in morphology with couple examples each</p> <p>c. You have the following three documents - <i>D1, D2, D3:</i></p> <p><i>D1: Natural language processing is becoming important since soon we will begin talking to our computers.</i></p> <p><i>D2: If computers understand natural language they will become much simpler to use.</i></p> <p><i>D3: Speech recognition is the first step to build computers like us.</i></p> <p>Answer the following with respect to the above set of 3 documents after text normalization (stop word removal etc) has been done on all 3 documents. Assume no chunking is done.</p> <ol style="list-style-type: none"> 1) List the stop-words present in the documents. 2) What is the vocabulary V? 3) What are the number of bigrams and trigrams in <i>D2</i>? Display all of them. 	6M	CO 1
Q.2	<p>a. What is the need of smoothing for language modeling? Explain with Laplace Smoothing and the drawbacks of the said method.</p> <p>b. Explain CKY Parsing Algorithm with example.</p>	10 M	CO 2
Q.3	<p>a. Brief about any four relations held between word senses in WordNet.</p> <p>b. Draw two different parse trees possible due to ambiguous structure for the following sentence and justify it. Sentence: <i>One morning I shot an elephant in my pajamas.</i></p>	4M	CO 3

NP

VP

PP

c.

Consider the following HMM model created for PoS tagging. As given below, E is the Emission Probability matrix and S is the State Transition Probability matrix.

The test sentence is : Time flies like an arrow

1. Mention and Justify the number of possible tag-sequences.
2. Find the correct PoS tag sequence using Viterbi Algorithm.

Emission Probability matrix E

	Time	flies	like	an	arrow
VB	0.1	0.2	0.2	0	0
NN	0.1	0.1	0	0	0.1
IN	0	0	0.25	0	0
DT	0	0	0	0.5	0

VB DT NN 3
VB IN 2
VB NN
NN VB
NN NN

NN NN

State Transition Probability matrix S

	VB	NN	IN	DT	<E>
<S>	0.2	0.8	0	0	0
VB	0	0.3	0.2	0.5	0
NN	0.4	0.5	0.1	0	0
IN	0	0.75	0	0.25	0
DT	0	1	0	0	0

Q.4 a. What do you mean by Word Sense Disambiguation (WSD)? Give an example. Explain any dictionary-based approach to achieve WSD task.

10 M CO 3

b. Brief the different types of ambiguities to be addressed in NLP and identify the ambiguities present in the following sentence.
" Flying planes can be dangerous. "

10 M CO 1

Q.5 a. Differentiate between the open-class words and closed-class words with examples.

5M CO 1

b. Give intuition about the vector semantics following the distributional hypothesis. Explain the word-to-word co-occurrence matrix in the said context with an example.

10 M CO 4

c. Suppose you have the following two 4-dimensional word vectors for two words w1 and w2, respectively:
 $w_1 = (0.2, 0.1, 0.3, 0.4)$ and $w_2 = (0.3, 0, 0.2, 0.5)$
What is the cosine similarity between w1 and w2? Are these words similar or dissimilar?

5M CO 4

Page 2/2



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE
[Central Technological Institute, Maharashtra State]
Matunga, Mumbai-400 019

SEMESTER EXAMINATION
SEMESTER & PROGRAM

Examination-Dec 2022
Sem-VII Final Year. B.Tech. (IT)

DATE : 08 /12/2022
TIME : 11:00 AM - 12:00 PM

TIME ALLOWED

COURSE (Course Code)

3.00 Hr

Cyber Security (R4IT4002T)

MARKS 100

Q.No	Question	Mark s	CO
Q.1	a. Justify the importance of Burpsuite in analysis of android application. b. Demonstrate in detail android adb commands which are related with process, package ,intent with example. c. Illustrate in detail IPSEC protocol	6 8 6	3 3 3
Q.2	a. Illustrate pretty good privacy protocol b. Illustrate message digest process of SHA . c. Illustrate meltdown attack in detail.	6 8 6	2 1 4
Q.3	a. Illustrate malware analysis using machine learning algorithms b. Differentiate in detail host based intrusion detection system and network based intrusion detection system c. Develop a program for reading SOAP request and SOAP response with example	6 6 8	3 3 3
Q.4	a. For all scheme, provide answer 1) Whether it allows Bob to decrypt messages from Alice. Be sure to specify what Bob has to do in order to decrypt the message and make sure that it came from Alice and has not been tampered with. Remember, Bob can only use Encryption /Decryption once and Hash once. 2) Which example below provides confidentiality, integrity.	6	1

Alice $E(N) \cdot E(M)$ Bob

page 14

2. E(M)
 3. E(H(M))
 4. H(M)

X. Illustrate in detail rounds of AES 6 1

$$040 \rightarrow 37 \mid 41 \quad 8 \quad 1$$

c. Justify the need and Illustrate fermat theorem by choosing prime number nearest to your last three digits of roll no.

- Q.5 a Justify the purpose of following IT Acts 6 3
 1. Section 503
 2. Section 499
 3. Section 463 email spoofing, forgery of e-records
 4. Section 383 web jacking
- b Justify different penetration tools used in kali linux with purpose 8 4
- c Illustrate virus attachment in detail 6 3

$$a^{\frac{p-1}{2}} \mod p = 1$$

$$3^{40} \mod 41$$

$$a^{40} \mod p = 1$$

$$81 \mod 41$$

$$a^{36} \mod p = 1$$

$$40^{13} \times 3 \mod 41$$

$$1^{13} \times 3 \mod 41$$

$$3 \mod 41$$



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE

Matunga, Mumbai-400 019
Autonomous Institute affiliated to University of Mumbai

EXAMINATION	MST February 2023	DATE OF EXAM	14 th March 2023
SEMESTER & PROGRAM	B Tech Semester-VIII, (Information Technology)	TIME	10:00 am to 11:30 am
TIME ALLOWED	1.30 HRS.	MARKS	40 Marks
COURSE NAME - (CODE)	Human Computer Interaction(R4IT4003S)		

Instructions

1. All questions are compulsory and carry equal marks.
2. Figures to the right justification indicate full marks.
3. Critical thinking with proper short justification carry full marks.

