Council for Technical Education and Vocational Training
Office of the Controller of Examinations
Sanothimi, Bhaktapur
Regular/Back Exam – 2081/2082 Chaitra/Baishakh

Program: **Diploma in Computer Engineering**   **Full Marks: 80**
Year/Part: **III/I (2022)**   © Arjun   **Pass Marks: 32**
Subject: **Data Mining and Data Warehousing**   **Time: 3 hrs.**

Candidates are required to give their answers in their own words as far as practicable. The figures in the margin indicate full marks.

🌐 **www.arjun00.com.np**

Attempt ALL questions.

1. What is data mining? Explain data mining system architecture with necessary diagram. **[2+6]**

2. Explain the preprocessing steps in detail. **[8]**

3. What are the characteristics of data warehouse? Differentiate between database and data warehouse. **[4+4]**

4. Explain about different OLAP operation. **[8]**

5. Find the frequent item sets and generate association rules on the following transaction using Apriori Algorithm. **[8]**

Transaction ID	Items
T1	I1, I2, I3
T2	I1, I2
T3	I1, I4, I5
T4	I5, I4
T5	I5, I3
T6	I1, I4, I5

Assume: Minimum support threshold = 33.33%
Minimum confident threshold = 60%

6. What is classification? Explain Bayesian classification with suitable example. **[2+6]**

7. Define clustering. Explain about agglomerative and divisive clustering. **[2+6]**

8. Compare classification with clustering. Describe text mining in details. **[3+5]**

Cont. ......

9. Why data mining is important? What are different aspects of security and privacy in data mining? [4+4]

10. Write short notes on: (any **TWO**) [2×4]
    a. OLAP servers
    b. Association rules
    c. K-means clustering
    d. Support and confidence

Good Luck !

Council for Technical Education and Vocational Training
Office of the Controller of Examinations
Sanothimi, Bhaktapur
Regular/Back/Scholarship Exam - 2081/2082 Chaitra/Baishakh
Program: **Diploma in Information Technology**   **Full Marks: 80**
Year/Part: **III/I (2022)** © A r j u n   **Pass Marks: 32**
Subject: **Data Mining (Elective I)**   **Time: 3 hrs.**

Candidates are required to give their answers in their own words as far as practicable. The figures in the margin indicate full marks.

🌐 **www.arjun00.com.np**

Attempt any **EIGHT** questions.

1. What is data mining? Explain the importance of data mining. [4+6]

2. What is data preprocessing? Explain different preprocessing steps in brief. [2+8]

3. Define OLTP. Explain the architecture of Relational ROLAP. [2+8]

4. What is data warehousing? Explain ELT process. [4+6]

5. A dataset has five transactions. Find the all frequent itemset and association rule using apriori algorithm (min support = 3 and min confidence = 70%). [10]

Transaction ID	Data Items
T_1	A, B, C, D
T_2	A, B, D
T_2	B, D
T_4	A, C, D
T_5	A, B, C

6. Find the cluster from the given datasets using k-means algorithm (k = 2). [10]

ID	X	Y
1	1	7
2	4	5
3	3	9
4	6	8
5	1	5

7. Explain the Bayesian method. Differentiate between supervised learning algorithm and unsupervised algorithm. [4+6]

8. How Support Vector Machine (SVM) is used to classify the different datasets? [10]

9. Write short notes on (any **TWO**) [2×5]
   a) Histogram analysis
   b) Cross validation
   c) Text mining
   d) Perceptron

Good Luck !

Council for Technical Education and Vocational Training
Office of the Controller of Examinations
Sanothimi, Bhaktapur
Regular Exam-2081 Jestha/Ashadh

Program:	Diploma in Computer Engineering	Full Marks: 80
Year/Part:	III/I (2022) © A r j u n	Pass Marks: 32
Subject:	Data Mining and Data Warehousing	Time: 3 hrs.

Candidates are required to give their answers in their own words as far as practicable. The figures in the margin indicate full marks.

www.arjun00.com.np

Attempt any FIVE questions.

