

**Seventh Semester B.E FASTTRACK Examination, AUGUST SEPTEMBER 2021**  
**DATA SCIENCE**

Time: 3 hrs

Max. Marks : 100

Instructions:- Answer any Five full Questions.

- 1a. Define Data science. Write the differences between data science and Big data. [2] [1] [2] [10]
- 1b. Describe the different steps of data science life cycle. [2] [1] [2] [10]
- 2a. Define Datafication. Explain the different application of data science. [2] [1] [2] [10]
- 2b. Write a note on history of data science. Illustrate how data science is used in different fields. [2] [1] [2] [10]
- 3a. Describe the importance of exploratory data analysis in data science. [2] [2] [1, 3] [6]
- 3b. Write the difference between sample and populations with an example. [2] [2] [2] [7]
- 3c. Define model. Explain stastical modeling with an example. [2] [1, 2] [2] [7]
- 4a. Illustrate the data science process with a neat diagram. [2] [2] [1] [10]
- 4b. What is hypothesis testing? Explain some hypothesis test methods. [2] [3] [4] [10]
- 5a. Explain linear regression for predicting the house price with an example code. [3] [3] [4] [10]
- 5b. Explain the classification of machine learning algorithms. [2] [2, 4] [2] [10]
- 6a. Explain KNN algorithm. List the different modeling assumptions to be made while using KNN algorithm. [2] [2] [2] [10]
- 6b. Describe K-means with an example diagram and suitable snippet. [3] [2, 3] [4] [10]
- 7a. Demonstrate how naïve bayes classifier is used to classify the messages as spam or ham. [3] [3] [4] [10]
- 7b. Explain the comparison between Naïve Bayes and KNN algorithm. [2] [2, 4] [4] [10]
- 8a. Describe logistic regression for classification with an example. [2] [4] [4] [10]
- 8b. Write a note on  
 i) newton's Method [2] [3] [4] [10]  
 ii) Stochastic Gradient Descent
- 9a. Demonstrate how mapreduce can be used in word frequency count problems. [3] [4] [4] [10]
- 9b. List and explain the drawbacks of hadoop. [2] [2] [3, 4] [10]
- 10a. Explain how modeling is done in data engineering.s [2] [2, 3] [2] [10]
- 10b. Write a note on - i) Hadoop ii) Cloudera [2] [3] [4] [10]

**Seventh Semester B.E. Makeup Examination, January 2020**  
**DATA SCIENCE**

Time: 3 Hours

Max. Marks: 100

- Instructions:* 1. Non-programmable calculators are allowed  
 2. Answer any one full question from each unit.

**UNIT - I**

- 1 a. Explain the different sources of Big Data deluge. (2) (1) (1) (06)
- b. Define Data Science. Explain its goal. Give any two real life examples for applications of data science. (2) (1) (1) (06)
- c. List the different steps of data science life cycle. Briefly explain each one of them. (2) (1) (1) (08)

**OR**

- a. Briefly explain the classification of big data with examples. (2) (1) (1) (07)
- b. What is Datafication? Explain it with any two examples. (2) (1) (1) (06)
- c. Define Data Scientist? Illustrate the five main sets of skill sets of Data Scientist. (2) (1) (1) (07)

**UNIT - II**

- 3 a. Explain the importance of exploratory data analysis in data science. (2) (1) (5) (06)
- b. Define the terms Populations and Samples. Explain the two with examples. (2) (1) (5) (06)
- c. What is the meaning of model? Explain the statistical model. How do you build a model? (2) (1) (5) (08)

**OR**

- 4 a. Explain the different methods of Sampling. (2) (1) (5) (06)
- b. Define the term Inference. Illustrate the different types of Inferential Statistics. (2) (1) (5) (06)
- c. Explain the steps in Data Science process with the help of a block diagram. (2) (1) (5) (08)

**UNIT - III**

- 5 a. Demonstrate the use of linear regression for predicting house price with suitable snippets. (3) (3) (4) (10)
- b. Illustrate K Nearest Neighbor (KNN) algorithm with an example. List the modeling assumptions considered while using KNN (3) (3) (4) (10)

**OR**

- 6 a. Illustrate the importance of least squared methods in fitting the model (2) (3) (2) (08)
- b. Write a note on K means clustering algorithm. (2) (3) (1) (08)
- c. What is Machine learning? Explain the classification of machine learning algorithms (2) (3) (1) (04)

**UNIT - IV**

7 a. Why linear regression cannot be used for spam filtering?

(1) (2) (2)

16IS6

b. Analyze the following raw dataset with  $y$  as response and  $x_1$  &  $x_2$  as predictor variables to estimate the three coefficients,  $\beta_0$ ,  $\beta_1$  and  $\beta_2$ , using logistic regression.

(06)

$x_1$	$x_2$	$y$
2.7810836	2.550537003	0
1.465489372	2.362125076	0
3.396561688	4.400293529	0
1.38807019	1.850220317	0
3.06407232	3.005305973	0
7.627531214	2.759262235	1
5.332441248	2.088626775	1
6.922596716	1.77106367	1
8.675418651	-0.2420686549	1
7.673756466	3.508563011	1

(4) (2) (4) (08)

c. Compare and contrast Logistic and Linear regression algorithms

(2) (3) (1) (06)

**OR**

8 a. Explain the Bayes Law with meaning of terms used.

(2) (1) (1) (06)

b. Explain the spam classification using Naïve Bayes technique.

(2) (3) (4) (08)

c. Apply Bayes law to find the probability that the patient is actually infected? Given that 1% of the population is infected. A highly sensitive and specific medical test shows 99% of sick patients test positive and 99% of healthy patients test negative.

(3) (1) (2) (06)  
L CO PO M

**UNIT - V**

9 a. Explain the MapReduce technique with a suitable example

(2) (1) (5) (10)

b. Explain the term Hadoop and how to get started with Hadoop.

(2) (1) (5) (10)

**OR**

10 a. Explain Word frequency problem in detail.

(2) (1) (2) (10)

b. Explain what MapReduce can do? And what it can't.

(2) (1) (5) (10)

**Seventh Semester B.E. Makeup Examination, January 2019****DATA SCIENCE**

Time: 3 Hours

Max. Marks: 100

**Instructions:**

1. Unit I and unit II are compulsory
2. Answer anyone full question from remaining units

**UNIT - I**

L	CO	PO	M
(2)	(1)	(2)	(08)

1. a. What is data science? Illustrate the relationship between big data and data science. (2) (1) (2) (08)
- b. Define datafication. Bring out the differences between a data analyst and data scientist. (2) (1) (2) (08)
- c. List and explain the different challenges that arise in processing big data. (2) (1) (4) (04)

**UNIT - II**

2. a. What is a model? Explain how modeling is done using exploratory data analysis. (2) (2) (2) (10)
- b. Illustrate the process of data science with a diagram. (2) (2) (1) (10)

**UNIT - III**

3. a. What is machine learning? Explain the machine learning algorithms. (2) (1) (1) (10)
- b. Compare between Supervised and Unsupervised machine learning. (3) (2) (4) (10)

**OR**

4. a. Briefly explain the K Nearest Neighbour algorithm. List out the modeling assumptions to be made while using KNN algorithm. (2) (2) (4) (10)
- b. Explain the logistic regression model. (2) (2) (2) (10)

**UNIT - IV**

5. a. Explain classifiers, run-time, interoperability with respect to logistic regression. (2) (2) (1) (10)
- b. Write a brief note on:
  - (i) Newton's method for maximum likelihood
  - (ii) Stochastic Gradient Descent

**OR**

6. a. Demonstrate how classification of spam messages is performed using Naïve Bayes algorithm. (3) (3) (2) (10)
- b. Explain why KNN and Linear regression algorithms cannot be used for filtering spam. (2) (2) (2) (10)

**UNIT - V**

7. a. Explain MapReduce and its framework with an example. (2) (2) (1) (10)
- b. Write a note on:
  - (i) Hadoop
  - (ii) Role of Data scientist in Data Analytics field

3-19

**OR**

- 8 a. Explain how data abundance and data scarcity is handled in data engineering (2) (2) (1) (06)
- b. Explain the word frequency problem. How is it tackled using MapReduce? (2) (2) (5) (06)
- c. Illustrate how modeling is done in data engineering. (2) (2) (5) (08)

DATA  
MINING  
AND  
DATA  
ENGINEERING  
TECHNOLOGY

**Seventh Semester B.E. Semester End Examination, Dec/Jan 2018-19****DATA SCIENCE**

Time: 3 Hours

Max. Marks: 100

**Instructions:** 1. Unit I and unit II are compulsory.  
2. Answer any one full question from remaining units.

**UNIT - I**

- 1 a. Define data science. List out the similarities and differences between Big data and Data science. (2) (1) (1) (10)
- b. Write a brief note on history of data science. Illustrate how data science is applied in different fields (3) (1) (1) (10)

**UNIT - II**

- 2 a. Write a note on  
 (i) Population  
 (ii) Sample  
 (iii) Parameter  
 (iv) Estimation of parameters (2) (1) (2) (10)
- b. Explain data science process with a neat diagram (2) (1) (1) (10)

**UNIT III**

- 3 a. Demonstrate how least squares method is used to best fit the linear regression algorithm (3) (2) (2) (10)
- b. Illustrate K means algorithm with an example (2) (2) (2) (10)

