8. Discuss the following clustering algorithm using examples : 14

  (a) DBSCAN

  (b) Web mining

  (c) Temporal mining

  (d) BIRCH algorithm

❖

FA—4/i (900)   (4)   UL(7)—DWH & DM

CR+IT

UL(7) — DWH & DM

2020(A)

Time : 3 hours

Full Marks : 70

Candidates are required to give their answers in their own words as far as practicable.

The figures in the margin indicate full marks.

Answer any *five* questions.

1. (a) Demonstrate the applications of data mining for financial analysis. 7

  (b) How classification of data mining system is done ? Explain them with example. 7

2. Examine the steps involved for the design and construction of data warehouse. 14

3. (a) Explain FP tree algorithm with an example. 7

  (b) Explain K-means algorithm with example. 7

FA—4/1   (Turn over)

4. (a) Explain the algorithm for constructing a decision tree from training samples. 7

(b) Explain the data structures and schema that support multidimensional data in data warehouse with suitable illustration. 7

5. Suppose that a data warehouse for big university consist of the following four dimensions : Student, Course, Semester and Instructor, and two measures count and avg_grade. When at the lowest conceptual level (Ex. For a given student, course, semester and instructor combination). the avg_grade measure stores the actual course grade of the student. At higher conceptual levels, avg_grade stores the average grade for the given combination.

(a) Draw a snowflake schema diagram for the data warehouse. 4

(b) Starting with the base cuboid [Student, Course, Semester, Instructor], what specific OLAP operations (eg - Roll-up from Semester to year) should one perform in order to list the average grade of CS for each big-university students. 5

(c) If each dimension has five levels (including all) such as Student<major<status< university<all, how many cuboids will this

cube contains (including the base and Apex cuboids). 5

6. (a) What is a spatial database ? Explain the methods of mining spatial databases ? 7

(b) Describe how OLAP technology helps in discovery driven exploration of data cubes. 7

7. Write an algorithm for constructing a decision tree. Construct a decision tree for the following data set using information gain. Predict the class label for a data point with values<Female, 2, standard, high : 14

Gender	Car ownership	Travel cost	Income level	Transport mode
Male	0	Cheap	Low	Bus
Male	1	Cheap	Medium	Bus
Female	0	Cheap	Low	Bus
Male	1	Cheap	Medium	Bus
Female	1	Expensive	High	Bus
Male	2	Expensive	Medium	Car
Female	2	Expensive	High	Car
Female	1	Cheap	Medium	Train
Male	0	Standard	Medium	Train
Female	1	Standard	Medium	Train

2020(A)

Time : 3 hours

Full Marks : 70

Candidates are required to give their answers in their own words as far as practicable.

The figures in the margin indicate full marks.

Answer any five questions.

1. (a) What is machine learning ? Write down the difference between Machine Learning and Data Mining.   7

   (b) Explain Learning System. Design the learning system for Checkers problem.   7

2. (a) What is lazy learner in machine learning, explain in brief ?   7

   (b) Consider following data set suppose you are given the following set of data with three input   7

(Turn over)

6. (a) Write down the difference between Artificial Neural Network and Biological Neural Network.   7

   (b) What are the different learning law in ANN, explain in brief ?   7

7. (a) What is linearly inseparable problem ? Show that Ex-OR and Ex-NOR are linearly inseparable.   7

   (b) Explain Genetic Algorithm. Illustrate with a simple example.   7

8. (a) What is the significance of ensemble learning in machine learning ? Explain with suitable example.   7

   (b) Explain logistic regression in machine learning. Explain with example.   7

———— ❖ ————

FA—6/1 (900)   (6)   UL(7)—Mach. Learn.

FA—6/1

Boolean variable a, b, c and a single variable K.

a	b	c	K
1	0	1	1
1	1	1	1
0	1	1	0
1	1	0	0
1	0	1	0
0	1	0	1
0	0	0	1
0	0	1	0

According to naive base classifier find the following:

(i) $P(k=1 \mid a=1 \wedge b\,1 \wedge c=0)$

(ii) $P(k=0 \mid a=1 \wedge b\,1)$

3. A smell of sulphur (S) can be caused either by rotten eggs (E) or as a sign of the doom brought by the Mayan Apocalypse (M). The Mayan Apocalypse also causes the oceans to boil (B).