1. a. What is knowledge discovery process (KDD)? Explain data mining system architecture. [2+6]

   b. What are data objects? Explain steps in data preprocessing. [2+6]

2. a. Differentiate between data warehouse and database Write characteristics of data warehouse. [5+3]

   b. What is data warehouse schemas? Explain snowflake and fact constellation schema. [2+6]

3. a. What is text mining? Explain the aspects of security and privacy in data mining. [2+6]

   b. Perform k-medoid algorithm. (K=2) [8]

i	x	y
x_1	2	6
x_2	3	4
x_3	3	8
x_4	4	7
x_5	6	2
x_6	6	4
x_7	7	3
x_8	7	4
x_9	8	5
x_{10}	7	6

4. a. Explain roll-up, drill down, slice and pivot operations in detail. [8]

www.arjun00.com.np

Cont. ......

b. What do you mean by market basket analysis? Explain apriori algorithm with example [2+6]

5. a. What is linear regression? Differentiate between classification and clustering. [4+4]

   b. Explain decision tree induction classification with example. [8]

6. Write short notes on: (any **FOUR**) [4×4]

   a. Support and confidence
   b. Divisive method
   c. Web mining
   d. K-means method
   e. ROLAP

Good Luck !

Council for Technical Education and Vocational Training
Office of the Controller of Examinations
Sanothimi, Bhaktapur
Regular/Back Exam- 2081 Jestha/Ashadh

Program:	Diploma in Information Technology	Full Marks: 80
Year/Part:	III/I (2022) © A r j u n	Pass Marks: 32
Subject:	Data Mining (Elective I)	Time: 3 hrs.

Candidates are required to give their answers in their own words as far as practicable. The figures in the margin indicate full marks.

🌐 **www.arjun00.com.np**

Attempt any **EIGHT** questions.

1. Define the term data mining. Explain the knowledge discovery process with diagram. [2+8]

2. Define data object. Explain the data preprocessing. [2+8]

3. Explain the OLAP operations. Differentiate between OLAP and OLTP. [5+5]

4. Differentiate between ETL and ELT. Write about market basket analysis. [6+4]

5. A dataset has five transactions. Find all frequent item set and association rule using apriori algorithm (min-support = 3 and confidence = 80%). [10]

TID	Data Items
1	A, C, D
2	B, C, E
3	A, B, C, E
4	B, E

6. Define artificial neural network (ANN). Explain back propagation in neural network in brief. [2+8]

7. Find the cluster from the given datasets using k-means algorithm (k = 2). [10]

Id	X	Y
1	4	10
2	5	8
3	2	4
4	5	9
5	4	2
6	10	12

8. Define SVM. Explain the privacy and security issues in data [10]
   mining.

9. Write short notes on: (any **TWO**)   [2×5]
   a. Data cleaning
   b. Tree pruning
   c. Visual data mining
   d. Data warehousing

   Good Luck !

Council for Technical Education and Vocational Training
Office of the Controller of Examinations
Sanothimi, Bhaktapur
Regular/Back Exam – 2080 Magh/Phagun

Program:	**Diploma in Computer Engineering**	**Full Marks: 80**
Year/Part:	**III/II (2018)**	**Pass Marks: 32**
Subject:	**Data Mining**	**Time: 3 hrs.**

© Arjun

Candidates are required to give their answers in their own words as far as practicable. The figures in the margin indicate full marks.

www.arjun00.com.np

Attempt any **FIVE** questions.

1. a. Define data mining with its application. [8]

    b. Draw block diagram of data mining system architecture and explain. [8]

2. a. Differentiate between database and data warehouse. [8]

    b. What is data warehouse schemas? Discuss schema model in data warehouse. [3+5]

3. a. What is OLAP? Explain OLAP operations. [2+6]

    b. Differentiate between OLAP and OLTP. [8]

4. a. What is data preprocessing? Discuss its major steps. [3+5]

    b. What is clustering and classification? Differentiate between them. [4+4]

5. a. Explain k-means algorithm with suitable example. [4+4]

    b. Discuss about information retrieval and deep learning in detail. [4+4]

6. Write short notes on: (any **FOUR**) [4×4]

    a. Data cube

    b. DMQL

    c. Frequency patterns

    d. Apriori algorithm

    e. Concept of entropy

Good Luck !