**OR**

- 4 a. Demonstrate KNN algorithm with an example and discuss the underlying assumptions made while using KNN (3) (3) (2) (10)
- b. Illustrate how predicting of house prices is done using linear regression model (2) (3) (4) (10)

**UNIT - IV**

- 5 a. Demonstrate how Naïve Bayes algorithm is used in classification of spam filters (3) (3) (2) (10)
- b. Discuss the merits and demerits of Linear Regression for Spam filtering. (2) (3) (2) (10)

**OR**

- 6 a. Compare and contrast between Naïve Bayes and KNN algorithm in classification of messages (4) (3) (1) (10)
- b. Demonstrate Linear regression model with an example (3) (3) (2) (10)

**UNIT - V**

- 7 a. Explain map reduce considering word frequency problem as an example (2) (2) (1) (08)
- b. Illustrate handling of data abundance and data scarcity in data engineering (2) (2) (1) (06)
- c. Explain how modeling is done in data engineering process (2) (2) (5) (06)

**OR**

- 8 a. Explain the Hadoop and MapReduce. (2) (2) (1) (10)
- b. Write a short note on: (2) (2) (5) (10)
- (i) Role of Data Scientist in Data Analytics field
  - (ii) Examples of MapReduce

**Sixth Semester B.E Makeup Examination, Sept.\_Oct.\_2020****DISTRIBUTED COMPUTING**

Time: 3 hrs

Max.Marks :100

Instructions :1. Answer any Five full Questions selecting at least One Full Question from Each Unit. 2. Each Question carry Equal Marks. 3. Missing Data may be suitably assumed. 4. Draw Figures wherever necessary.

**MODULE 1**

L CO PO M

1a. Define Distributed System &amp; discuss its characteristics.

[2] [1] [1] [5]

1b. List the various challenges in Distributed Systems. Explain any four in brief.

[2] [1] [1] [10]

1c. With appropriate example explain Distributed system with a neat diagram.

[2] [1] [1] [5]

**OR**

2a. Illustrate with an example &amp; explain the following; i) Web cache &amp; proxy server ii) Thin clients

[2] [1] [1] [10]

2b. Summarize the following design requirements for Distributed Architectures

i) Performance Issues ii) Dependability issues

[2] [1] [1] [10]

**MODULE 2**

3a. Explain Characteristics and issues related to TCP stream communication

[2] [2] [1] [10]

3b. Define Marshalling. Construct a marshaled form that represents an Organization with instance variable values :{ 'KLSGIT', 'BELGAUM', 1979} by using CORBA-CDR &amp; Java Serialization.

[3] [2] [2] [10]

**OR**

4a. Discuss RMI invocation semantics and also tabulate failure handling mechanism for each.

[2] [2] [1] [10]

4b. Define RPC and how it can be implemented.

[1] [2] [1] [10]

**MODULE 3**

5a. With a neat diagram explain the components of file service architecture in brief w. r. t. following:i) Flat File Service ii) Directory Service iii) Client Module

[2] [3] [1] [10]

5b. List and explain out the transparencies in a distributed file system.

[2] [3] [1] [10]

**OR**

6a. Write the steps of RSA Algorithm. Illustrate with an example given P=3,Q=11 &amp; message(m)=3.

[4] [3] [3] [10]

6b. Analyze the following uses of Cryptography with suitable scenarios.

i) Secrecy and integrity ii) Digital Signatures

[4] [3] [3] [10]

**MODULE 4**

7a. Explain Berkeley algorithm for internal synchronization.

[2] [4] [1] [10]

~~QUESTION~~ ~~CHARTER~~

7b. With the neat diagram. Explain the concept of synchronization subnet in an NTP implementation. [2] [4] [2] [10]

OR

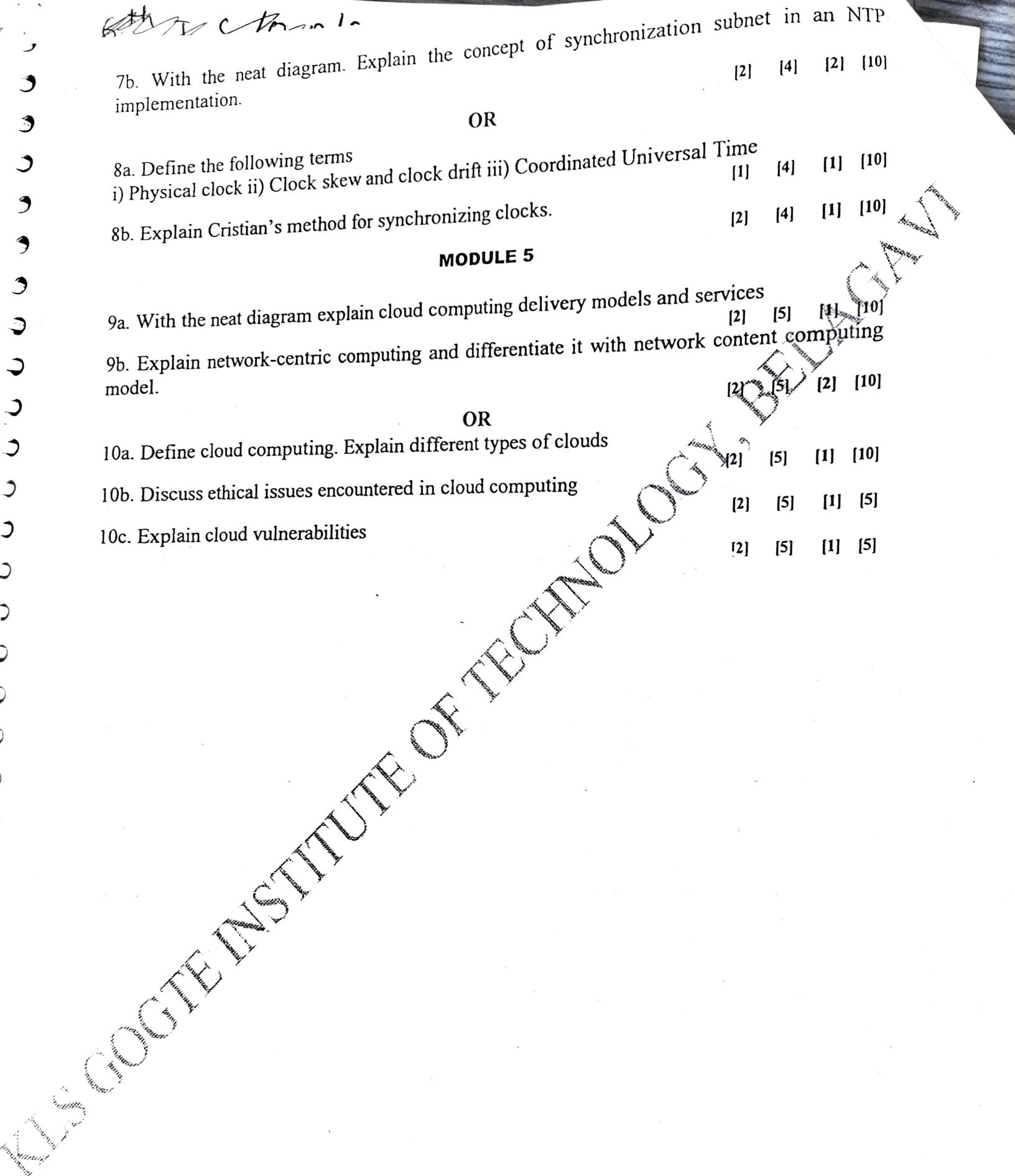
- 8a. Define the following terms  
i) Physical clock ii) Clock skew and clock drift iii) Coordinated Universal Time [1] [4] [1] [10]
- 8b. Explain Cristian's method for synchronizing clocks. [2] [4] [1] [10]

**MODULE 5**

- 9a. With the neat diagram explain cloud computing delivery models and services [2] [5] [1] [10]
- 9b. Explain network-centric computing and differentiate it with network content computing model. [2] [5] [2] [10]

OR

- 10a. Define cloud computing. Explain different types of clouds [2] [5] [1] [10]
- 10b. Discuss ethical issues encountered in cloud computing [2] [5] [1] [5]
- 10c. Explain cloud vulnerabilities [2] [5] [1] [5]



**Sixth Semester B.E. Makeup Examination, June 2018**  
**DISTRIBUTED COMPUTING**

Time: 3 Hours

Max. Marks: 100

*Instructions:* 1. Unit I and III are compulsory.  
2. Draw diagrams neatly wherever applicable.

