

GUJARAT TECHNOLOGICAL UNIVERSITY

DIPLOMA ENGINEERING – SEMESTER –6 (NEW)- EXAMINATION –WINTER-2022

Subject Code: 3361604**Date: 15-12-2022****Subject Name: Data Mining And Warehousing****Time: 02:30 PM TO 05:00 PM****Total Marks: 70****Instructions:**

1. Attempt all questions.
2. Make Suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.
4. Use of simple calculators and non-programmable scientific calculators are permitted.
1. English version is authentic.

Q.1 Answer any seven out of ten. દરમાંથી કોઈપણ સાતના જવાબ આપો. 14

1. What is a data mining?
2. List any four applications of data mining.
3. Define continuous attributes in data mining.
4. List major tasks involved in data preprocessing.
5. Write full form of ARFF and WEKA.
6. What is clustering in data mining?
7. What is redundancy in data integration?
8. What is classification in data mining?
9. What is data mart?
10. What is WEKA?

Q.2 (a) What are the different techniques of data mining? Explain in brief. 03
અનુ. 2 (અ) ડેટા માર્ફિનિંગની વિવિધ તકનીકો શું છે? ટૂંકમાં સમજવો. 03

OR

- (a) Briefly explain types of data in data mining. 03
- (અ) ડેટા માર્ફિનિંગમાં ડેટાના પ્રકારોને સંક્ષિપ્તમાં સમજવો. 03
- (b) What are the major applications of data mining? Explain in brief. 03
- (અ) ડેટા માર્ફિનિંગના મુખ્ય એપ્લિકેશનો શું છે? ટૂંકમાં સમજવો. 03

OR

- (b) Briefly explain challenges of data mining. 03
- (અ) ડેટા માર્ફિનિંગના પડકારોને સંક્ષિપ્તમાં સમજવો. 03

- (c) What is standard deviation and variance in statistical calculations? Explain with example. **04**
- (ક) આંકડાકીય ગણતરીઓમાં standard deviation અને variance શું છે? ઉદાહરણ સાથે સમજાવો. **04**

OR

- (c) What is range and quartile in statistical calculations? Explain with example. **04**
- (ક) આંકડાકીય ગણતરીમાં range અને quartile શું છે? ઉદાહરણ સાથે સમજાવો. **04**
- (d) Explain nominal and binary attributes with example. **04**
- (દ) nominal અને binary એટ્રિબ્યુટ્સ ઉદાહરણ સાથે સમજાવો. **04**

OR

- (d) Explain ordinal and numeric attributes with example. **04**
- (દ) ordinal અને numeric એટ્રિબ્યુટ્સ ઉદાહરણ સાથે સમજાવો. **04**

- Q.3** (a) What is data cleaning and why is it important? **03**
- પ્રશ્ન. (અ) ડેટા ક્લિનિંગ શું છે અને તે શા માટે મહત્વપૂર્ણ છે? **03**
- 3

OR

- (a) What is correlation analysis? How is correlation analysis used to detect redundancies? **03**
- (અ) correlation analysis શું છે? રિનડન્સી શોધવા માટે correlation analysis નો ઉપયોગ કેવી રીતે થાય છે? **03**
- (b) How do we perform data visualization in WEKA? Explain in brief. **03**
- (અ) આપણે WEKA માં ડેટા વિજ્યુલાઈઝન કેવી રીતે કરી શકીએ? ટૂંકમાં સમજાવો. **03**

OR

- (b) How do we perform data filtering in WEKA? Explain in brief. **03**
- (અ) આપણે WEKA માં ડેટા ફિલ્ટરિંગ કેવી રીતે કરી શકીએ? ટૂંકમાં સમજાવો. **03**
- (c) Explain missing values and noisy data in data cleaning. **04**
- (ક) ડેટા ક્લિનિંગમાં missing values અને noisy data સમજાવો. **04**

OR

- (c) Explain tuple duplication and identification problem in data integration. **04**
- (ક) ડેટા એકીકરણમાં ટપલ દુપ્લિકેશન અને ઓળખની સમસ્યા સમજાવો. **04**
- (d) Explain the concept of probability-based solving. **04**
- (દ) probability-based solving નો concept સમજાવો. **04**

OR

- (d) What is clustering? Explain its types. **04**
- (દ) ક્લાસ્ટરિંગ શું છે? તેના પ્રકારો સમજાવો. **04**

- Q.4** (a) What is a data warehouse? What are the benefits of a data warehouse? **03**
- પ્રશ્ન. ૪ (અ) ડેટા વેરહાઉસ શું છે? ડેટા વેરહાઉસના ફાયદા શું છે? **03**

OR

- (a) Describe any three differences between operational database systems and data warehouse. **03**
- (અ) ઓપરેશનલ ડેટાબેઝ સિસ્ટમ્સ અને ડેટા વેરહાઉસ વર્ચેના કોઈપણ ત્રણ તકાવતોનું વર્ણન કરો. **03**
- (b) Explain enterprise data warehouse model. **04**
- (અ) એન્ટરપ્રાઇઝ ડેટા વેરહાઉસ મોડલ સમજાવો. **04**

OR

- (b) Explain virtual data warehouse model. **04**

(ભ)	વર્ણુઅલ ડેટા વેરહાઉસ મોડલ સમજવો.	૦૪
(ચ)	What is decision tree in data mining? Explain ID3 decision tree algorithm with example.	૦૭
(ક)	ડેટા માઈનિંગમાં decision tree શું છે? ઉદાહરણ સાથે ID3 decision tree અધ્યોરિધમ સમજવો.	૦૭
Q.5	(ા) How does the data warehousing and data mining work together?	૦૪
પ્રશ્ન. ૫	(અ) ડેટા વેરહાઉસિંગ અને ડેટા માઈનિંગ એકસાથે કેવી રીતે કાર્ય કરે છે?	૦૪
	(બ) What is the difference between OLAP and OLTP?	૦૪
	(ઝ) OLAP અને OLTP વચ્ચે શું તક્ષાવત છે?	૦૪
	(ચ) Define following terms. 1. Mean 2. Median 3. Mode	૦૩
	(ક) નીચેના શરૂઆતી વ્યાખ્યાપિત કરો. 1. Mean 2. Median 3. Mode	૦૩
	(દ) What is ARFF file in WEKA? Can we use csv file in WEKA?	૦૩
	(એ) WEKA માં ARFF ફાઈલ શું છે? શું આપણે WEKA માં csv ફાઈલનો ઉપયોગ કરી શકીએ?	૦૩