Q.1 Principles for usability

[CO1]

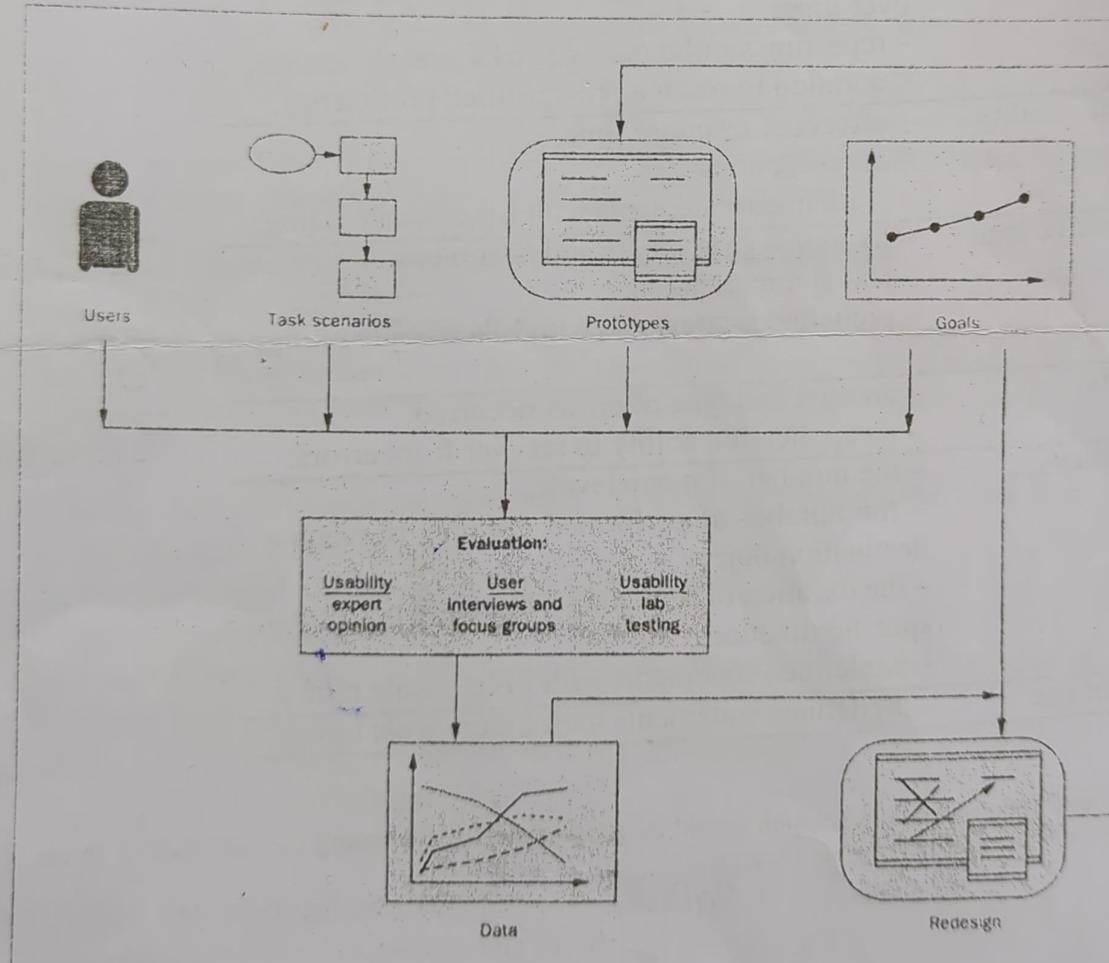
- repeatable design for usability relies on maximizing benefit of one good design by abstracting out the general properties which can direct purposeful design → *semantic web?*
- The success of designing for usability requires both creative insight (new paradigms) and purposeful principled practice → ? *govt. websites*

Justify the above usability principle with proper case study.

Q.2

[CO1]

(10)



The above figure shows an overview of rapid-prototyping methodology. It shows how end-users, task scenarios, proto-types, and project-team usability goals are used to evaluate and redesign a prototype. A decision to retain or redesign a UI prototype is made depending on end-users' needs and final evaluation results. Justify it with proper case study.

ISLC + Case study

Q.3 Usability attributes are given in table below. Rate your mobile phone applications with proper justification for measures and evaluation. [CO2] (10)

Attribute	Measures
efficiency	<ul style="list-style-type: none"> - duration spent on each screen, - duration to complete task, - user's error rate,
satisfaction	<ul style="list-style-type: none"> - predefined statements with Likert-scale rating,
effectiveness	<ul style="list-style-type: none"> - number of successfully completed tasks, - number of steps required to complete a task, - number of taps related to an app usage, - number of taps unrelated to an app usage, - number of times that a back button was used;
learnability	<p>first-time:</p> <ul style="list-style-type: none"> - number of attempts to solve a task, - number of assists during performing a task, - number of errors performed by a user; <p>over time:</p> <ul style="list-style-type: none"> - repeating similar pairs of tasks in each session, - duration to reach a pre-specified proficiency,
memorability	<ul style="list-style-type: none"> - effects of response time, - duration of pauses, - predefined statements with Likert-scale rating;
cognitive load	<ul style="list-style-type: none"> - fixations, gaze points and heat maps, - text summarization, - predefined statements with Likert-scale rating;
errors	<ul style="list-style-type: none"> - amount and type of errors occurred, - an application ability to recover from errors;
simplicity	<ul style="list-style-type: none"> - the number of menu levels, - the number of performed gestures to reach a destination object - the duration of searching a button to perform a specific function, - predefined statements with Likert-scale rating;
ease of use	<ul style="list-style-type: none"> - predefined statements with Likert-scale rating,

