Paper / Subject Code: 52707 / Human Machine Interaction

Time: 3 hours   **Max.Marks:80**

Note: 1. Question 1 is compulsory.
2. Attempt any 3 from Q2 to Q6.
3. Indicate your answer with various sketches whenever necessary.

Q1. Attempt any **four**.   **[20]**

(a) List pros and cons of any one modern device in design of a tutor for kids.

(b) List techniques in qualitative research.

(c) Differentiate between mental model and implementation model.

(d) Explain effect on data structures when appropriate data controls are not used.

(e) What are the four interfaces that give name WIMP.

Q2. (a) Explain briefly software evolution process.   **[10]**

(b) Design a UI for any multimedia application.   **[10]**

Q3. (a) Provide a systematic design analysis for municipal corporation's mobile app that   **[10]**
includes all the area details in terms of ward etc. Also give the interface guidelines.

(b) Explain seven stages of action and three levels of processing.   **[10]**

Q4   (a) A UI designer wants to design an application for people with difficulty using   **[10]**
keyboard. Comment on selecting proper device based controls.

(b) Differentiate between quantitative and qualitative research in knowing the users.   **[10]**

Q5   (a) State and explain principles of Gestalts theory.   **[10]**

(b) Provide suitable analysis and Interface design for state road transportation system.   **[10]**

Q6   Write Short notes on following.   **[20]**

(a) Statistical Graphics
(b) Guidance and Feedback
(c) Interview Questions
(d) Goal directed Design

_____

55380

Paper / Subject Code: 52701 / Elective- III 1) Machine Learning

(3Hrs)

Max Marks: 80

N.B.: (1) Question No. **1 is compulsory**.

(2) Attempt any **three** of remaining **five** questions.

(3) Assume any suitable **data** if necessary and clearly state it.

1	(a)	Define well posed learning problem. Hence, define robot driving learning problem.	[05]
	(b)	Explain, in brief, Bayesian Belief networks.	[05]
	(c)	Write short note on Temporal Difference Learning.	[05]
	(d)	Explain procedure to construct decision trees.	[05]
2.	(a)	Explain how support vector machine can be used to find optimal hyperplane to classify linearly separable data. Give suitable example.	[10]
	(b)	Explain procedure to design machine learning system.	[10]
3.	(a)	What is linear regression? Find the best fitted line for following example:	[10]

i	x_i	y_i	$\hat{y}_i$
1	63	127	120.1
2	64	121	126.3
3	66	142	138.5
4	69	157	157.0
5	69	162	157.0
6	71	156	169.2
7	71	169	169.2
8	72	165	175.4
9	73	181	181.5
10	75	208	193.8

	(b)	What is decision tree? How you will choose best attribute for decision tree classifier? Give suitable example.	[10]
4	(a)	Explain K-mean clustering algorithm giving suitable example. Also, explain how K-mean clustering differs from hierarchical clustering.	[10]
	(b)	What is kernel? How kernel can be used with SVM to classify non-linearly separable data? Also, list standard kernel functions.	[10]

57922

5. (a) What is Q-learning? Explain algorithm for learning Q. [10]

    (b) Explain following terms with respect to Reinforcement learning: delayed rewards, exploration, and partially observable states. [10]

6   Write short notes on

    (a) Soft margin SVM [05]

    (b) Radial Basis functions [05]

    (c) Independent Component Analysis [05]

    (d) Logistic Regression [05]

****

57922

(3 Hours)   **Total Marks: 80**

N.B.: (1) Question No.1 is compulsory.

     (2) Attempt any three questions from the remaining five questions.

     (3) Make suitable assumptions wherever necessary but justify your assumptions.

1.	(a) What is hacking? Who are the different types of hackers?	**05**
	(b) What is incident and what are the goals of incident response?	**05**
	(c) What volatile data can be obtained from investigation of routers?	**05**
	(d) What are the challenges in evidence handling?	**05**

2.   (a) Classify the different categories of cyber crime with examples of each. Identify the   **10**
type of cyber-crime for each of the following situations:
   i) Hacking into a Web server and defacing legitimate Web pages
   ii) Introducing viruses, worms, and other malicious code into a network or computer
   iii) Unauthorized copying of copyrighted software, music, movies, art, books.

   iv) Internet gambling and trafficking

   (b) Briefly explain the role of the following tools in digital forensics: i) netstat   **10**
   ii) psloggedon iii) tcptrace iv) netcat v) cryptcat

3.   (a) Briefly explain the process of collecting the volatile data in Windows system.   **10**

   (b) Briefly explain each of the following: Qualified forensic duplicate, restored image,   **10**
mirror image.

4.   (a) Explain e-mail forensic investigation methods.   **10**

   (b) Discuss the steps for investigating routers.   **10**

5.   (a) Briefly explain the role of Windows registry in collecting forensic evidence.   **10**

   (b) Explain different bodies of law.   **10**

6.   Write a short note on: **20**

     (1) CFAA

     (2) Storage layer of File system

****

(Time: 03 Hours)   **(Total Marks: 80)**

Note: 1. Question 1 is compulsory
        2. Answer any three out of remaining questions.

Q1 a) Information requirements are recorded for "Hotel occupancy" considering **[10]** dimensions like Hotel, Room and Time. Few Facts recorded are vacant rooms, occupied rooms, number of occupants, etc.
     Answer the following questions for this problem:
       i.   Design the star schema.
       ii.   Can you convert this star schema to a snowflake schema? If yes, justify and draw the snowflake schema.

     b) Explain Data mining as a step in KDD .Illustrate the architecture of typical data **[10]** mining system.