Council for Technical Education and Vocational Training
Office of the Controller of Examinations
Sanothimi, Bhaktapur
Regular/Back Exam-2080, Bhadra

Program:	Diploma in IT Engineering	Full Marks: 80
Year/Part:	III/I (2016) © A r j u n	Pass Marks: 32
Subject:	Data Mining (Elective I)	Time: 3 hrs.

Candidates are required to give their answers in their own words as far as practicable. The figures in the margin indicate full marks.

Attempt Any Eight questions. 🌐 **www.arjun00.com.np**

1. Explain the concept of data mining. Define database and highlight the main characteristics of database approach. [4+2+4]

2. What is data warehouse? explain partitioning in data warehouse physical design. [2+8]

3. Define data mart. Explain star schema in detail. [2+8]

4. Explain extraction, transformation and load (ETL) process. [10]

5. What do you mean by OLAP? Describe OLAP architecture. [2+8]

6. Explain decision tree. Describe perception. [5+5]

7. Explain Hierarchical clustering. [10]

8. What do you mean by text mining and web mining? Explain. [10]

9. What are data mining standards? Write down privacy and security issues in data mining applications. [4+6]

10. Write short notes on: (**Any Two**) [2x5=10]

   a) SQL

   b) Neural network

   c) Limitations of traditional processing

Good Luck!

Council for Technical Education and Vocational Training
Office of the Controller of Examinations
Sanothimi, Bhaktapur
Regular/Back Exam-2079 Chaitra/2080 Baishakh

Program: **Diploma in Computer Engineering**   **Full Marks: 80**
Year/Part: **III/II (2018)** © A r j u n   **Pass Marks: 32**
Subject: **Data Mining**   **Time: 3 hrs.**

Candidates are required to give their answers in their own words as far as practicable. The figures in the margin indicate full marks.

🌐 **www.arjun00.com.np**

Attempt any **FIVE** questions.

1. a. What is data warehouse? Explain data mining system architecture. [2+6]

   b. Mention data mining functionality, classification, prediction, clustering and evolution. [8]

2. a. How does a snowflake schema differ from a star schema? List any two advantages and disadvantages of snowflake schema. [5+3]

   b. Explain concept of time series data and analysis. [8]

3. a. Explain term data cleaning, data integration and data transformation. [8]

   b. What do you mean by DMQL? Explain DMQL with syntax and proper example. [2+6]

4. a. What is agglomerative clustering? Explain concept of divisive clustering. [4+4]

   b. Define decision tree. Explain entropy and information gain in detail. [2+6]

5. a. Cluster the following instances of given data with the help of K-mean algorithm (K=2). [8]

Instances	X	Y
1	1.0	2.5
2	1.0	4.5
3	2.5	3.0
4	2.0	1.5
5	4.5	1.5
6	4.0	5.0

   b. Explain FP-growth algorithm with the properties. [8]

6. Write short notes on: [4×4]
   a. Advanced Data Mining
   b. DBMS Vs Data Warehouse
   c. Application of Data Mining
   d. KDD

Good Luck !

Council for Technical Education and Vocational Training
Office of the Controller of Examinations
Sanothimi, Bhaktapur
Regular/Back Exam-2078/2079, Chaitra/Baishakh

Program: Diploma in Computer Engineering　　　**Full Marks: 80**

Year/ Part: III/II (2018 New)　　　**Pass Marks: 32**

Subject:　Data Mining　　© Arjun　　　**Time: 3 hrs.**

Candidates are required to give their answers in their own words as far as practicable. The figures in the margin indicate full marks.

🌐 www.arjun00.com.np

Attempt Any Five Questions.