UNIT - I

- Ques 1** a. Explain client-server and peer to peer architecture with a neat diagram for each case.  
(Level [2], CO [1], PO [1])  
b. With suitable diagrams explain intranet, mobile and ubiquitous computing.  
(Level [2], CO [1], PO [1])

UNIT - II

- 2 a. Explain the characteristics of Inter process communication. ( Level [2], CO [2], PO [3] ) 06 M

b. Explain the UDP datagram communication, along with its issues and failure model. ( Level [2], CO [2], PO [1] ) 08 M

c. Illustrate the use of events with an example of dealing room system. ( Level [2], CO [2], PO [2] ) 06 M

OR

- 3 a. Explain the distributed object model. ( Level [2], CO [2], PO [1,2] ) 06 M

b. Explain the TCP stream communication, along with the related issues and failure model. ( Level [2], CO [2], PO [1] ) 08 M

c. Summarize the role of client and server stub procedures in RPC. ( Level [2], CO [2], PO [1, 2] ) 06 M

UNIT - III

- 4 a. Solve using RSA algorithm to obtain the Cipher texts using encryption and get back the plain text using decryption for the Plain text numbers  $P_1=18$  and  $P_2=19$ . Assume  $P=5$ ,  $Q=11$ ,  $d=7$ . Compute  $n$ ,  $e$ ,  $C$  (cipher text) and  $P$  (Plain text). ( Level [L3], CO [3], PO [2] )

b. With the help of a neat diagram explain the file service architecture and the flat file service operations involved. ( Level [L2], CO [3], PO [1] )

## **UNIT - IV**

- 5 a. ~~Explain the architecture of Network Time Protocol with suitable diagram.~~ ( Level [2], CO [3], PO [2] ) 10 M

b. What is a logical clock? Explain how to synchronize logical clocks in Lamport's algorithm. ( Level [2], CO [3], PO [2] ) 10 M

OR

- 6 a. Explain the Cristian's method for synchronizing clocks. ( Level [2], CO [3], PO [2,12] ) 10 M

b. What is an Election algorithm? Explain a ring-based Election algorithm. ( Level [2], CO [3], PO [2] ) 10 M

**UNIT -V**

- 7 a. Explain the characteristics of network centric computing and network centric content. ( Level [L2], CO [5], PO [1] ) 10 M
- b. Distinguish between different types of clouds based on their infrastructure. List the reasons for the success of cloud computing. ( Level [L4], CO [5], PO [1] ) 10 M
- OR**
- 8 a. With the help of a neat diagram of the structure of the 3 delivery models, compare the services of the 3 models SaaS, PaaS and IaaS. ( Level [L4], CO [5], PO [1] ) 10 M
- b. Discuss the ethical issues and major challenges in cloud computing. ( Level [L2], CO [5], PO [3] ) 10 M

## Sixth Semester B.E. Semester End Examination, May / June 2018

**DISTRIBUTED COMPUTING**

Max. Marks: 100

Time: 3 Hours

**Instructions:** 1. Unit I and III are compulsory.  
2. Answer any one full question from remaining Units.

**UNIT - I**

- 1 a. Define distributed system and explain its significant characteristics or consequences. (Level [1, 2], CO [1], PO [1])

04 M

- b. Analyze the design requirements for distributed architectures. (Level [4], CO [1], PO [1, 3])

06 M

- c. Explain different types of system architectural models with suitable diagrams. (Level [2], CO [1], PO [1])

10 M

**UNIT - II**

- 2 a. With the help of a neat diagram explain client server communication involving request reply protocol along with the message identifiers. (Level [2], CO [2], PO [1])

10 M

- b. Explain TCP stream communication with respect to its message size, lost messages, Flow control, message duplication and ordering, message destinations and its failure model. (Level [2], CO [2], PO [2])

10 M

**OR**

- 3 a. Explain the term marshalling and unmarshalling. Write the CORBA CDR message (flattened form) for the Person struct with value {'Smith', 'London', 1934}. (Level [L2,L3], CO [2], PO [2])

10 M

- b. Explain the implementation of RMI in detail with a neat diagram. (Level [L2], CO [2], PO [2])

10 M

**UNIT - III**

- 4 a. What are threats and attacks related to distributed system security. (Level [1], CO [3], PO [1])

06 M

- b. Illustrate the use of cryptography in the implementation of secure systems. (Level [2], CO [3], PO [1,2])

08 M

- c. Explain the distributed file system requirements. (Level [2], CO [ ], PO [1])

06 M

**UNIT - IV**

- 5 a. Explain the Network Time Protocol with a neat diagram. (Level [L2], CO [4], PO [1])

10 M

- b. Explain and distinguish the following terms  
a) Clock Skew and Clock drift b) External and Internal Synchronization. (Level [L2], CO [4], PO [1])

10 M

**OR**

- 6 a. Explain ring based election algorithm with a neat diagram. (Level [L2], CO [4], PO [1])

10 M

- b. Explain and compare Christian's method of Synchronizing clocks and Berkeley algorithm. 10 M

( Level [L2,L4], CO [4], PO [1] )

- 7 a. Analyze the most obvious obstacles of Cloud Computing. 06 M

( Level [4], CO [4], PO [2] )

- b. With a neat diagram explain the structure of the three delivery models of Cloud Computing. 08 M

( Level [2], CO [4], PO [1] )

- c. What are the major challenges faced by Cloud Computing? 06 M

( Level [1], CO [4], PO [3, 12] )

**OR**

- 8 a. Analyze the reasons for the success of Cloud Computing. 06 M

( Level [4], CO [4], PO [3, 12] )

- b. With the help of a neat diagram explain the Cloud Computing reference models and services. 08 M

( Level [2], CO [4], PO [1] )

- c. Explain the ethical issues in Cloud Computing. 06 M

( Level [2], CO [4], PO [1] )

**Sixth Semester B.E MAKEUP Examination, AUGUST OCTOBER 2021****ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

Time: 3 hrs

Max. Marks :100

Instructions :1. Answer any five full questions.

1a. What Is Artificial Intelligence? Explain in brief its importance.

L CO PO M

[2] [1] [1] [5]

1b. Compare Strong AI and Weak AI methods.

[3] [1] [1] [7]

1c. Demonstrate the Semantic Net with an example.

[3] [1] [2] [8]

2a. For a problem of Missionaries and Cannibals in AI write the Operators and draw the Search tree without cycles having solution for the same.

[3] [2] [2] [10]

2b. Write a note on Combinatorial Explosion and Problem reduction.

[2] [1] [1] [10]

3a. Explain how searching helps in problem solving emphasizing on types of searches in AI.

[2] [1] [1] [10]

3b. Demonstrate the working of Depth first search and Breadth first search algorithms by the use of either algorithm or pseudo-code for the same.

[3] [2] [2] [10]

4a. List and Explain the properties of Search Methods.

[2] [1] [1] [10]

4b. Explain with figure the three problems that could be faced by a HILL climbing algorithmic techniques.

[2] [2] [2] [10]

5a. Demonstrate the use of Game trees in solving the tic-tac-toe problem. Draw the partial game tree for the same.

[3] [2] [2] [10]

5b. Write a note on Alpha-beta Pruning emphasizing on its effectiveness and its implementation.

[2] [1] [1] [10]

6a. State the deduction theorem and apply the same to prove the following.

$$\{A \rightarrow B\} \vdash A \rightarrow (C \rightarrow B)$$

[3] [2] [2] [10]

6b. Write a note on Soundness, Completeness, Decidability, Monotonicity.

[2] [1] [1] [10]

7a. What is need of training in Machine Learning? Using a simple learning method derive a final hypothesis which is consistent for following training data:

<slow, wind, 30ft, 0, evening, cold>  
<slow, rain, 20ft, 0, evening, warm>  
<slow, snow, 30ft, 0, afternoon, cold>

[3] [3] [2] [10]

7b. Explain the candidate elimination technique and Meaning of Inductive bias.

[2] [3] [1] [10]

8a. Explain in brief the three types of learning methodologies in Artificial Neural Networks.

[2] [3] [1] [10]

8b. Demonstrate the working of simple perceptron to represent the learning of logical OR function for maximum 3 epochs.

[3] [3] [2] [10]

9a. Explain with an example the working of Probabilistic Reasoning and Joint Probability Distributions.

[2] [3] [1] [10]

9b. In the city of Cambridge, there are two taxi companies. One taxi company uses yellow taxis, and the other uses white taxis. The yellow taxi company has 90 cars, and the white taxi company has just 10 cars. A hit-and-run incident has been reported, and an eye witness has stated that she is certain that the car was a white taxi. Further suppose that experts have asserted that given the foggy weather at the time of the incident, the witness had a 75% chance of correctly identifying the taxi.

Given that the lady has said that the taxi was white, what is the likelihood that she is right?

[4] [3] [2] [10]

10a. Explain how learning happens in Simple Bayesian Concept Learning?

[2] [3] [1] [10]

10b. Write a note on Bayesian Belief Networks and The Noisy-V Function.

[2] [3] [1] [10]

**Seventh Semester B.E FASTTRACK Examination, AUGUST SEPTEMBER 2021****ARTIFICIAL INTELLIGENCE**

Time: 3 hrs

Max. Marks :100

**Instructions :** Data, if necessary, may be assumed. 3 Precise and accurate answers expected. 1 Sketches, when required, may be drawn. Answer any Five full Questions.

L CO PO M

1a. Discuss the difference between strong methods and weak methods. Explain the Turing test in detail. [2] [1] [1] [10]

1b. Provide a brief overview of PROLOG and LISP languages and explain how they are used in Artificial Intelligence research. [3] [3] [2] [10]

2a. Explain simple semantic net with an example. Also represent the same using Frames demonstrating the inheritance. [2] [1] [1, 2] [10]

2b. Explain the properties of Semantic tree. Define the search tree. Build the search tree for the given problem.

Three missionaries and three cannibals are on one side of a river, with a canoe. They all want to get to the other side of the river. The canoe can only hold one or two people at a time. At no time should there be more cannibals than missionaries on either side of the river, as this would probably result in the missionaries being eaten.