Q2 a) The college wants to record the Marks for the courses completed by students using **[10]** the dimensions: I) Course, II) Student, III) Time & a measure Aggregate marks .
Create a Cube and perform following OLAP operations :
i) Rollup   ii) Drill down   iii) Slice   iv) Dice   v) Pivot.

     b) Apply the Naive Bayes classifier algorithm to classify an unknown sample **[10]**

X (outlook = sunny, temperature = cool, humidity = high, windy = false)
The sample data set is as follows:

Outlook	Temperature	Humidity	Windy	Class
Sunny	Hot	High	False	N
Sunny	Hot	High	True	N
Overcast	Hot	High	False	P
Rain	Mild	High	False	P
Rain	Cool	Normal	False	P
Rain	Cool	Normal	True	N
Overcast	Cool	Normal	True	P
Sunny	Mild	High	False	N
Sunny	Cool	Normal	False	P
Rain	Mild	Normal	False	P
Sunny	Mild	Normal	True	P
Overcast	Mild	High	True	P
Overcast	Hot	Normal	False	P
Rain	Mild	High	True	N

Q3 a) Discuss Data Warehouse design strategies in detail?   **[10]**

     b) Discuss the types of attributes and data visualization for data exploration.   **[10]**

Q4 a) Discuss various OLAP models and their architecture. **[10]**

b) Find clusters using *k*-means clustering algorithm, if we have several objects **[10]** (4 types of medicines) and each object have two attributes or features as shown in table below. The intention is to group these objects into $k = 2$ group of medicine based on the two features (pH and weight index).

Object	Attribute 1 (X) Weight Index	Attribute 2 (Y) pH
Medicine A	1	1
Medicine B	2	1
Medicine C	4	3
Medicine D	5	4

Q5 a) Discuss the process of extraction, transformation and loading with a neat **[10]** and labelled diagram.

b) A database has five transactions. Let minimum support = 40% and minimum **[10]** confidence = 60%
   i) Find all frequent patterns using Apriori Algorithm.
   ii) List strong association rules.

Transaction-Id	Items
A	1, 3, 4, 6
B	2, 3, 5, 7
C	1, 2, 3, 5, 8
D	2, 5, 9, 10
E	1, 4

Q6 Write short note on the following (*Answer any FOUR*) **[20]**

i) Applications of Data Mining (*minimum two in detail*)

ii) Data pre-processing

iii) FP Tree

iv) Updates to dimension tables

v) Meta data with example

* * *

3 Hours   **Total Marks:80**

N.B.   **Question No: 1 is Compulsory**
Attempt any three from the remaining
Assume suitable data wherever necessary

1	**a**	Compare Big data analytics with traditional data mining and warehousing system.	**5**
	b	With respect to data stream querying, give example of a) one time Queries b)continues queries c)predefined Queries d)ad-hoc queries	**5**
	c	How Big Data Analytics can be useful in the development of Digital India	**5**
	d	What are distance measures? Brief any two distance measures.	**5**
2	**a**	Explain K mean algorithms for large data set. Explain its significance as compared to other clustering algorithms.	**10**
	b	How Bloom filter is useful for big data analytics. Explain with one example.	**10**
3	**a**	Explain DGIM algorithm for counting ones in a window.	**10**
	b	Elaborate collaborative filtering system. How is the system different from a content based system.	**10**
4	**a**	Explain Hadoop ecosystem with core components? Explain the Physical Architecture of Hadoop. State its limitations	**10**
	b	What is the MapReduce? Explain the role of combiner with the help of an example.	**10**
5	**a**	Differentiate between a RDBMS and a No-Sql Database	**10**
	b	Explain the Clique Percolation Method (CPM) used in direct discovery of communities in a Social Graph with example.	**10**
6	**a**	List down the steps in HITS Algorithm with one example.	**10**
	b	Explain Park-Chen-Yu algorithm. How memory mapping is done in PCY.	**10**

_____

(3 hrs)   **Marks: 80**

N.B.   (1)   Question one is Compulsory.
       (2)   Attempt any 3 questions out of the rest.
       (3)   Assume suitable data if required.

Q1. a.   What are various system models of distributed system?   **(05)**

     b.   Prove that a k-stage linear pipeline can be at-most k times faster than that of a   **(05)** non-pipelined serial processor.

     c.   Compare parallel and distributed Systems by giving real time examples for each   **(05)**

     d.   The time required to execute a task with single processor is 1200ms and with 8   **(05)** processors it takes 200ms. Find the efficiency of parallel computing.

Q2. a.   Illustrate 4-stage pipeline architecture.   **(10)**

     b.   Differentiate between Message oriented & Stream oriented communications   **(10)**

Q3. a.   Describe any one method of Logical Clock synchronization with the help of an   **(10)** example.

     b.   Illustrate the parallel algorithm for sorting numbers in ascending order with an   **(10)** example and analyze the performance of this algorithm in terms of parallel run time and communication cost.

Q4. a.   What is the need for process migration and explain the role of resource to process   **(10)** and process to resource binding in process migration.

     b.   Illustrate the implementation details of pipelined floating-point adder.   **(10)**

Q5. a.   Discuss and differentiate various client-centric consistency models by providing   **(10)** suitable example application scenarios.

     b.   Discuss Ricart-Agrawala's algorithm and Justify how this algorithm optimized   **(10)** the message overhead in achieving mutual exclusion.

Q6.   Write a short note on any two   **(20)**

     a.   File cashing schemes

     b.   An architecture of Information System

     c.   Load balancing techniques

****

57559

9D52F0F6A8FDAEC3592C926912B34621