7. Draw decision tree for the following data sets. Use entropy as a node selection mechanism. 14

RID	age	income	student	credit–rating	Class: buys-computer
1	youth	high	no	fair	no
2	youth	high	no	excellent	no
3	middle-aged	high	no	fair	yes
4	senior	medium	no	fair	yes
5	senior	low	yes	fair	yes
6	senior	low	yes	excellent	no
7	middle-aged	low	yes	excellent	yes
8	youth	medium	no	fair	no
9	youth	low	yes	fair	yes
10	senior	medium	yes	fair	yes
11	youth	medium	yes	excellent	yes
12	middle-aged	medium	no	excellent	yes
13	middle-aged	high	yes	fair	yes
14	senior	medium	no	excellent	no

8. (*a*) Explain the working principle of DBSCAN with example. 7

(*b*) Explain the applications of the data warehousing and data mining in Government. 7

2018

Full Marks : 70

Time : 3 hours

Answer any five questions.

The figures in the right-hand margin indicate marks.

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

1. (*a*) How is a data warehouse different from a database ? How are they similar ? 7

(*b*) Describe the steps involved in data mining when viewed as a process of knowledge discovery. 7

2. Briefly compare the following concepts. You may use an example to explain your point(s).

(*a*) Snowflake schema, fact constellation, starnet query model 7

(*b*) Data cleaning, data transformation, refresh 7

3. (*a*) In real-world data, tuples with missing values for some attributes are a common occurrence. Describe various methods for handling this problem. 7

(*b*) Consider the following data (in increasing order) for the attribute age: 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70. 7

(*i*) Use smoothing by bin means to smooth these data, using a bin depth of 3. Illustrate your steps. Comment on the effect of this technique for the given data.

(*ii*) How might you determine outliers in the data ?

(*iii*) What other methods are there for data smoothing ?

4. A database has five transactions. Let min sup D 60% and min conf D 80%.

TID	items-bought
T 100	{M, O, N, K, E, Y}
T 200	{D, O, N, K, E, Y}
T 300	{M, A, K, E}
T 400	{M, U, C, K, Y}
T 500	{C, O, O, K, I, E}

Find all frequent itemsets using Apriori and FP-growth, respectively. Compare the efficiency of the two mining processes. 14

5. (*a*) Briefly outline the major steps of decision tree classification. 7

(*b*) What is web content mining ? How is it different from web structure mining ? 7

6. (*a*) Briefly describe and give examples of each of the following approaches to clustering: partitioning methods, hierarchical methods, density-based methods, and grid-based methods. 7

(*b*) Suppose that the data mining task is to cluster points (with (x, y) representing location) into three clusters, where the points are

$A_1(2, 10)$, $A_2(2, 5)$, $A_3(8, 4)$, $B_1(5, 8)$, $B_2(7, 5)$, $B_3(6, 4)$, $C_1(1, 2)$, $C_2(4, 9)$.

The distance function is Euclidean distance. Suppose initially we assign A_1, B_1, and C_1 as the center of each cluster, respectively. Use the *k*-means algorithm to show only the three cluster centers after the first round of execution. 7

2018

Full Marks : 70

Time : 3 hours

Answer any five questions.

The figures in the right-hand margin indicate marks.

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

1. (*a*) Explain in detail the Network Security Model with the help of a neat and labelled diagram. 7

   (*b*) Sam wants to send the message 'HOW ARE YOU' to John. Use Substitution Cipher to encrypt and decrypt the message. 7

2. (*a*) Write down the steps involved in the Diffie-Hellman Key Exchange Algorithm. 7

   (*b*) Mathew wants to send the message 'CALL ME AT FOUR' to Paul. Show how a Transposition Cipher can be used to secure the message. 7

3. (*a*) Explain Public Key cryptography with the help of a suitable diagram. What are its pros and cons ? 7

(*b*) What purpose does a Digital Signature serve ? Explain the steps invovled in the creation of a Digital Signature. 7

4. (*a*) How does Pretty Good Privacy work ? What are the basic services that are provided by PGP ? 7

(*b*) What is Kerberos ? Describe the Kerberos authentication process in detail. 7

5. (*a*) What do you mean by IP Security ? Briefly discuss the IP security architecture. 7

(*b*) What does X.509 Certificate mean ? With the help of a diagram discuss the format of standard X.509 certificate. 7

6. (*a*) What is Secure Sockets Layer ? Explain the steps involved in an SSL handshake. 7

(*Continued*)

(*b*) What are the basic requirements in a Secure Electronic Transaction ? Who are the participants in a SET ? Give an example to support your answer. 7

7. (*a*) What is the importance of Passwords in providing security ? What are the basic things that should be kept in mind while creating a Password ? 7

(*b*) Explain Network based Intrusion Detection System in detail. 7

8. Write short notes on the following :

(*a*) Macro Virus 7

(*b*) Packet Filtering Firewalls. 7

———————

UL (7)-LPA

2018

Full Marks : 70

Time : 3 hours

Answer any **five** questions.