A M Comments
0/10

Zomato
Twitter

UTS
Classroom

Q.4 Choose 3 evaluation techniques which would be appropriate for evaluating the interface of an air traffic control system. Justify your choices. [CO3]

① Assumptions ③aApply ICA - SOAR, MT-R (10)

② Domain Goals / Tasks ③^{**} evaluate using Norman

③ c GOMS? empirical testing a must



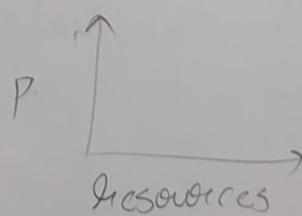
VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE
Matunga, Mumbai-400 019
Autonomous Institute affiliated to University of Mumbai

EXAMINATION	MST - March 2023	DATE OF EXAM	15 th March 2023
SEMESTER & PROGRAM	Sem-VIII, Final Year BTech Information Technology	TIME	10.00 am to 11:30 am
TIME ALLOWED	1.5 HR.	MARKS	40
COURSE NAME - (CODE)	PE-4 Advanced Database Engineering (R4IT4402S)		

Instructions

1. All questions are compulsory and figures to the right indicate full marks.
2. Assume and mention necessary data wherever required.
3. Draw diagrams wherever necessary.
4. Write answers to each main question on a separate page.
5. Do not write anything on the question paper.

- Q.1 (a) Define 'network latency' and elaborate on two optimization strategies adopted by the parallel database systems to reduce the effects of network latency. (04M)
[CO1]
- (b) i. Derive Amdahl's law with respect to the parallel database systems. [CO1]
ii. Draw a linear graph (on X-Y plane) to show the maximum speedup possible if a database transaction is being executed five times on a parallel database system assuming 10%, 20%, 30%, 40%, and 50% part of that transaction is sequential in the respective execution. (04M)
- Q.2 Consider the database relation 'cloth_sales' (given on the 2nd page) that needs to be stored on a distributed database system. [CO1]
i. Perform vertical partitioning on the original relation.
(Show only relations without data) (01M)
ii. Create three horizontal partitions using hash-partitioning on the original relation.
(Show data) (06M)
iii. Write SQL queries to fetch all the records from the partitions created above.
(One query per partitioning type) (03M)
horizontal & vertical
- Q.3 Consider the horizontal partitions created by you in Q2. [CO1]
i. Perform parallel range-partitioning sort to find the top three popular types of cloths by their sold units.
(07M)
ii. Write a SQL query to answer the above question (after sorting). (01M)
iii. How do you achieve parallelism in this sorting operation ? (02M)



Q.4 (a) A 'data warehouse' is a repository (or archive) of information gathered from multiple sources, stored under a unified schema, at a single site. [05M] [CO2]

Describe five ways to preprocess the data before storing it in a data warehouse using suitable examples.

(b) Construct a cross-tab of the database relation 'cloth_sales' (given on the 2nd page) by cloth name and size for color 'light'. [05M] [CO2]

11.2

121
222
221

312
212

gp

then merge sort?

external merge sort? X

0, 1, 2

R₀

R₁

R₂

cloth_id	cloth_name	cloth_color	cloth_size	units_sold
111 2	dress	dark	medium	555 —
112 0	dress	dark	large	141 —
121 0	dress	light	medium	821 ✓
122 1	dress	light	large	24
211 2	shirt	dark	medium	55
212 0	shirt	dark	large	696
221 0	shirt	light	medium	987
222 1	shirt	light	large	370 —
311 2	pants	dark	medium	777
312 0	pants	dark	large	3
321 0	pants	light	medium	499 —
322 1	pants	light	large	0

h₀
312
112
321
212
121
221

3
141
499
696
821
987

h₁
322
122
222
24
370

h₂
211
111
311
55
555
777

[56, 556, 988]

0-100 101-600
601-1000



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE

Matunga, Mumbai-400 019

Autonomous Institute affiliated to University of Mumbai

EXAMINATION	Mid Semester Examination (MST) March 2023	DATE OF EXAM	16 th March 2023
SEMESTER & PROGRAM	Sem-VIII, Final Year B Tech (Computer Engineering/ Information Technology)	TIME	10 AM to 11:30 PM
TIME ALLOWED	1.5 HRS.	MARKS	40
COURSE NAME - (CODE)	Big Data Analytics (R4CO4011T/R4IT4011T)		

- Instructions
1. All questions carry equal marks.
 2. Figures to the right indicate full marks.
 3. Assume suitable data wherever necessary and give its justification.

Q.1 a. Write a Map Reduce code for the following input and output

Input:

(10)

Suppose we are working with two tables, Table 1 is employee table with emp_id as a primary key and dept_id as a foreign key. Table 2 is department table with dept_id as a primary key.