6/1

FA – 6/1   (2)   Contd.
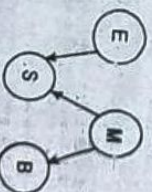

The Bayesian network and corresponding conditional probability tables for this situation are shown below. For each part, you should give either a numerical answer (e.g. 0.81) or an arithmetic expression in terms of numbers form the tables below (e.g. 0.9, 0.9).   14

P(E)	
+e	0.4
-e	0.6

P(M)	
+m	0.1
-m	0.9

P(S / E, M)			
+e	+m	+s	1.0
+e	+m	-s	0.0
+e	-m	+s	0.8
+e	-m	-s	0.2
-e	+m	+s	0.3
-e	+m	-s	0.7
-e	-m	+s	0.1
-e	-m	-s	0.9

P(B I M)		
+m	+b	1.0
+m	-b	0.0
-m	+b	0.1
-m	-b	0.9

The find :

(a) What is the probability that the oceans boil ?

FA– 6/1   (3)   (Turn over)

(b) What is the probability that the Mayan Apocalypse is occurring, given that there is a smell of sulphur, the oceans are boiling, and there are rotten eggs ?

(c) What is the probability that the Mayan Apocalypse is occurring, given that the occeans are boiling ?

4. Imagine that you have given following set of training examples. Each feature can take up to three nominal values a, b, and c.   14

F_1	F_2	F_3	Class
a	c	a	+
c	a	c	+
a	a	c	−
b	c	a	−
c	c	b	−

How would the Naive system classify the following test example :

$F_1 = a$, $F_2 = c$, $F_3 = b$.

FA – 6/1   (4)   Contd.

5. (a) Consider the following data set and perform KNN classification and predict the class for X(P1 = 2, P3 = 7), For K = 3.   7

P_1	P_2	Class
7	7	False
7	4	False
3	4	True
1	4	True

(b) Use the k-means algorithm and Euclidean distance to cluster the following 8 examples into 3 cluster :   7

Point- $A_1 = (2, 10)$, $A_2 = (2, 5)$, $A_3 = (8, 4)$ $A_4 = (5, 8)$, $A_5 = (7, 5)$, $A_6 = (6, 4)$, $A_7 = (1, 2)$, $A_8 = (4, 9)$.

Suppose that the initial seeds (centers of each cluster) are A_1, A_4 and A_7. Run the k-means algorithm for 1 each only. At the end of this epoch show : The new cluster (i. e. the examples belonging to each cluster).

FA – 6/1   (5)   (Turn over)

UL(7)—Soft. Engg.

CS+IT

2020(A)

Time : 3 hours

Full Marks : 70

Candidates are required to give their answers in their own words as far as practicable.

The questions are of equal value.

Answer any five questions.

1. (a) What do you understand by black box and white box testing ? Show the difference between black box and white box testing.

   (b) Describe the Non-functional requirement in software development.

2. (a) What do you understand by UML ? Describe it's uses in design phase.

   (b) What are the differences between verification and validation in software development ?

(Turn over)

FA – 8/1

---

8. Write short notes on any two of the following :

   (a) Regression Testing

   (b) Sequence Diagram

   (c) Iterative Waterfall Model

———— ❖ ————

FA – 8/1 (900)   (4)   UL(7) — Soft. Engg.

3. (a) Describe the difference between Alpha testing and Beta testing.

(b) A software company has won the contract to build the Software for an embedded system. Assume that 200000 lines of code is required for the software. Compute the Effort, Time and Persons required to develop the software when basic COCOMO estimation model is used.

4. (a) What do you understand by Agile methodology in the software project development ? What are the different types of agile models ? How these models handle the challenges faced in traditional software development model ?

(b) Define the terms MTTR, MTBF and MTTF in software maintenance. Write down the formula to calculate these values. How it is helpful in software maintenance ?

5. Explain the different cohesion and coupling method used in software development. Also discuss the types of cohesion and coupling are good for software development and why ?

6. (a) Discuss functional and non-functional requirements in software engineering in detail.

(b) What do you understand by feasibility study ? What are the different types of feasibilities analyses done before going through any project ? Discuss them with example.

7. (a) Whether any development life cycle models are suitable for any type of project ? Justify your answer.

(b) Describe the different phases involve in prototype model and evolutionary model.

UL(7) — Crypto.

2020(A)

Time : 3 hours

Full Marks : 70

Candidates are required to give their answers in
their own words as far as practicable.