[3] [1] [2] [10]

3a. Demonstrate constructing a goal tree for Towers of Hanoi problem with four disks. [2] [2] [2] [10]

3b. List and explain the Properties of Search Methods. [2] [1] [2] [10]

4a. Discuss the following in brief (i) Hill Climbing (ii) Beam searching [2] [2] [2] [10]

4b. Write the pseudocode for implementing of depth-first search and Breadth first search functions. [2] [2] [2] [10]

5a. Prove the following:  $(A \rightarrow B) \rightarrow (B \rightarrow C \rightarrow C \rightarrow D \rightarrow A \rightarrow D)$  [5] [2] [1] [10]

5b. Explain the Deduction Theorem, with the rule: if  $A \cup B \vdash C$  then  $A \vdash (B \rightarrow C)$  [2] [2] [1] [10]

6a. Translate from English to Logic notations, making use of symbols R=raining, T=Tuesday, N=New York, S=sick, T=tired:

1.“It is raining and it is Tuesday”

2.“It is raining in New York”

3.“It is raining in New York, and I’m getting sick or just very tired”

[5] [2] [1] [10]

6b. Discuss briefly

- 1.Soundness
- 2.Completeness
- 3.Decidability
- 4.Monotonicity

[2] [2] [1] [10]

7a. Solve the phrase to get Conjunctive Normal Form (CNF):(A  $\rightarrow$  B)  $\rightarrow$  C

[3] [2] [2] [10]

**Seventh Semester B.E FASTTRACK Examination, AUGUST SEPTEMBER 2021**

**ARTIFICIAL INTELLIGENCE**

Time: 3 hrs

Max. Marks :100

Instructions :Data, if necessary, may be assumed. 3 Precise and accurate answers expected. 1 Sketches, when required, may be drawn. Answer any Five full Questions.

1a. Discuss the difference between strong methods and weak methods. Explain the Turing test in detail. L CO PO M

1b. Provide a brief overview of PROLOG and LISP languages and explain how they are used in Artificial Intelligence research. [2] [1] [1] [10]

2a. Explain simple semantic net with an example. Also represent the same using Frames demonstrating the inheritance. [3] [3] [2] [10]

2b. Explain the properties of Semantic tree. Define the search tree. Build the search tree for the given problem. [2] [1] [1, 2] [10]

Three missionaries and three cannibals are on one side of a river, with a canoe. They all want to get to the other side of the river. The canoe can only hold one or two people at a time. At no time should there be more cannibals than missionaries on either side of the river, as this would probably result in the missionaries being eaten.

3a. Demonstrate constructing a goal tree for Towers of Hanoi problem with four disks. [3] [1] [2] [10]

3b. List and explain the Properties of Search Methods. [2] [1] [2] [10]

4a. Discuss the following in brief (i) Hill Climbing (ii) Beam searching [2] [2] [2] [10]

4b. Write the pseudocode for implementing of depth-first search and Breadth first search functions. [2] [2] [2] [10]

5a. Prove the following  $\vdash (A \rightarrow B) \rightarrow (B \rightarrow C \rightarrow C \rightarrow D \rightarrow A \rightarrow D)$  [5] [2] [1] [10]

5b. Explain the Deduction Theorem, with the rule:if  $A \cup B \sqsubseteq C$  then  $A \sqsubseteq (B \rightarrow C)$  [2] [2] [1] [10]

6a. Translate from English to Logic notations, making use of symbols R=raining, T=Tuesday, N=New York, S=sick, T=tired:

1.“It is raining and it is Tuesday”

2.“It is raining in New York”

3.“It is raining in New York, and I’m getting sick or just very tired”

[5] [2] [1] [10]

6b. Discuss briefly

- 1.Soundness
- 2.Completeness
- 3.Decidability
- 4.Monotonicity

[2] [2] [1] [10]

7a. Solve the phrase to get Conjunctive Normal Form (CNF):(A → B ) → C

[3] [2] [2] [10]

**Seventh Semester B.E FASTTRACK Examination, AUGUST SEPTEMBER 2021****ARTIFICIAL INTELLIGENCE**

Time: 3 hrs

Max. Marks :100

**Instructions :**Data, if necessary, may be assumed. 3 Precise and accurate answers expected. 1 Sketches, when required, may be drawn. Answer any Five full Questions.

L CO PO M

1a. Discuss the difference between strong methods and weak methods. Explain the Turing test in detail. [2] [1] [1] [10]

1b. Provide a brief overview of PROLOG and LISP languages and explain how they are used in Artificial Intelligence research. [3] [3] [2] [10]

2a. Explain simple semantic net with an example. Also represent the same using Frames demonstrating the inheritance. [2] [1] [1, 2] [10]

2b. Explain the properties of Semantic tree. Define the search tree. Build the search tree for the given problem. [3] [1] [2] [10]

Three missionaries and three cannibals are on one side of a river, with a canoe. They all want to get to the other side of the river. The canoe can only hold one or two people at a time. At no time should there be more cannibals than missionaries on either side of the river, as this would probably result in the missionaries being eaten. [3] [1] [2] [10]

3a. Demonstrate constructing a goal tree for Towers of Hanoi problem with four disks. [2] [2] [2] [10]

3b. List and explain the Properties of Search Methods. [2] [1] [2] [10]

4a. Discuss the following in brief (i) Hill Climbing (ii) Beam searching [2] [2] [2] [10]

4b. Write the pseudocode for implementing of depth-first search and Breadth first search functions. [2] [2] [2] [10]

5a. Prove the following: $\vdash (A \rightarrow B) \rightarrow (B \rightarrow C \rightarrow C \rightarrow D \rightarrow A \rightarrow D)$  [5] [2] [1] [10]

5b. Explain the Deduction Theorem, with the rule:if  $A \cup B \sqsubseteq C$  then  $A \sqsubseteq (B \rightarrow C)$  [2] [2] [1] [10]

6a. Translate from English to Logic notations, making use of symbols R=raining, T=Tuesday, N=New York, S=sick, T=tired:

1.“It is raining and it is Tuesday”

2.“It is raining in New York”

3.“It is raining in New York, and I’m getting sick or just very tired”

[5] [2] [1] [10]

6b. Discuss briefly

1.Soundness

2.Completeness

3.Decidability

4.Monotonicity

[2] [2] [1] [10]

7a. Solve the phrase to get Conjunctive Normal Form (CNF):(A  $\rightarrow$  B)  $\rightarrow$  C

[3] [2] [2] [10]

7b. How to build a medical expert system using backward chaining in Rule-Based Expert System. [3] [2] [2] [10]

8a. Demonstrate Proof by Reputation using an example. [2] [2] [1] [10]

8b. Explain the Resolution Rule and using the Resolution Rule resolveA, B, ( $\neg$ B, C) [2] [2] [1] [10]

9a. Using block schematic, discuss why the blackboard architecture is an effective way to combine information from a number of knowledge sources. Also explain the main components of the blackboard architecture. [2] [1] [1] [10]

9b. Explain the architecture of the Copycat system. Solve the analogies using copycat system. [2] [1] [1] [10]

i. AABB is to AACC as JJKK is to what? [3] [1] [2] [10]

ii. ABB is to ABCC as JKK is to what? [2] [1] [2] [10]

iii. AABC is to AABD as IJKK is to what? [2] [1] [2] [10]

iv. is to Z as EFG is to what? [2] [1] [2] [10]

v. FSP is to SFS as ABBBC is to what? [2] [1] [2] [10]

10a. Define an agent? List and explain five properties of agents. Explain briefly five types of agents. [2] [1] [2] [10]

10b. Explain with block diagram three-layer subsumption architecture for an agent. [2] [1] [2] [10]

[2] [1] [2] [10]

**Seventh Semester B.E. Makeup Examination, January 2020****ARTIFICIAL INTELLIGENCE**

Time: 3 Hours

Max. Marks: 100

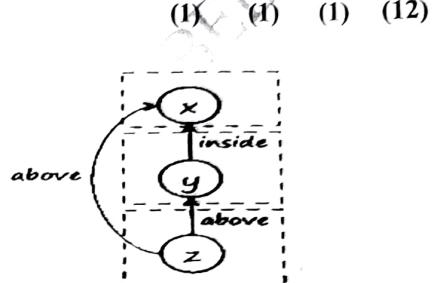
S63

- Instructions:**
1. Precise and accurate answers expected.
  2. Data, if necessary, may be assumed.
  3. Sketches, when required, may be drawn.

**UNIT - I**

L CO PO M

- 1 a. Define (i) Intelligence (ii) Artificial intelligence (iii) Strong Methods and Weak Methods  
And discuss the Chinese Room Experiment in detail.
- c. Design Semantic net  
with appropriate frames  
for the following  
objects having relations to  
each other:



(1) (1) (1) (12)

**OR**

(3) (3) (2) (08)

- 2 a. Construct a simple semantic net, for the following: Bob as builder, Fido as dog, and Fang as cat and Mice. Also represent the same using Frames demonstrating the inheritance.

(3) (3) (2) (10)

- b. Explain the properties of Semantic tree. Define the search tree. Build the search tree for the given problem.

Three missionaries and three cannibals are on one side of a river. They all want to get to the other side of the river. The canoe can only hold one or two people at a time. At no time should there be more cannibals than missionaries on either side of the river, as this would probably result in the missionaries being eaten.