Employee Table

Emp_id	Emp_name	street	city	Emp_contact	Salary	Dept_id
101	jone	althan	Surat	1111111	20000	10001
102	cartin	udhna	Surat	2222222	15000	20001
103	krish	ajava	Vadodara	3333333	30000	20001
104	dhiru	ramnagar	Vadodara	8888888	36000	30001
105	om	althan	Surat	7777777	22000	30001
106	adi	vesu	Navsari	2323232	35000	10001
107	annant	shivnagar	Navsari	5555555	34000	10002
108	yogi	althan	Surat	8989898	25000	10002
109	muskan	vesu	Vadodara	9999999	18000	10001
110	rudra	kashi	hazira	1212121	31000	20001

Department Table

Dept_id	Dept_name	Dept_location
10001	Account	Surat
20001	Sales	Hazira
30001	Finance	Vadodara

10002	Marketing	Surat
-------	-----------	-------

Output

Name	Street	City	Contact	Salary	Department
rudra	Kasha	Hazira	1212121	31000	Sales

- Q.2 a. Which and how Big data tool will help in developing smart cities? Give justification. (05)
 b. How single point of failure problem is addressed in different versions of Hadoop? (05)
- Q.3 a. Write a Scala or Python script to count non-duplicate words from input file. (05)
 b. How lazy evaluation helps Spark to enhance the performance? Give suitable example.
- Q.4 a. Cite a situation in which Cassandra is preferred over MySQL (05)
 b. Describe CAP theorem (05)



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE

Matunga, Mumbai-400 019

Autonomous Institute affiliated to University of Mumbai

Instructions 1. Figures to the right indicate full marks.

EXAMINATION	Mid sem Semester Examination March 2023	DATE OF EXAM	17/3/2023
SEMESTER & PROGRAM	SemVIII- Cloud Computing	TIME	10:00 AM to 11:30 AM
TIME ALLOWED	1.5 HRS.	MARKS	40
COURSE NAME – (CODE)	Cloud Computing- R4IT012T		

Q.No.	Question	Marks	Co
Q1.	a. Illustrate cloud computing Architecture. b. List private cloud products.	8	2
Q2.	a. Differentiate between private and public cloud <input checked="" type="checkbox"/> b. Draw Microsoft Hyper V architecture	2	1
Q3.	a. Illustrate XEN hypervisor b. List different types of Virtualization	6	3
Q4.	a. Illustrate platform as service implementation b. Draw infrastructure as service implementation	6	2,3



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE
 Matunga, Mumbai-400 019
Autonomous Institute affiliated to University of Mumbai

EXAMINATION	ESE - May 2023	DATE OF EXAM	17/05/2023
SEMESTER & PROGRAM	Sem-VIII, Final Year BTech Information Technology	TIME	09.30 am to 12:30 pm
TIME ALLOWED	3.0 Hrs.	MARKS	100 - 21
COURSE NAME - (CODE)	PE-4 Advanced Database Engineering (R4IT4402S)		

- Instructions
1. All questions are compulsory and figures to the right indicate full marks.
 2. Assume and mention necessary data wherever required.
 3. Draw diagrams wherever necessary.

Q.1 (a) Distinguish between the following indices, used in a distributed database system, with the help of suitable examples: a global primary index and a global secondary index. (05M) [CO1]

(b) Elaborate on the deadlock-detection technique used in a distributed database system using a suitable example using a suitable example(s)/diagram(s). (05M) [CO1]

(c) Justify that 'the semi-join strategy, used in query processing across multiple data sources, is particularly advantageous when relatively few tuples of a relation contribute to the join', using a suitable example with sample data. (Compare the effective data transfer done in semi-join and normal-join) (10M) [CO1]

How is the Bloom filter used for efficient overapproximation of the semijoin result ?

Q.2 (a) Validate the following schedule of two transactions under the timestamp-ordering protocol: (10M) [CO1]

T ₂₅	T ₂₆
read(C); C := "from Mr.GM !"; read(B);	
read(A);	read(B); B := "to All "; write(B);
display(C)	read(A); write(C);
write(C); display(A+B+C)	A := "Best Wishes "; write(A); display(A+B)

103	1	Dr Shinde	CE IT
104	1	Dr. Samde	CE IT
102	1	Bose	FY
301	3	Dr. Kazi	EE
201	2	Panchal	EE
202	2	Faculty	phd
	1	Shinde	yes
	1	Kazi	yes
	1	Bose	No

Assume initial transaction-timestamps (TS) as follows: ('R' is 'read' and 'W' is 'write')

$$R_{TS}(A) = W_{TS}(A) = R_{TS}(B) = W_{TS}(B) = R_{TS}(C) = W_{TS}(C) = 0$$

$$TS(T_{25}) = 25 \text{ and } TS(T_{26}) = 26$$

Show the updated timestamps at each operation and whether that particular operation is allowed to execute or not (with reason).

- (b) In the distributed database system how do you ensure that all the nodes in which a transaction T executed must agree on the final outcome of the execution. Describe the complete execution of the two-phase commit (2PC) protocol for a transaction T executed on three nodes if one of the nodes cannot execute its portion of T successfully. (Assume that all three nodes are up and running throughout the execution of 2PC protocol) [10M] [CO1]