The figures in the margin indicate full marks.

Answer any five questions.

1. (a) Explain the block cipher and stream cipher.

    8

    (b) Explain double DES. What kind of
    attack on double DES makes it use less ?

    6

2. (a) Find all multiplicative inverses in z_{10}.   7

    (b) Solve the following equation :   7

    $10x \equiv 2 \pmod{15}$

FA — 5/1

(Turn over)

7. Write short notes on any two of the following :

7+7 = 14

(a) Diffie Hellman

(b) Key Management

(c) SHA – 512

(d) Digital Signature

———— ❖ ————

3. (a) Explain the type of Intrusion Detection Systems. 7

(b) Compare conventional encryption and public-key encryption. 7

4. (a) State and explain some advantages and some disadvantages of static and dynamic s-boxes. 7

(b) Compare DES and AES. Which one is bit-oriented ? 7

5. (a) Find the result of the following $(21)^{24}$ mod 8. 7

(b) State and explain the advantages and disadvantages of symmetric-key and asymmetric-key crypto system. 7

6. (a) List some features of the whirlpool crypto-graphic function. What kind of compression function is used in whirlpool ? 7

(b) Explain security goals. 7

2020(A)

Time : 3 hours

Full Marks : 70

Candidates are required to give their answers in their own words as far as practicable.

The figures in the margin indicate full marks.

Answer any *five* questions.

1. (a) What do you understand by AI ? What are AI problems ? What is the AI technique to solve an AI problem ?   7

   (b) Suppose you are given with two jugs, a 5 liters one and a 3 liters one. Neither has any marker on it. There is a pump that can be used to fill the jugs with water. How can you get 2 liters of water in the 5-liter jug ? Write the production rules to get the state space representation of the given problem.   7

FA–2/1   (Turn over)

2. (a) Consider the following set of axioms : 10

   (i) Sham likes easy courses.

   (ii) All courses in Arts department are easy.

   (iii) All courses in Science department are not easy.

   (iv) Physics is a Science course.

   (v) Sketching is an Arts course.

   Translate these sentences into predicate form and find using resolution principle "Which course does Sham like ?"

   (b) Define functions in LISP to do the following : 4

   (i) A function that finds maximum among three numbers.

   (ii) A recursive function that implements member function.

3. (a) Explain hill climbing method of solving a problem. What are the problem in hill climbing search methods due to which they may fail to find the solutions ? 7

(b) Solve the given traveling salesman problem using **branch and bound** method. Assume the source city is A and the travelling salesman is to go through a found trip from city A to city A, visiting all other three cities B, C and D exactly once, covering minimum possible distance path : 7

	A	B	C	D
A	-	200	125	75
B	200	-	50	75
C	125	50	-	50
D	75	75	50	-

4. (a) Explain A* search algorithm. Differentiate between best first search and A* algorithm. 7

(b) Find the solution to the 8 puzzle problem using A* search algorithm. State the heuristic function used : 7

Initial State

2		
1	8	4
7	6	5

Goal State

1	2	3
8		4
7	6	5

5. (a) What do you understand by knowledge representation and mappings ? What are the various approaches to knowledge representation ?   7

   (b) How can we represent knowledge using frames ? What are the reasoning actions that can be performed using frames ?   7

6. (a) Solve the cryptarithmetic puzzle :   7

   (i) LOGIC + LOGIC = PROLOG

   (ii) BASE + BALL = GAME

   (b) Draw a semantic network to represent : Every teacher likes intelligent students.   7

7. (a) What are the advantages and disadvantages of Expert System ? Discuss different existing Expert Systems.   10

   (b) Write PROLOG program for :   4

   (i) Revesring a list.

   (ii) Implementing quick sort

8. Differentiate all the following (Mention any two differences) :   2×7 = 14

   (a) Breadth First Search Vs Depth First Search.

   (b) Simple Hill Climbing Vs Steepest ascent Hill Climbing.

   (c) Simulated Annealing Vs Hill Climbing.

   (d) A* searching teachnique Vs AO* searching technique.

   (e) Procedural knowledge Vs Declarative Knowledge.

   (f) Conventional Computer System Vs Expert System.

   (g) Supervised learning Vs Unsupervised learning.

   —————— ❖ ——————

FA – 2/1   (4)   Contd.

FA – 2/1 (900)   . (5)   UL(7)—Arti. Int.

12