(3) (3) (2) (10)

**UNIT - II**

L CO PO M

- 3 a. Explain DFS and BFS in brief. Also explain the properties of search methods. Build a search tree for Searching for a Gift.

(3) (1) (1) (10)

- b. Write the pseudo code for implementing of depth-first search. And analyze the given tree in Fig.1 using DFS.

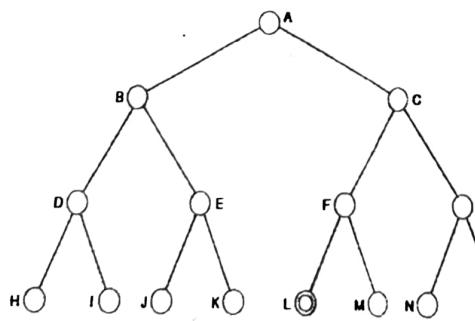


Fig.1. Tree

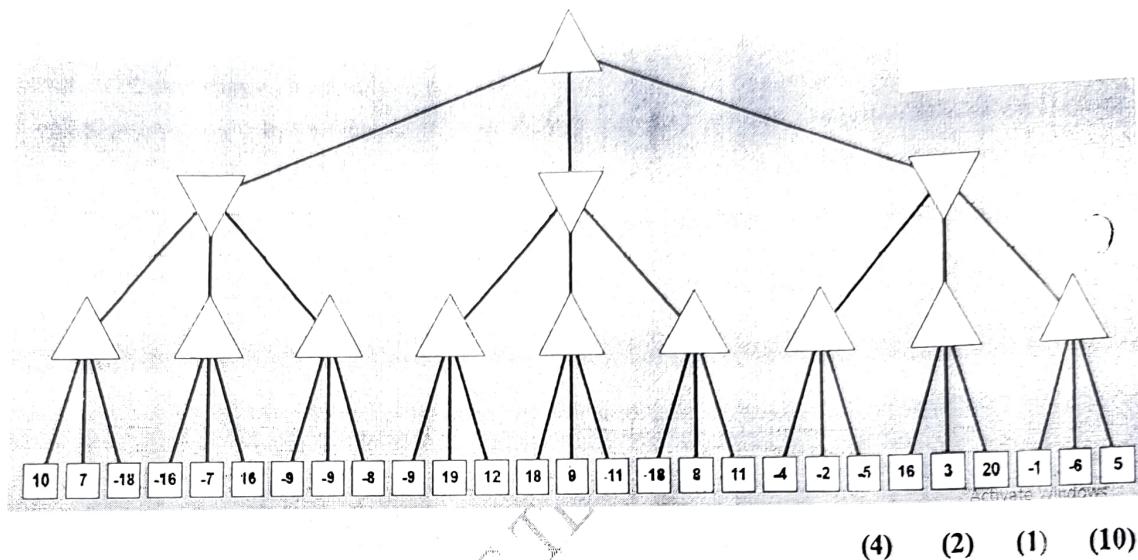
(3) (1) (1) (10)

**OR**

- 4 a. List and explain the Properties of Search Methods. (2) (1) (1) (08)
- b. Develop the various heuristics approaches for 8-puzzle. The puzzle consists of a 3 X 3 grids, with the numbers 1 through 8 on tiles within the grid and one blank square. Tiles can be slid about within the grid, but a tile can only be moved into the empty square if it is adjacent to the empty square. (3) (2) (2) (12)
- L CO PO M

**UNIT - III**

- 5 a. Apply and analyze alpha beta pruning algorithm on the given graph



(4) (2) (1) (10)

- b. Explain deduction theorem with example. Explain the relationship between universal and existential quantifiers. (2) (3) (2) (10)

**OR**

- 6 a. Explain with relevant examples why the alpha-beta procedure will always generate the same answer as Minimax without pruning. Why is it useful? (2) (3) (2) (10)

- b. Prove the following

$$\begin{array}{l} \text{i)} \quad (A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow ((C \rightarrow D) \rightarrow (A \rightarrow D))) \\ \text{ii)} \quad (\neg A \rightarrow B) \rightarrow (\neg B \rightarrow A) \end{array}$$

(3) (3) (2) (10)  
L CO PO M

**UNIT - IV**

- 7 a. Solve the phrase to get Conjunctive Normal Form (CNF):  $(A \rightarrow B) \rightarrow C$

(3) (2) (2) (06)

- b. Explain the Resolution Rule and using the Resolution Rule resolve  $\{A, B\}, \{\neg B, C\}$

(2) (1) (1) (06)

- c. With block diagram, explain the architecture of an Expert System.

(2) (1) (1) (08)  
L CO PO M

**OR**

- 8 a. Solve the phrase to get Conjunctive Normal Form (CNF):  $A \leftrightarrow (B \wedge C)$

(3) (2) (1) (06)

- b. Illustrate "Proof by Reputation" using an example.

(2) (2) (1) (06)

- c. Explain how to build a medical expert system using backward chaining in Rule-Based Expert System.

(2) (2) (1) (08)

**UNIT - V**

L CO PO M

- 9 a. Using block schematic, discuss why the blackboard architecture is an effective way to combine information from a number of knowledge sources. Also explain the main components of the blackboard architecture. (2) (1) (1) (10)

- b. Using block schematic, explain various kinds of problems the Copycat Architecture can solve. (2) (1) (1) (10)

L CO PO M

**OR**

- 0 a. Define an agent? List and explain five properties of agents. Explain briefly five types of agents. (2) (1) (1) (10)

- b. Explain with block diagram three-layer subsumption architecture for an agent. (2) (1) (1) (05)

- c. Compare Horizontal and Vertical agent architectures. (2) (1) (1) (05)

DO NOT WRITE

**Seventh Semester B.E. Semester End Examination, Dec./Jan. 2019-20****ARTIFICIAL INTELLIGENCE**

Time: 3 Hours

Max. Marks: 100

*Instructions: 1. Answer one full question each from the Units***UNIT - I**

L CO PO M

- 1 a. Explain Turing test and Chinese room argument experiment. Compare and contrast. (2) (3) (2) (10)
- b. Convert the following information into semantic nets and frames:  
 Tom is a cat. Tom caught a bird. Tom is owned by John. Tom is ginger in color. Cats like cream.  
 The cat sat on the mat. A cat is a mammal. A bird is an animal. All mammals are animals. Mammals have fur. (3) (2) (2) (10)

**OR**

- 2 a. Explain the relationship between graphs, semantic nets, semantic trees, search spaces, and search trees. (2) (3) (2) (10)

- b. Design a suitable representation and draw the complete search tree for the following problem.  
 A farmer is on one side of a river and wishes to cross the river with a wolf, a chicken, and a bag of grain. He can take only one item at a time in his boat with him. He can't leave the chicken alone with the grain, or it will eat the grain, and he can't leave the wolf alone with the chicken, or the wolf will eat the chicken. How does he get all three safely across to the other side? (5) (3) (2) (10)

L CO PO M

**UNIT - II**

- 3 a. Explain how you implement Depth First Search and Breadth First Search with example code. Compare both search techniques. (2) (3) (2) (10)
- b. Explain with example how you use heuristics for search. Explain the criteria for selecting a good heuristic. (2) (3) (2) (10)

**OR**

- 4 a. Explain different techniques to identify optimal paths. (2) (3) (2) (10)
- b. Implement a greedy-search algorithm. How well does it perform compared with the other methods you have implemented? Invent a 0-1 knapsack problem, and use your search tree implementation to model this problem. Can you model the fractional knapsack problem using a search tree? (5) (2) (1) (10)

L CO PO M

**UNIT - III**

- 5 a. Explain the following terms (i) Game Trees (ii) Minimax (ii) Alpha beta pruning. (2) (2) (3) (12)
- b. What is Logic? Explain Why Logic is used in Artificial Intelligence and explain Logical Operators. (2) (2) (3) (08)

**OR**

- 6 a. Explain the concepts of Translating between English and Logic Notation and explain the following Truth Tables of Not, And, Or, Implies, if, Complex Truth Tables. (2) (3) (2) (12)
- b. Explain deduction Theorem with an example. (2) (4) (4) (08)

## **UNIT - IV**

L CO PO

- 7 a. Consider the following axioms and convert them to clausal form

- a) Every coyote chases some roadrunner.
- b) Every roadrunner who says "beep-beep" is smart.
- c) No coyote catches any smart roadrunner.
- d) Any coyote who chases some roadrunner but does not catch it is frustrated.
- e) (Conclusion) If all roadrunners say "beep-beep", then all coyotes are frustrated.

Prove the conclusion If all roadrunners say "beep-beep", then all coyotes are frustrated

(3) (3) (2) (10)

- b. Explain with examples resolution in propositional logic.

(2) (3) (2) (10)

### **OR**

- 8 a. Explain with examples resolution in predicate logic.

(2) (3) (2) (10)

- b. Explain with example backward chaining.

(2) (3) (2) (10)

## **UNIT - V**

L CO PO M

- 9 a. Explain black board architecture with its implementation.

(2) (3) (2) (10)

- b. Explain with example Dempster-Shafer theory of evidence.

(2) (3) (2) (10)

### **OR**

- 10 a. List and explain properties of agents.

(2) (1) (2) (10)

- b. List and explain types of agents.