- Q.3 (a) i. Write an SQL query that computes a cross-tab on the database relation 'sales' (given on the 4th page) aggregating ~~revenue~~ values as per the ~~cloth size~~ ~~quarter~~. Show the output of the query. [CO2] (03M)
- ii. Write an SQL query with 'rollup' operation on the database relation 'sales' (given on the 3rd page). The empty/null values in the results must be replaced by some alias value. Show the output of the query. (Show only aggregating columns while writing on paper) [04M] (03M)
- iii. What is the limitation/restriction of 'rollup' and 'cube' operations and how do you overcome it ? Write an appropriate SQL query on the database relation 'sales' (given on the 3rd page) and show its output. [04M] (03M)
- (b) Consider the relations of the 'university' database schema (given on the 4th page): [CO2]
- Construct a bitmap index on attribute 'salary', dividing values into 4 ranges: below 50,000; 50,000 to below 75,000; 75,000 to below 1,00,000; and 1,00,000 and above. [02M]
 - Consider a query that requests all instructors in the 'Comp. Sci.' department with a salary of 1,25,000 or more. Outline the steps in answering the query using bitmaps, and show the final and intermediate bitmaps constructed to answer the query. [03M]
 - Find out the instructors from the departments other than the 'Comp. Sci.' using only a single bitmap. [02M]
- Find out the students from the 'Comp. Sci.' department using a Bloom filter. [03M]

- Q.4 Consider following relational schema of 'COMPANY' database:

PROJ (PNO, PNAME, BUDGET); EMP (ENO, ENAME); ASG (PNO, ENO, ROLE)
Assume that the company has at least one project, one employee, and one assignment; and PNO, ENO are the mandatory attributes.

- (a) i. Create a DTD schema. [05M]
- Give a sample XML representation. [04M]
- Consider a query to find the names of all the managers working on the projects with a budget more than 100K. Write a XQuery for the same displaying project name and manager names in the ascending order of projects. [03M]
- (b) Write following queries in SQL using SQL/XML (SQLX) functions to convert the relational data to the XML representation: [08M]
- Get project name and budget of a project number 'P202305' using xmlelement function.
 - Get assignment details with project number 'P202305' and employee number 'E20230517' using xmlelement function.

- iii. Get project details of a project using `xmlforest` function.
 iv. Get project details of a project using `xmlagg` function. (Include root tag)

~~Q.5(a)~~ Solve the following query on two relations 'r' and 's' using the MapReduce paradigm:

(10M)
[CO4]

$$r:C,s:D \nabla_{sum(s.E)}(r \bowtie_{r.A=s.B} s)$$

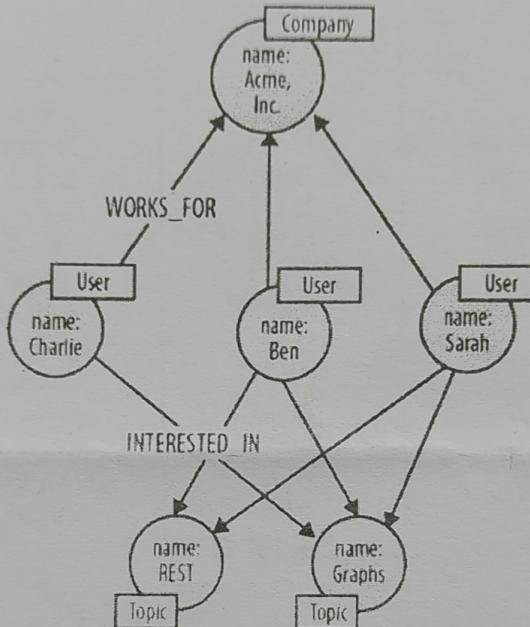
Assuming that both the relations are not partitioned, show the query processing diagrammatically with all the input and output data for map-nodes and reduce-nodes.

The solution should be generalized i.e. do not assume any specific data (consider data values as v_1, v_2, \dots, v_n) and the number of nodes can be anything.

(Hint to solve the query: think of multiple MapReduce steps !)

(b) Write following queries in Cypher query language for the given sample graph :

(10M)
[CO4]



- i. Create the graph - *wrong because Charlie → REST does not exist*
~~ii. Find users having common interest as that of 'Charlie'.~~
 iii. Show and delete the entire graph.

Given : database relations with sample data

'university' db	'department'	dept_id	dept_name	building
		BL	Biology	Watson
		CS	Comp. Sci.	Taylor
		EE	Elec. Eng.	Taylor
		FN	Finance	Painter
		HT	History	Painter
		MS	Music	Packard
		PH	Physics	Watson

'university' db	'instructor'	iid	iname	dept_name	salary
		10101	Katz	Comp. Sci.	75000
		12121	Wu	Finance	100000
		15151	Mozart	Music	35000
		22222	Einstein	Physics	175000
		32343	El Said	History	50000
		33456	Gold	Physics	87000
		45565	GM	Comp. Sci.	45000
		58583	Califieri	History	62000
		76543	Singh	Finance	80000
		76766	Crick	Biology	72000
		83821	Brandt	Comp. Sci.	105000
		98345	Kim	Elec. Eng.	40000

'university' db	'student'	sid	sname	dept_id
		00128	Zhang	CS
		12345	Shankar	CS
		19991	Brandt	HT
		23121	Chavez	FN
		44553	Peltier	PH
		45678	Levy	PH
		54321	Williams	CS
		55739	Sanchez	MS
		70557	Snow	PH
		76543	Brown	CS
		76653	Aoi	EE
		98765	Bourikas	EE
		98988	Tanaka	BL

'sales'	region	year	quarter	revenue
	north	2021	1	555
	north	2021	2	141
	north	2021	3	821
	north	2021	4	24
	north	2022	1	55
	north	2022	2	696
	north	2022	3	987
	north	2022	4	370
	south	2021	1	777
	south	2021	2	30
	south	2021	3	499
	south	2021	4	110
	south	2022	1	1234
	south	2022	2	2023
	south	2022	3	1705
	south	2022	4	999



