2021 (A) -NEW

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 right-hand margin indicate full marks.

Answer any **Five** questions.

1. (a) Consider the following problem. [2×4 = 8]
   - John likes all kinds of food.
   - Apples are food.
   - Chicken is food.
   - Anything anyone eats and isn't killed by is food.
   - Bill ate peanuts and still alive.
   - Sue eats everything Bill eats.

   i. Convert the formulas into clause form.
   ii. Prove that "John likes peanuts" using resolution.
   iii. Translate these sentences into formulas in predicate logic.
   iv. Prove that John likes peanuts using backward chaining.

   (b) Explain Multi-layer perceptron (MLP) with back propagation with schematic block diagram. [6]

2. (a) Why is the use of Nearest Neighbor classifiers problematic in high dimensional input spaces. [7]
   (b) Describe Alpha-Beta pruning and give the other modifications to min max procedure to improve its performance. [7]

3. (a) Explain the following terms: - [2×4 = 8]
   i. Ideal Rational Agent
   ii. Q-learning.
   iii. Uniformed search Strategies
   iv. Greedy best-first search

   (b) Explain how Bayesian statistics provides reasoning under various kinds of uncertainty. [6]

4. (a) Explain about Tic-Tac-Toe game problem by assuming one player is X the other one can be either human or a computer by taking 3X3 grid space. [9]
   (b) Describe in detail about Dempster-Shafer theory. [5]

5. (a) Prove that the A* heuristic search algorithm is optimal when applied in conjunction with a monotonic heuristic. State the conditions under which the algorithm is also complete, and explain why this is the case. [8]
   (b) Explain the concept of planning with state space search. How is it different from partial order planning? [6]

6. (a) Explain the constraint satisfaction procedure to solve the crypt arithmetic problem. [6]
   (b) Why does search in game-playing programs always proceed forward from the current position rather than backward from the goal? [8]

7. (a) Differentiate the following: [4×2 = 8]
   i. Statistical Learning Methods and Reinforcement Learning.
   ii. Forward Chaining and Backward Chaining.

   (b) Illustrate swap mutation, scramble mutation and inversion mutation with the help of examples. [6]

8. Write short notes on any **two** of the following: [7×2 = 14]
   (a) Probabilistic Reasoning.
   (b) Recursive best-first search (RBFS)
   (c) Decision Tree Learning
   (d) Fuzzy Logic System.

2021 (A) -NEW

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 right-hand margin indicate full marks.

Answer any **Five** questions. [2×5 = 10]

1. (a) Explain the following term with proper example:
   i. Aggregation
   ii. Sampling
   iii. Dimensionality reduction
   iv. Feature subset selection
   v. Discretization and binarization

   (b) Explain why ETL must deal with dirty data when extracting information from the source systems. [4]

2. (a) Suppose a data warehouse of a hospital consists of four dimensions, namely time, doctor, Location and patient with the concept hierarchy as follows: Time: day, month, quarter, year; Doctor: doctor, specialization (e.g., pediatrician, etc.); Patient: patient, category (e.g. outdoor, indoor), address; Location: location, address. There are two measures, namely count and charge, where charge is doctor's consultation fees. Draw a star and snowflake schema for the above data warehouse. [8]

   (b) Explain the motivating challenges in development of data mining. [3]

   (c) Why is the entity-relationship modelling technique not suitable for the data warehouse. [3]

3. (a) How Data Mining is useful in Educational System? [4]

   (b) Discuss the importance of similarity metric clustering? Why is it difficult to handle categorical data for clustering? [6]

   (c) Illustrate the various OLAP operations in Data Warehouse? [4]

4. (a) "Data preprocessing is necessary before data mining process". Justify your answer [4]

   (b) How to cluster the data sets using k-medoid clustering algorithm? [5]

   (c) Explain the architecture of data warehouse with proper example. [5]

5. (a) Explain statistical based, distance based, deviation-based outlier detection? [4]

   (b) How do you extract structures from unstructured textdata? What features are extracted in this process Explain. [4]

   (c)Make a comparison of Apriori and ECLAT algorithms for frequent item set mining in transactions databases. Apply these algorithms to the following data:[6]

| TID | LIST OF ITEMS |
|---|---|
| 1. | Bread, Milk, Sugar, TeaPowder, Cheese, Tomato |
| 2. | Onion, Tomato, Chillies, Sugar, Milk |
| 3. | Milk, Cake, Biscuits, Cheese, Onion |
| 4. | Chillies, Potato, Milk, Cake, Sugar, Bread |
| 5. | Bread, Jam, Mik, Butter, Chilles |
| 6. | Butter, Cheese, Paneer, Curd, Milk, Biscuits |
| 7. | Onion, Paneer, Chilies, Garlic, Milk |
| 8. | Bread, Jam, Cake, Biscuits, Tomato |

6. (a) Build a Decision Tree for classification using the training data in the table given below. Divide the Height attribute into 3 ranges as follows: Less than 1.6, 1.6-1.8, Greater than 1.8 **[9]**

| Gender | Height | Class |
|---|---|---|
| F | 1.58 | Tall |
| M | 1.58 | Medium |
| M | 1.7 | Medium |
| F | 1.65 | Tall |
| F | 1.85 | Tall |
| F | 1.4 | Short |
| M | 1.4 | Short |
| M | 1.7 | Medium |
| F | 1.75 | Tall |
| M | 1.82 | Tall |
| F | 1.6 | Tall |

(b) Discuss Bagging and boosting methods for increasing the accuracy. **[5]**

7. (a) Differentiate the following: **[4×2 = 8]**

i. Direct Query Answering and Intelligent Query Answering

ii. Web Content Mining and Text Mining.