(2) (1) (2) (10)

**Seventh Semester B.E. Makeup Examination, January 2019**  
**ARTIFICIAL INTELLIGENCE**

Time: 3 Hours

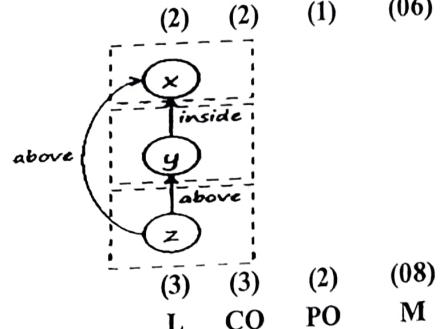
Max. Marks: 100

**Instructions:** 1. Unit-I and Unit-II are compulsory  
 2. Answer any one full question from each of the remaining units.

**UNIT - I**

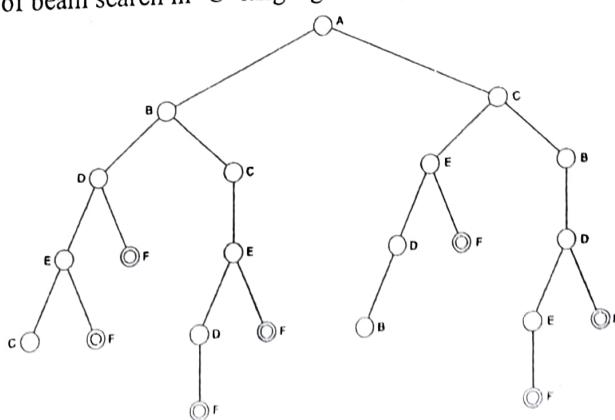
L	CO	PO	M

- 1 a. Define
- Intelligence
  - Artificial intelligence
  - Strong Methods and Weak Methods
- (1) (1) (1) (06)
- b. Explain the Chinese Room argument, and
- present some of the arguments against it, and
  - the counter-arguments.
  - Which do you find most convincing? How does this affect your view on the overall worth of the study of Artificial Intelligence?
- (2) (2) (1) (06)
- c. Design Semantic net with appropriate frames for the following objects having relations each other:

**UNIT - II**

L	CO	PO	M

- 2 a. Explain the differences and similarities between depth-first search and breadth-first search. Give examples of the kinds of problems where each would be appropriate. Mention the time and space complexity of depth-first search and breadth-first search.
- (2) (3) (2) (10)
- b. Write implementation of beam search in 'C' language. Analyze beam search for the following tree.



L	CO	PO	M

**UNIT - III**

- 3 a. Prove the following: $\vdash (\forall \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow ((C \rightarrow D) \rightarrow (\forall \rightarrow D)))$
- (3) (2) (1) (06)

b. Explain the Deduction Theorem, with the rule: if  $A \cup \{B\} \vdash C$  then  $A \vdash (B \rightarrow C)$

(2) (1) (1)

c. Explain briefly

- a. Soundness
- b. Completeness
- c. Decidability
- d. Monotonicity

(2) (1) (2) (08)

### OR

- 4 a. i. What is Logic?  
ii. Why Logic is used in Artificial Intelligence?  
iii. List five Logical Operators.
- b. Translate from English to Logic notations, making use of symbols R=raining, T=Tuesday, N=New York, S=sick, T=tired:  
i. "It is raining and it is Tuesday"  
ii. "It is raining in New York"  
iii. "It is raining in New York, and I'm getting sick or just very tired"

(1) (1) (2) (08)

- c. Identify the meaning of Deduction, for the logical expressions:  
a.  $\frac{A \quad B}{A \wedge B}$  b.  $\frac{A \wedge B}{A}$  c.  $\frac{A \quad A \rightarrow B}{B}$

(2) (2) (1) (06)  
(3) (2) (2) (06)

### UNIT - IV

L CO PO M

- 5 a. Explain Skolemization and Unification with examples  
b. List rules to convert a wff to prenex normal form. Also list the rules to move quantifiers to the front.

(2) (3) (2) (10)  
(1) (3) (2) (10)

### OR

- 6 a. Explain with example forward chaining and backward chaining.  
b. Write a short note on  
i) CLIPS  
ii) CYC

(2) (3) (2) (10)

### UNIT - V

L CO PO M

- 7 a. Explain why the blackboard architecture is an effective way to combine information from a number of knowledge sources, include its main components and explain the block schematic.  
b. Explain what kinds of problems, the Copycat architecture can solve, using block schematic.

(2) (1) (1) (10)  
(2) (1) (1) (10)

### OR

- 8 a. i. Define an agent?  
ii. List and explain five properties of agents.  
iii. Explain briefly five types of agents.
- b. i. Explain with block diagram a three-layer subsumption architecture for an agent.  
ii. Compare Horizontal and Vertical agent architectures.

(2) (1) (1) (10)  
(2) (1) (1) (10)

Seventh Semester B.E. Semester End Examination, Dec/Jan 2018-19  
**ARTIFICIAL INTELLIGENCE**

Max. Marks: 100

Time: 3 Hours

- Instructions:**
1. Unit I and II are compulsory.
  2. Readable question's numbers must be in allocated space.
  3. Data, if necessary, may be assumed.
  4. Sketches, when required, may be drawn.

L CO PO M

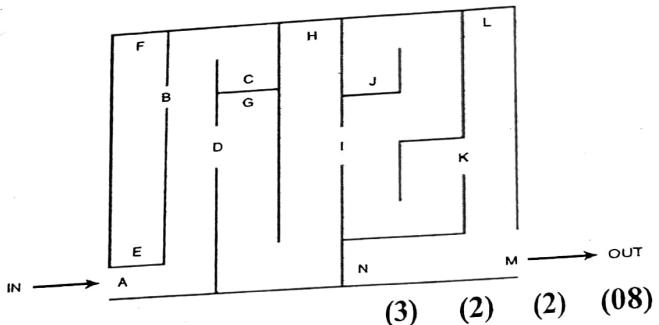
**UNIT - I**

- 1 a. Differentiate between  
 i) Strong methods and weak methods (3)  
 ii) Strong AI and weak AI (3) (2) (05)
- b. Convert the following information into:  
 a) a semantic net  
 b) a frame based representation (3) (3) (2) (10)  
 A Ford is a type of car. Bob owns two cars. Bob parks his car at home. His house is in California, which is a state. Sacramento is the state capital of California. Cars drive on the freeway, such as Route 101 and Highway 81.
- c. Differentiate between a top-down approach and a bottom-up approach to problem solving. In what kind of situations might each be appropriate (3) (3) (2) (05)

L CO PO M

**UNIT - II**

- 2 a. Refer to maze, for entry, exit, dead nodes, decision nodes – identify path from entry to exit, and build an appropriate search tree.



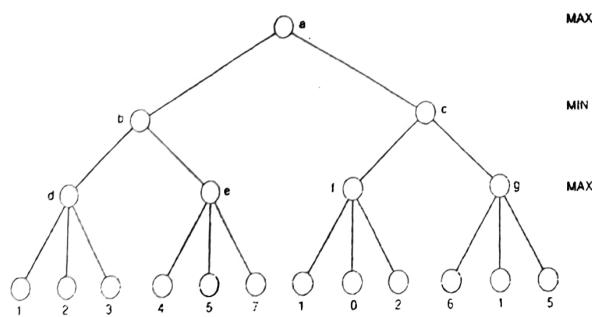
(3) (2) (2) (08)

- b. Summarize the Properties of Search Methods. (1) (1) (1) (04)
- c. Solve the eight-queens problem, using Constraint Satisfaction Search method, list out the steps and chess-board sketch, too. (3) (2) (2) (08)

L CO PO M

**UNIT - III**

- 3 a. Apply and analyze alpha-beta pruning algorithm for the following game tree.



(4) (2) (2) (10)

# Seventh Semester B.E. Semester End Examination, Dec/Jan 2018-19

## ARTIFICIAL INTELLIGENCE

Max. Marks: 100

Time: 3 Hours

**Instructions:** 1. Unit I and II are compulsory.

2. Readable question's numbers must be in allocated space.

3. Data, if necessary, may be assumed.

4. Sketches, when required, may be drawn.

### **UNIT - I**

I.	CO	PO	M
1. Differentiate between i) Strong methods and weak methods ii) Strong AI and weak AI	(3)	(3)	(2)

b. Convert the following information into:

a) a semantic net

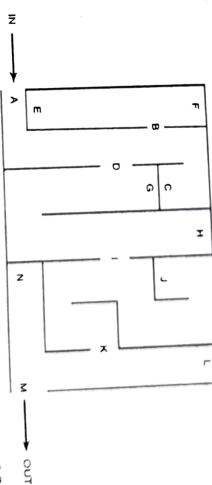
b) a frame based representation  
A Ford is a type of car. Bob owns two cars. Bob parks his car at home. His house is in California. A Ford is a type of car. Bob owns two cars. Cars drive on the freeway, such as which is a state. Sacramento is the state capital of California. Cars drive on the freeway, such as Route 101 and Highway 81.

c. Differentiate between a top-down approach and a bottom-up approach to problem solving. In what kind of situations might each be appropriate

(3) (3) (2) (05)

### **UNIT - II**

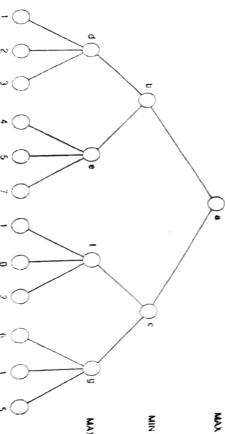
2. a. Refer to maze, fo entry, exit, dead nodes decision nodes – identify path from entry to exit, and build an appropriate search tree.