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE

Matunga, Mumbai-400 019
Autonomous Institute affiliated to University of Mumbai

EXAMINATION	End Semester Examination 2023	DATE OF EXAM	17 th May 2023
SEMESTER & PROGRAM	Semester-VIII, B Tech (COMP)	TIME	9:30 am to 12:30 pm
TIME ALLOWED	03:00 HRS.	MARKS	100 Marks
COURSE NAME - (CODE)	Human Computer Interaction - (R4CO4403S)		

Instructions	1. All questions are compulsory and carry equal marks. 2. Figures to the right indicate full marks. 3. Assume suitable data wherever necessary		
Q.1	a) How to get to know the system users? Explain various methods adopted in user-centered design. What are the people directly or indirectly affected by a student registration system.	[CO1]	10
	b) Understanding users and their behavior is an important factor influencing user-interface design. An automatic syringe is designed to administer proper dose of medicine to the patient. Create a prototype user interface for the same that can set the dose (4 digits numeric) with minimal human error while setting the dose. Justify your design.	[CO1]	10
Q.2	a) A scenario is an idealized but detailed description of a specific instance of human-computer interaction (HCI). Scenarios specify how users carry out their tasks in a specified context. Write scenarios for purchasing an airline ticket. Note - Generate scenarios to cover a wide range of situations, not just the most common ones. Include problem-situations that will test the system concept, not just straight forward scenarios.	[CO2]	10
	b) What is the definition of usability as per ISO 9241 standard? Effective applications are both consistent within themselves and consistent with one another. Discuss this in context of Microsoft Office products.	[CO1, CO2]	10
Q.3	<p>a) The cognitive walkthrough is a formalized way of imagining people's thoughts and actions when they use an interface for the first time. During a cognitive walkthrough the evaluator needs to ask four questions as below</p> <ul style="list-style-type: none"> i. Is the effect of the action the same as the user's goal at that point? ii. Will users see that the action is available? iii. Once users have found the correct action, will they know it is the one they need? iv. After the action is taken, will users understand the feedback they get? <p>Given below is an action sequence for creating a customized voicemail message on an iPhone.</p> <ol style="list-style-type: none"> 1. Tap Voicemail. 2. Tap Greeting. 3. Tap Custom. 4. Tap Record and speak your greeting. 5. When you finish, tap Stop. 6. To listen to your greeting, tap Play. 7. To re-record, repeat steps 4 and 5. 8. Tap Save. 	[CO3]	10

ai 30
11-10

Page 1/2

	Imagine an iPhone interface and create a report of the cognitive walkthrough for the above mentioned task in context with the review questions		
	b Design an experiment to test whether adding color coding to an interface will improve accuracy. Identify your hypothesis, participant group, dependent and independent variables, experimental design, task and analysis approach.	[CO3]	10
Q.4	<p>a KLM (key-store-level) model predicts expert error-free task completion time (human performance) with interactive computing systems. Total predicted time for a task is given by the equation. $t_{EXECUTE} = t_K + t_P + t_H + t_D + t_M + t_R$</p> <p>What does each of the above timing represent?</p> <p>Develop a KLM model and predict time for the completion of the task "Change font and style for the word "KLM" to bold, Arial" using mouse only.</p>	[CO3]	10
	b Create a GOMS description of the task of photocopying a paper from a Journal. Discuss the issue of closure in terms of your GOMS description.	[CO3]	10
Q.5	<p>a Consider the activity of making a telephone call. Record the actions in an HTA diagram or textually. Start off simply, assuming you know the number to dial, but then add more complicated situations : finding the number in an address book, or what to do when the number is engaged.</p> <p>b There are four main translations involved in the interaction framework viz. articulation, performance, presentation and observation.</p> <ul style="list-style-type: none"> i. The compact disk player has a button for power off. However, its remote control does not have a power off button. ii. It is difficult in a command line interface to determine the result of copying and moving files in a hierarchical file system. iii. User is unable to figure out which switches from the bank to turn on to lit the front portion of a classroom. iv. The user is unable to know whether the voice recorder is in playing or recording state. <p>Specify in each of the above four cases which of the interaction framework translations are in effective.</p>	[CO3, CO4]	10



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE
[Central Technological Institute, Maharashtra State]
Matunga, Mumbai-400 019

SEMESTER EXAMINATION	ESE-MAY 2023	DATE OF EXAM	22/05/2023
SEMESTER & PROGRAM	VIII B.TECH(IT)	TIME	9.30AM - 12.30 PM
TIME ALLOWED	3 HRS.	MARKS	100
COURSE (CourseCode) :	Cloud Computing -R4IT4012T		

Instructions 1. Figures to the right indicate full marks.
 2. Assume suitable data wherever necessary.

Q. No.	Question	Mar ks	C
Q.1	a. Illustrate Esxi hypervisor architecture	8	2
	b. Write down the description of the following code.	6	2
	Vm vm = new Vm(vmid, brokerId, mips, pesNumber, ram, bw, size, vmm, new CloudletSchedulerTimeShared()); vmlist.add(vm); broker.submitVmList(vmlist); cloudletList = new ArrayList<Cloudlet>(); int id = 0; long length = 560000; long fileSize = 800; long outputSize = 900; UtilizationModel utilizationModel = new UtilizationModelFull();		
	c. Describe in detail Google Bigtable .	6	2
Q.2	a. Differentiate between Google file system and HDFS.	6	4
	b. Justify different terms in Cloud storage SLA.	8	5
	c. If your cloud providers server x has i)Two outages each year lasting two hours each for the cloud system ii)Also it has 21% chance of failure and if site is down 24 hours Calculate a) Internet connection availability of system b) Overall availability of cloud system	6	4
Q.3	a. Illustrate architecture of Openstack	6	1
	b. Private Cloud is deployed at VJTI. What kind of IAAS ,PAAS ,SAAS services can be started at VJTI?	8	1
	c. Justify different cloud scaling methods.	6	1