The figures in the right-hand margin indicate marks.

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

1. (*a*) How a file can be described in UNIX environment? List and explain about the various types of files in UNIX. 7

   (*b*) Give the reason why Linux commands are divided into internal and external commands. 7

2. (*a*) Differentiate between fork() and vfork(). 7

   (*b*) Write the syntax of six versions of exec functions and also explain how these functions differ from each other. 7

3. (*a*) Explain the arithmetic operators in shell. Write a shell program to find the factorial of a given number. 7

(*Turn Over*)

(*b*) Explain the number, string, and file comparison operators of shell in Linux. 7

4. (*a*) Explain similarities and dissimilarities between the semaphore and shared memory IPC mechanisms. 7

(*b*) Define grep command. Write a grep command to display the lines which does not matches the entire given pattern. 7

5. (*a*) Explain the synchronization of threads by using mutexes. 7

(*b*) Discuss in detail about the POSIX thread API. 7

6. (*a*) Explain the APIs associated for creating and destroying a shared memory with example. 7

(*b*) What are the drawbacks of System V IPC mechanisms ? 7

7. (*a*) What is zombie process ? Explain how zombie process can be removed from a system. 7

(*b*) Explain how pipes are used as a standard input and output. 7

(*Continued*)

8. (*a*) What is firewall ? Explain the different types of firewall in brief. 7

(*b*) Explain the following system calls : 7

(*i*) fcntl

(*ii*) lseek

(*iii*) read

(*iv*) write

———————

7. (*a*) Describe the architecture of the fuzzy logic systems. 7

(*b*) Explain inductive and deductive learning model with suitable example. 7

8. Write short notes on the following: 5 + 5 + 4

(*a*) Associative networks

(*b*) Simulated annealing

(*c*) LISP

VI Sem. CS+IT

2018

Full Marks : 70

Time : 3 hours

Answer any five questions.

The figures in the right-hand margin indicate marks.

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

1. (*a*) What is AI ? What are the different approaches in defining artificial intelligence ? 7

(*b*) Discuss the state space for the following problem:

You are given three jugs, a 12-gallon, 8-gallon, 3-gallon. None of the jugs have any measuring markers on it. You can use pump to fill jugs with water. How can you get exactly one gallon of water in any of the jugs ? Properly formulate the production rule to get the answer. 7

2. Consider the following sentences:

Anyone whom Mary loves is a football star.

Any student who does not pass does not play. John is a student. Any student who does not study does not pass. Anyone who does not play is not a football star. (Conclusion) If John does not study, then Mary does not love John.

(*a*) Translate the following sentences into first order predicate logic. 7

(*b*) Prove the conclusion using resolution. 7

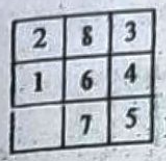
3. (*a*) Write a program in PROLOG to find out the GCD of 'n' numbers. 7

(*b*) Discuss the following search technique with the help of an example. Also discuss the benefits and shortcoming of each.

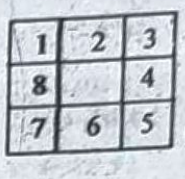
(*i*) Breadth First Search.

(*ii*) Depth First Search. 7

4. (*a*) What do you mean by the term 'Heuristic Function'? What heuristics can be used in case of 8-puzzle problem ? Calculate these heuristics values for the following initial and goal state. 7



Initial State



Goal State

(*b*) Explain the difference between Brute-force, heuristic and local search strategies. 7

5. (*a*) Describe Bayesian inference method using suitable example. 7

(*b*) Suppose an automobile insurance company classifies a driver as good, average, or bad. Of all their insured drivers, 25% are classified as good, 50% are average, and 25% are bad. Suppose for the coming year, a good driver has a 5% chance of having an accident, and average driver has 15% chance of having an accident, and a bad driver has a 25% chance. If you had an accident in the past year what is the probability that you are a good driver ? 7

6. (*a*) Describe the main components of expert system. 7

(*b*) Briefly explain the working of artificial neural network. 7