- b. Summarize the Properties of Search Methods.  
c. Solve the eight-queens problem, using Constraint Satisfaction Search method, list out the steps and chess-board sketch, too.

### **UNIT - III**

3. a. Apply and analyze alpha-beta pruning algorithm for the following game tree.



(3)	(2)	(2)	(08)
L CO PO M	(4)	(2)	(2)

b. Translate the following sentences in to predicate logic:

- i) Marcus was a man
- ii) All Pompeians were Romans
- iii) Everyone is loyal to someone
- iv) Marcus tried to assassinate Caesar
- v) All men are people

(2) (2) (2) (10)

**OR**

- 4 a. Explain effectiveness of alpha-beta pruning algorithm. Explain how alpha-beta pruning algorithm is implemented with code.

(2) (3) (2) (10)

- b. Prove the following:

- a)  $(\neg A \rightarrow B) \rightarrow (\neg B \rightarrow A)$
- b)  $(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow ((C \rightarrow D) \rightarrow (A \rightarrow D)))$

**UNIT - IV**

- 5 a. Solve the phrase to to get Conjunctive Normal Form (CNF):  $(A \rightarrow B) \rightarrow C$

(3) (2) (2) (10)

- b. Explain the Resolution Rule and using the Resolution Rule resolve  $\{A, B, (\neg B, C)\}$

(2) (1) (1) (10)

- c. Illustrate with block diagram and explain the architecture of an Expert System.

(2) (1) (1) (10)

**OR**

- 6 a. Solve the phrase to to get Conjunctive Normal Form (CNF):  $A \leftrightarrow (B \wedge C)$

(3) (2) (1) (06)

- b. Explain the method of Proof by Reputation using an example.

(2) (2) (1) (05)

- c. How to build an medical expert system using backward chaining in Rule-Based Expert System.

(1) (2) (1) (08)

**UNIT - V**

- 7 a. Write a note on

- i) Blackboard Architecture
- ii) Copycat Architecture

(2) (3) (2) (10)

- b. Explain Dempster Shafer Theory of evidence with an example

(2) (3) (2) (10)

**OR**

- 8 a. Explain properties of agents

(2) (1) (2) (10)

- b. Explain with block diagram a three-layer subsumption architecture for an agent with suitable examples

(2) (1) (2) (10)

*B.Tech S.P.Q.P/R, 2018*

## Sixth Semester B.E. Semester End Examination, May/June 2018-19

### DATA MINING

Time: 3 Hours

Max. Marks: 100

- Instructions:**
1. Unit I and Unit II are compulsory.
  2. Answer any one full question from remaining units
  3. Use suitable examples wherever needed.

**UNIT – I (Compulsory)**

- 1 a. What is data warehouse? List and explain the general guidelines for implementing data warehouse. (1) (1) (1) (10)
- b. Explain the process of ETL with an example (2) (1) (1) (10)

**UNIT – II (Compulsory)**

- 2 a. What is an attribute? List and explain different attribute types (2) (2) (1) (05)
- b. Describe data preprocessing. Summarize the different strategies/techniques available for data preprocessing (2) (2) (1) (10)
- c. Compute the cosine similarity for the given document vectors  
 $X = (3, 2, 0, 5, 0, 0, 0, 2, 0, 8)$   $Y = (1, 0, 0, 0, 0, 1, 0, 1, 0, 2)$  (3) (2) (2) (05)

**UNIT – III**

- 3 a. Summarize how to extract association rules efficiently from a given frequent dataset. (3) (3) (2) (10)
- b. Explain the FP growth algorithm with an example (3) (3) (1) (10)

- 4 a. For the transaction data set,

TID	Items
1	{a,b}
2	{b,c,d}
3	{a,c,d,e}
4	{a,d,e}
5	{a,b,c}
6	{a,b,c,d}
7	{a}
8	{a,b,c}
9	{a,b,d}
10	{b,c,e}

Construct the FP tree and explain the steps followed in constructing it. (3) (3) (3) (05)

- b. Explain how rule generation is done in Apriori algorithm with a pseudocode for the same (2) (3) (1) (10)
- c. Discuss the concept of Support & Confidence (2) (3) (1) (05)

**UNIT – IV**

- | Customer ID | Gender | Car Type | Shirt Size  | Class |
|-------------|--------|----------|-------------|-------|
| 1           | M      | Family   | Small       | C0    |
| 2           | M      | Sports   | Medium      | C0    |
| 3           | M      | Sports   | Medium      | C0    |
| 4           | M      | Sports   | Large       | C0    |
| 5           | M      | Sports   | Extra Large | C0    |
| 6           | M      | Sports   | Extra Large | C0    |
| 7           | F      | Sports   | Small       | C0    |
| 8           | F      | Sports   | Small       | C0    |
| 9           | F      | Sports   | Medium      | C0    |
| 10          | F      | Luxury   | Large       | C0    |
| 11          | M      | Family   | Large       | C1    |
| 12          | M      | Family   | Extra Large | C1    |
| 13          | M      | Family   | Medium      | C1    |
| 14          | M      | Luxury   | Extra Large | C1    |
| 15          | F      | Luxury   | Small       | C1    |
| 16          | F      | Luxury   | Small       | C1    |
| 17          | F      | Luxury   | Medium      | C1    |
| 18          | F      | Luxury   | Medium      | C1    |
| 19          | F      | Luxury   | Medium      | C1    |
| 20          | F      | Luxury   | Large       | C1    |
- Note: L (level), CO (Course Outcome), PO (Programme Outcome), M (Marks)**
- 5 a. What is decision tree? Construct a decision tree for mammal classification problem. (3) (3) (2) (08)
- b. Consider the training examples shown in Table for a binary classification problem.

Customer ID	Gender	Car Type	Shirt Size	Class
1	M	Family	Small	C0
2	M	Sports	Medium	C0
3	M	Sports	Medium	C0
4	M	Sports	Large	C0
5	M	Sports	Extra Large	C0
6	M	Sports	Extra Large	C0
7	F	Sports	Small	C0
8	F	Sports	Small	C0
9	F	Sports	Medium	C0
10	F	Luxury	Large	C0
11	M	Family	Large	C1
12	M	Family	Extra Large	C1
13	M	Family	Medium	C1
14	M	Luxury	Extra Large	C1
15	F	Luxury	Small	C1
16	F	Luxury	Small	C1
17	F	Luxury	Medium	C1
18	F	Luxury	Medium	C1
19	F	Luxury	Medium	C1
20	F	Luxury	Large	C1

- a. Compute the Gini index for the overall collection of training examples.
- b. Compute the Gini index for the Customer ID attribute.
- c. Compute the Gini index for the Gender attribute.
- d. Compute the Gini index for the Car Type attribute.
- e. Compute the Gini index for the Shirt Size attribute using multiway split.
- f. Which attribute is better, Gender, Car Type, or Shirt Size?
- g. Explain why Customer ID should not be used as the attribute test condition even though it has the lowest Gini.

- c. Write a note on Gain ratio. (3) (3) (2) (08)

6 a. Explain the k-nearest neighbor algorithm. Explain 1,2,3- nearest neighbor of an instance (3) (3) (1) (10)

- b. Illustrate Hunt's algorithm to induce decision tree with an example (3) (3) (1) (10)

7 a. List and explain the features of cluster analysis (3) (3) (1) (10)

- b. With a neat diagram explain taxonomy of cluster analysis methods (3) (4) (3) (10)

- OR  
L CO PO M

8 a. Explain K-means algorithm with an example (3) (3) (1) (10)

- b. Write a note on Quality and validity of cluster analysis methods (2) (3) (2) (10)

**OR**

KLS COLLEGE INSTITUTE

## Sixth Semester B.E. Makeup Examination, May/June 2018-19

### **DATA MINING**

Max. Marks: 100

Time: 3 Hours

**Instructions:** 1. Unit I and Unit II are compulsory  
2. Answer any one full question from remaining units  
3. Use suitable examples wherever needed

#### **UNIT - I (Compulsory)**

1 a. What is ODS? Explain typical structure of ODS with a neat diagram

b. Differentiate between ODS and DW

c. Explain the implementation steps for building a data warehouse

#### **UNIT - II (Compulsory)**

2 a. What is data mining? Explain the KDD process with a neat diagram (2) (1) (1) (08)

b. List and explain the different data mining tasks with a suitable examples (2) (2) (1) (08)

c. Compute the Jaccard's coefficient and simple matching coefficient for the binary vectors.

$$X = (1, 0, 0, 1, 0, 1, 0, 1, 0, 1), Y = (1, 1, 1, 1, 0, 0, 1, 0, 1)$$

#### **UNIT - III**

3 a. What is frequent itemset generation? Explain frequent itemset generation of Apriori algorithm with a pseudocode. (3) (2) (2) (10)

b. List and explain the factors that affects computational complexity of Apriori algorithm (2) (1) (10)

**OR**

4 a. Illustrate different methods for generating frequent item set. (2) (2) (1) (10)

b. Write a note on

(a) Maximal frequent itemset (b) Closed itemset

#### **UNIT - IV**

5 a. Explain Bayesian Classifiers for the data set

Example No.	Color	Type	Origin	Stolen?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

- b. Explain how decision tree can be constructed using hunt's algorithm. Use mammal classification as an example. (3) (2) (2) (10)

Note: L (level), CO (Course Outcome), PO (Programme Outcome), M (Marks)

**OR**

- 6 a. Consider the training examples shown in Table for a binary classification problem.