	X Justify Identity and Access Management (IAM) in cloud computing	8	5
	X Differentiate between paravirtualization and full virtualization	6	3
Q.4	a. Explain the Components of OpenNimbus Cloud .	6	1
	b. Describe in detail google app engine architecture	6	2
	c. Explain architecture of XEN server	8	2
Q.5	a Justify security in cloud computing	6	6
	b Write a program using google API which sets latitude and longitude	8	4
	c Convert following schema into Hbase table Where XYZ are last three digits of your rollno .	6	4

Kcomputer supercomputer	Spec{'core':'10302' , 'RMAX':'5042X'} Chipset {' Processor':'Sparc64' , 'RPEAK' :33XZ Y'}
Tianhe2 supercomputer	Spec{'core':'3,60,640' , 'RMAX':'67XYZ'} Chipset {'Processor':'IntelXeon' , 'RPEAK':'21ZXY'}
Seqioa supercomputer	Spec{'core':'5,10,640' , 'RMAX':'47XY'} Chipset {'Processor':'IntelXeonE3' , 'RPEAK':'87 8XYZ'}



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE

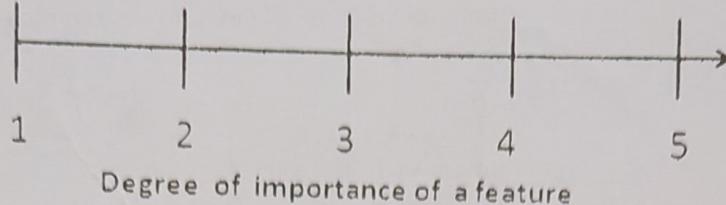
Matunga, Mumbai-400 019

Autonomous Institute affiliated to University of Mumbai

EXAMINATION		End Semester Examination 2023 (MAY 2023)	DATE OF EXAM	24 th May 2023
SEMESTER & PROGRAM		Semester-VIII, B.Tech (IT)	TIME	9:30 am to 12:30 pm
TIME ALLOWED		03:00 HRS.	MARKS	100 Marks
COURSE NAME - (CODE)		Human Computer Interaction - (R4IT44021)		
Instructions		1. All questions are compulsory and carry equal marks. 2. Figures to the right indicate full marks 3. Assume suitable data wherever necessary		
Q 1	a	Design the GUI for a web based ticket booking system for Indian Railway	[CO1]	10
	b	Design Concur Task Tree (CTT) – an engineering task model for the user tasks in operating a mobile phone. Assume that we are interested in modeling the tasks related to making a phone call only (not taking photos or listening to music)	[CO1]	10
Q.2	a	Design a GUI for unorganized farm workers with HCI perspective.	[CO2]	10
	b	Identify a short word processing task to be accomplished using Microsoft Word. Also identify alternate methods of achieving same task in Word. Draw the GOMS task structure model for both methods and use KLM to predict the task completion times. Compare the two methods on the basis of task completion time. Now redesign the interaction and/or interface and try to reduce the task completion time by effectively utilizing KLM model.	[CO1] [CO2]	10
Q.3	a	You have to design a menu structure for ordering house-hold items from a mall directly to your home through a mobile phone interface. Categorize the items in whatever way you wish into menus and submenus. Design 3 alternative screens or menus for effectively ordering items from listed categories, making use of Hicks Law. Compare your designs by actually measuring the reaction time for various test cases and also using the Hick's Law expression.	[CO3]	10
	b	Listed below are 15 features of an e-learning interface in random order 1. Course objectives enlisted 2. Multiple formats of quiz available 3. Consistency in look and feel for all screen layouts 4. Different levels of learning tasks exist to adapt 5. Learner has control over pace of learning 6. Navigable Topic Index 7. Previous button to navigate to previous learning page 8. Emphasis of used hyperlinks (topics) 9. Undo - For error recovery 10. Limited length of page to avoid scrolling 11. Chat-rooms facility 12. GUI enjoyable 13. Breadcrumb trail - keeping track of location within application 14. Print facility 15. GUI simple and intuitive Can you group these features into some categories? Can you arrange these features by descending order of their importance? How? Prepare a survey sheet with each feature and its degree of importance in five point Likert scale. You will need to conduct a survey of 15 students, 5 from each age group mentioned below and get their responses to the survey sheet. a) 13 to 17 (school) b) 18 to 22 (college) c) 23 to 27 (professionals)	[CO3]	10

Not Essential

Most Essential



Do you find differences in the preferences of features in these groups?
Try to analyze your data and find the reasons for the difference.
If possible, create prototype GUI for each of these groups.