(b) How can the data cube be efficiently constructed for discovery-driven Exploration? Explain various operations of a Data Cube. **[6]**

8. Write short notes on any **two** of the following: **[7×2 = 14]**

(a) Hierarchical Clustering

(b) Decision tree construction with pre sorting

(c) Data warehouse schema

(d) Issues and challenges of Data warehouse

UL(7) - Soft. Engg.

2021 (A) NEW

Time: 3 hours

Full Marks: 70

Candidates are required to give their answer 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 mean by software process? Explain reusability in software engineering. 7
   (b) Define SDLC in software engineering? Explain the spiral life cycle model in brief. 7

2. (a) Define feasibility study? Explain functional and non-functional requirement in terms of software engineering. 7
   (b) What is SRS? Which qualities are requiring for SRS? 7

3. (a) Describe modularization? Explain software architectural design with block diagram. 7
   (b) What is the need of coupling and cohesion. Explain with its type. 7

4. (a) What is Unit testing? Discuss about testing strategy for conventional software. 7
   (b) Explain the difference between software error, fault and failure with suitable example? 7

5. (a) What is Risk management? Explain principle of risk management in brief. 7
   (b) Define software quality assurance in software engineering? Explain major software quality assurance activities. 7

6. (a) What is Risk assessment? Explain RMMM plan in details. 7
   (b) Explain the object-oriented design notation in brief? 7

7. (a) What are the different reliability metrics, explain in brief? 7
   (b) Why we need UML diagram? Describe its usage in design phase. 7

8. Write short notes- (any two) 2x7= 14
   i. Formal technical review
   ii. Walkthrough
   iii. ISO 9000 quality standards

UL ( 7 ) - Machine Learning

2021 ( A ) NEW

Full Marks: 70

Time - 3 Hours

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

The figures in the right hand margin indicate full marks.

**Answer Any 5 ( FIVE ) Questions :**

1. ( a ) Write down the major differences between K-means clustering and hierarchical clustering. [ 7 ]

b ) Define Hidden Markov Model. What is meant by evaluation problem and how is this solved? [ 7 ]

2. (a) Explain feature selection and feature extraction method for dimensionality reduction. [ 7 ]

(b) What is the significance of optimal separating hyperplane in SVM? [ 7 ]

3. ( a ) What is a Perceptron? Explain the working of a perceptron with a neat diagram. [ 7 ]

(b ) Describe the random forest algorithm to improve classifier accuracy. [ 7 ]

4. (a) With suitable equations, explain any two types of activation functions used in neural networks. [ 7 ]

(b) Use K Means clustering to cluster the following data into two groups. Assume cluster centroid are m1=2 and m2=4. The distance function used is Euclidean distance. { 2, 4, 10, 12, 3, 20, 30, 11, 25 } [ 7 ]

5. (a) Differentiate between supervised and unsupervised training. Explain with suitable examples. [ 7 ]

(b ) Describe the concept on density based clustering and write the steps involved in DBSCAN algorithm. [ 7 ]

6. ( a ) Briefly describe the concept on model selection and generalisation. [ 7 ]

( b ) State the mathematical formulation of the SVM problem. Give an outline of the method for solving the problem. [ 7 ]

7. ( a ) Describe the significance of soft margin hyperplane and explain how they are computed. [ 7 ]

( b ) What is the role of radial basis function i separating nonlinear patterns. [7]

8. Write short Notes on **Any Two** of the following : [ 7 x 2 = 14 ]

( a ) Reinforcement Learning

( b ) Gradient Decent and Delta Rule

( c ) Genetic Programming

( d ) Benefits of pruning in decision tree induction

(1) a b
(2) a b
(6) a
(7) a b
(9) a a





2021 (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) Compare Active attacks and Passive Attacks. [7]
   b) What is meant by cryptography? Discuss about encryption. [7]

2. a) Draw the general structure of DES. Explain the encryption and decryption process. [8]
   b) Discuss various transformation functions of AES. [6]

3. a) Identify the possible threats for RSA algorithm and list their counter measures. [8]
   b) Briefly explain Deffie Hellman key exchange with an example. [6]

4. a) Explain in detail about Network-Based Intrusion Detection Systems. [7]
   b) Find all multiplicative inverses in $Z_{11}$. [7]

5. Users A and B use the Diffie - Hellman key exchange technique, a common prime q=11 and a primitive root alpha=7.
   a) What is the shared secret key? Also write the algorithm. [7]
   b) How man in middle attack can be performed in Diffie Hellman algorithm. [7]

6. a) Describe IP security Architecture. [7]
   b) What is transport mode and tunnel mode authentication in IP? Describe how ESP is applied both these modes. [7]

7. a) Discuss the various types of viruses. [7]
   b) Explain the Security of hash functions. [7]

8. Write short notes on any two of the following: [7 + 7]
   a) Key Management
   b) SHA-512.
   c) Asymmetric-key crypto system
   d) Types of Firewalls.

21015550