1. a) Define data mining with its functionalities. Mention important application of data mining. [5+3]

   b) What is fact constellation schemas? Explain star schema with its advantages and disadvantages. [8]

2. a) Describe data warehouse. Differentiate data warehouse with DBMS. [3+5]

   b) What are the important characteristics of OLTP Differentiate OLAP with OLTP [3+5]

3. a) What is clustering? Explain linear and Non-linear regression. [8]

   b) Define decision tree. Explain entropy and in formation gain in detail. [8]

4. a) What are the drawbacks of K-mean algorithm? Explain Agglomerative clustering in brief. [8]

   b) Explain DMQL with its syntax and example. [8]

5. a) What do you mean by slice and dice, drill up and drill down in multidimensional data? [8]

   b) Explain advance data mining with its important features. [8]

6. Write short notes on : **(Any Four)** [4x4=16]

   i) K-me doid algorithm

   ii) Deep Learning

   iii) Fp-growth Algorithm

   iv) MOLAP

   v) DMQL

Good Luck !

Council for Technical Education and Vocational Training
Office of the Controller of Examinations
Sanothimi, Bhaktapur
Regular/Back Exam-2077, Chaitra

Program: Diploma in Computer/IT Engineering	Full Marks: 80	
Year/Part: III/II (2013, 2008) © Arjun	Pass Marks: 32	
Subject: Data Mining & Warehousing (Elective -II)	Time: 3 hrs	

Candidates are required to give their answers in their own words as far as practicable. The figures in the margin indicate full marks.

www.arjun00.com.np

Attempt Any EIGHT questions.

1. Explain data mining. Describe the importance of data mining for social network. [4+6]

2. Write short notes on data warehousing. Describe the Architecture of data warehouse. [4+6]

3. Explain the star schemas, fact tables, dimension tables and dimension hierarchies of data warehouse with example. [10]

4. Differentiate between OLAP and OLTP. Explain the drill up and slice operation in data warehouse. [5+5]

5. Explain the partitioning, parallelism and compression in data warehouse. [10]

6. Define data warehouse construction process. Explain the ETL process in data warehouse. [5+5]

7. Explain the Neural network in data mining. [10]

8. Explain the cluster analyses. Write down the steps to cluster the given data using k-mean algorithm. [5+5]

9. Explain text mining and visual data mining. [5+5]

10. Write short notes on : (Any TWO) [2×5=10]
    a) Data mart
    b) Privacy and security issues in data mining
    c) Data manipulation language
    d) Man square deviation

Good Luck !

Council for Technical Education and Vocational Training
Office of the Controller of Examinations
Sanothimi, Bhaktapur
Regular/Back Exam-2076, Shrawan/Bhadra

Program: Diploma in Computer/IT Engineering **Full Marks: 80**
Year/Part: III /II (2010, 2008) © Arjun **Pass Marks: 32**
Subject: Elective II: Data Mining & Data Warehouse *Time: 3 hrs*

Candidates are required to give th...
practicable. The figures in the ma...

🌐 **www.arjun00.com.np**

Attempt **Any Eight** questions.

1.	Explain about Data Manipulation Language with examples.	[10]
2.	What do you mean by data warehouse? Explain star schemas, fact tables and dimension tables.	[4+6]
3.	Explain partitioning and compression in data warehouse.	[10]
4.	Define neural network. Explain perceptron and back propagation.	[2+8]
5.	Explain entropy and information gain in decision tree. What is the purpose of information gain?	[8+2]
6.	Explain K means and hierarchical clustering.	[10]
7.	Why do you need data mining? Explain any two techniques for data mining.	[3+7]
8.	Explain data cleaning, data integration and data transformation in data preprocessing.	[4+6]
9.	Explain about knowledge Discovery in Database with suitable diagram & examples.	[10]
10.	Write Short Notes on. **(Any Two)**	[2x5=10]

  a) Data Mart
  b) OLAP and OLTP
  c) Text mining
  d) Structured Query Language (SQL)

Good Luck!