Customer ID	Gender	Car Type	Shirt Size	Class
1	M	Family	Small	C0
2	M	Sports	Medium	C0
3	M	Sports	Medium	C0
4	M	Sports	Large	C0
5	M	Sports	Extra Large	C0
6	M	Sports	Extra Large	C0
7	F	Sports	Small	C0
8	F	Sports	Small	C0
9	F	Sports	Medium	C0
10	F	Luxury	Large	C0
11	M	Family	Large	C1
12	M	Family	Extra Large	C1
13	M	Family	Medium	C1
14	M	Luxury	Extra Large	C1
15	F	Luxury	Small	C1
16	F	Luxury	Small	C1
17	F	Luxury	Medium	C1
18	F	Luxury	Medium	C1
19	F	Luxury	Medium	C1
20	F	Luxury	Large	C1

- a. Compute the Gini index for the overall collection of training examples.

- b. Compute the Gini index for the Customer ID attribute.

- c. Compute the Gini index for the Gender attribute.

- d. Compute the Gini index for the Car Type attribute using multiway split.

- e. Compute the Gini index for the Shirt Size attribute using multiway split.

- b. Explain the characteristics of decision tree induction

**UNIT V**

- 7 a. What is agglomerative clustering? Explain in detail with an example

- b. List and explain the cluster analysis methods

**OR**

- 8 a. Explain the density-based clustering method in detail

- b. Explain how the clustering methods are evaluated in terms of quality and validity

(2) (3) (1) (10)

(2) (3) (1) (10)

(2) (3) (1) (10)

(2) (3) (1) (10)

**Seventh Semester B.E. Semester End Examination, Dec/Jan 2018-19****DATA MINING**

Time: 3 Hours

Max. Marks: 100

**Instructions:** 1. UNIT II & V are Compulsory.  
 2. Answer any one full question from remaining each UNITS.

**UNIT - I**

- 1 a. What is data mining? Discuss the motivating challenges in the development of data mining. (1,2) (1) (1) (8)
- b. Explain the different tasks of data mining with an example. (2) (1) (1) (8)
- c. Discuss the four types of data attributes with suitable examples. (2) (1) (1) (4)

**OR**

- 2 a. Why data preprocessing is required in data mining? Explain the various steps involved in data preprocessing. (2) (1) (1) (8)
- b. Consider the following two binary vectors.  
 $X=(1,0,0,1,0,0,1,0,0,1)$   $Y=(1,0,0,1,0,0,1,0,0,1)$   
 Find:  
 i) Simple matching coefficient(SMC)      ii) Jaccard coefficient  
 iii) Cosine similarity      iv) Extended Jaccard coefficient (3) (1) (4) (8)
- c. Discuss the various applications of data mining. (2) (1) (1) (4)

**UNIT - II**

- 3 a. Define classification. Explain the general approach of solving a classification problem. (2) (3) (1) (8)
- b. What is decision tree? Explain Hunt's algorithm for decision tree construction with an example. (2) (3) (4) (8)
- c. Briefly explain the Rule-based classifier with an example. (2) (3) (1) (4)

**UNIT - III**

- 4 a. Consider the following training data tuples from the AllElectronics customer database:

S.No	Age	Income	Student	Credit_rating	Class: Buys_computer
1	<=30	High	No	Fair	No
2	<=30	High	No	Excellent	No
3	31..40	High	No	Fair	Yes
4	>40	Medium	No	Fair	Yes
5	>40	Low	Yes	Fair	Yes
6	>40	Low	Yes	Excellent	No
7	31..40	Low	Yes	Excellent	Yes
8	<=30	Medium	No	Fair	No
9	<=30	Low	Yes	Fair	Yes
10	>40	Medium	Yes	Fair	Yes

11	<=30	Medium	Yes	Excellent	Yes
12	31..40	Medium	No	Excellent	Yes
13	31..40	High	Yes	Fair	Yes
14	>40	Medium	No	Excellent	No

Given a sample  $X = (\text{age} = \text{"}<=30\text{"}, \text{income} = \text{"medium"}, \text{student} = \text{"yes"}, \text{credit\_rating} = \text{"fair"})$ , predict the class label for **Buys\_computer** using Naïve Bayes classifier. (3) (3) (4) (10)

- b. Discuss the different methods for estimating accuracy of classification methods. (2) (3) (1) (10)

OR

- 5 a. Discuss Perceptron learning algorithm. Explain with a neat diagram multi layer feed forward Artificial neural network. (2) (3) (1) (10)
- b. Illustrate the basic idea behind support vector machine and discuss the characteristics and types of SVM. (2) (3) (1) (10)

#### UNIT - IV

- 6 a. What is the need of support and confidence in association rule mining? Discuss the limitations of support-confidence framework. (2) (4) (1) (5)
- b. Illustrate the Apriori algorithm for frequent itemset generation for the following data set.

TID	Items bought
T100	{K,A,D,B}
T200	{D,A,C,E,B}
T300	{C,A,B,E}
T400	{B,A,D}

Assume min\_sup = 60% and min\_conf=80%. List all frequent patterns and strong association rules. (3) (4) (4) (10)

- c. Briefly explain the factors affecting the computational complexity of Apriori algorithm. (2) (4) (1) (5)

OR

- 7 a. Explain FP-growth algorithm for generating frequent itemsets and its limitations. (2) (4) (4) (10)

b.

Consider the following transaction data set:

TID	1	2	3	4	5	6	7	8	9	10
Items	a, b	b,c, d	a,c, d,e	a,d, e	a,b, c	a,b, c,d	a	a,b, c	a,b, d	b,c, e

Construct the FP-tree. Show the trees separately after reading each transaction. (3,4) (4) (4) (10)

#### UNIT - V

- 8 a. What is cluster analysis? Explain different types of clustering. (2) (5) (1) (10)
- b. Write a note on DBSCAN algorithm with an example. (2) (5) (1) (10)

**Seventh Semester B.E. Makeup Examination, January 2019**  
**DATA MINING**

Time: 3 Hours

Max. Marks: 100

**Instructions:** 1. Answer five complete questions, Units II and V are compulsory. Answer any one question from each of the remaining units.

2. Provide examples, wherever needed.

**UNIT - I**

1. a. What is data mining and explain the process of knowledge discovery in databases with a neat diagram.  
 (2) (1) (1) (08)

b. Explain various tasks of data mining with an example.  
 (2) (1) (1) (08)

c. Explain different variations of Record data.  
 (2) (1) (1) (04)

**OR**

2. a. What is data preprocessing? List and explain the various tasks of data preprocessing.  
 (2) (1) (1) (08)

b. List the various applications of data mining.  
 (2) (1) (1) (04)

c. Explain similarity and dissimilarity for simple attributes. Consider the following two binary vectors.

$$X=(1,1,1,1,0,0,0,0,0,0) \quad Y=(1,1,0,0,0,0,1,0,0,1)$$

Find: i) Jaccard coefficient    ii) Cosine similarity    (3) (1) (4) (08)

**UNIT - II**

3. a. Define classification. Explain the general approach for solving a classification problem.  
 (1,2) (3) (1) (08)

b. Explain the algorithm for decision tree induction and also discuss each step in detail.  
 (2) (3) (1) (08)

c. Discuss the characteristics of Nearest-Neighbor classifiers.  
 (2) (3) (1) (04)

**UNIT - III**

4. a. With a neat diagram explain multilayer artificial neural network and also discuss the characteristics of artificial neural network.  
 (2) (3) (1) (10)

b. Explain support vector machine with suitable examples.  
 (2) (3) (1) (10)

**OR**

5. a. What are naïve bayes classifiers? With an example explain how naïve bayes classifier works.  
 (2) (3) (1) (10)

b. Discuss the different methods for estimating accuracy of classification methods.  
 (2) (3) (1) (10)

**UNIT - IV**

- 6 a. Consider the following transaction data set:

TID	T100	T200	T300	T400	T500	T600	T700	T800	T900
List of Item IDs	11,12, 15	12,14	12,13	11,12, 14	11,13	12,13	11,13	11,12, 13,15	11,12, 13

- i) Construct the FP tree.  
ii) Generate frequent itemsets, for the given transaction data set, using Apriori algorithm.

(3) (4) (4) (12)

- b. What is association analysis? Discuss the Apriori algorithm in the context of frequent itemsets generation for association rules.

OR

- 7 a. Discuss the necessity of support and confidence in association rule mining along with its limitations. Briefly explain the different measures used for evaluation of association patterns.

(2) (4) (4) (10)

- b. Explain FP-growth algorithm for discovering frequent itemsets. Discuss its limitations.

(2) (4) (1) (10)

**UNIT - V**

- 8 a. What is cluster analysis? How does it differ from classification? Briefly explain the basic agglomerative hierarchical clustering algorithm.

(1,2) (5) (1) (10)

- b. Explain basic DBSCAN algorithm of clustering with an example.

(2) (5) (4) (10)