	Q 4	<p>a You have identified a market opportunity for home media players that would cater for older members of the population. Many older people have difficulty understanding the operating principles of devices such as MP3 players, "internet radios" for streaming audio, and personal video recorders and players. Describe design and evaluation processes that could be used by a start-up company to improve the usability of such devices for this population. You should consider several different stages of the product design cycle, and describe five different user interface design techniques that would be relevant at those different stages.</p> <p>b Draw an 'interaction model' based on Norman's model for the following Interface:</p> <p>Assume all data.</p> <p>An interface for checking number of Leaves (absence with permission) availed off by a student and their type (Medical, Vacation; Conference visits;). Refer to the student leave rules of your institution for necessary constraints and other relevant data. Assume the data whenever required.</p>	[CO3]	10
			[CO3]	10
	Q 5	<p>a Consider the activity of making a telephone call. Record the actions in an HTA diagram or textually. Start off simply, assuming you know the number to dial, but then add more complicated situations : finding the number in an address book, or what to do when the number is engaged.</p> <p>b There are four main translations involved in the interaction framework viz. articulation, performance, presentation and observation.</p> <ul style="list-style-type: none"> i. The compact disk player has a button for power off. However, its remote control does not have a power off button ii. It is difficult in a command line interface to determine the result of copying and moving files in a hierarchical file system. iii. User is unable to figure out which switches from the bank to turn on to lit the front portion of a classroom. iv. The user is unable to know whether the voice recorder is in playing or recording state. <p>Specify in each of the above four cases which of the interaction framework translations are in effective.</p>	[CO3, CO4]	10
			[CO3, CO4]	10



VEERMATA JIJABAI TECHNOLOGICAL INSTITUTE

Matunga, Mumbai-400 019

Autonomous Institute affiliated to University of Mumbai

EXAMINATION	End Semester Examination (ESE) Nov 2022	DATE OF EXAM	06/12/2022
SEMESTER & PROGRAM	Sem-VII, Tech (Information Technology)	TIME	11:00 am to 2:00 pm
TIME ALLOWED	3.0 HRS.	MARKS	100
COURSE NAME - (CODE)	Data Mining and Data Warehousing (R4IT4001T)		

Instructions : All questions are compulsory

Figures to the right indicate full marks

Assume suitable data if required.

- Q.1 a Differentiate between Relational Modeling and Dimension Modeling. 05 CO1
b How can we solve 'missing values' problem? 05 CO1
c Why is a document-term matrix is an example of a data set that has asymmetric discrete features. How is the similarity of documents measured? 05 CO1
d Explain Principal Component Analysis 05 CO1
- Q.2 a Suppose that a data warehouse consists of the three dimensions *time*, *doctor*, and *patient*, and the two measures *count* and *charge*, where *charge* is the fee that a doctor charges a patient for a visit 10 CO1
i) Draw a star schema diagram for the above data warehouse
ii) Starting with the base cuboid [*day*; *doctor*; *patient*], what specific *OLAP operations* should be performed in order to list the total fee collected by each doctor in 2004?
iii) To obtain the same list, write an SQL query assuming the data is stored in a relational database with the schema *fee* (*day*, *month*, *year*, *doctor*, *hospital*, *patient*, *count*, *charge*).
Can we use the same dimension tables for Schedules of Doctors. What is the schema called?
b i) Use single-link agglomerative clustering and Euclidean distance to cluster the following 8 examples:
 $A_1=(2,10)$, $A_2=(2,5)$, $A_3=(8,4)$, $A_4=(5,8)$, $A_5=(7,5)$, $A_6=(6,4)$, $A_7=(1,2)$, $A_8=(4,9)$. Show the dendograms.
ii) Differentiate between agglomerative and divisive clustering
- Q.3 a What are the parameters you need to decide for solving a classification problem using neural networks? Also write the steps for working for neural network. Enumerate the disadvantages of neural networks over decision trees. 10 CO2
b Consider the data set shown in the Table given below 05 CO2
i) Compute the support for itemsets $\{e\}$, $\{b, d\}$, and $\{b, d, e\}$ by treating each customer as a market basket.

Each item should be treated as a binary variable (1 if an item appears in at least one transaction bought by the customer, and 0 otherwise.)

- ii) Compute the confidence for the association rules

$$\{b, d\} \rightarrow \{e\} \text{ and } \{e\} \rightarrow \{b, d\}$$

25% 100%

Customer ID	Transaction ID	Items Bought
1	0001	{a, d, e}
1	0024	{a, b, c, e}
2	0012	{a, b, d, e}
2	0031	{a, c, d, e}
3	0015	{b, c, e}
3	0022	{b, d, e}
4	0029	{c, d}
4	0040	{a, b, c}
5	0033	{a, d, e}
5	0038	{a, b, e}

e - 2

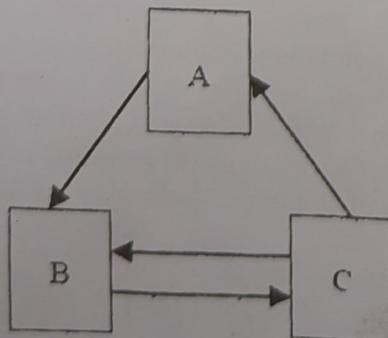
bd - 2

bde - 2

- c. How is data generalization for characterization performed using attribute oriented induction? 05 CO1

- Q.4 a. ABC company which sells different electronic products has decided to Personalize their website to give users personalized experience based on their preferences. 10 CO3
- i) List the data that needs to be collected from the users
 - ii) List the features that can be provided for Personalization.
 - iii) List the data mining techniques and algorithms required for implementing the features
- b. In what format is the data used for Spatial Data mining acquired? 10 CO3
Enumerate the data structures used for accessing spatial data

- Q.5 a. Give three application examples of spatiotemporal data streams. What kind of interesting knowledge can be mined from such data streams, with limited time and resources. What are the challenges in spatio temporal mining 10 CO4
- b. Why is Page Rank important for search engines? 10 CO3
For the following graph of web pages , find the pagerank given the damping factor of 0.7.



$$C(A) = 1$$

$$C(B) = 1$$

$$C(C) = 2$$

$$PR(N) = 1 - d + d \left(\frac{PR(A)}{C(A)} \right)$$

