

**GUJARAT TECHNOLOGICAL UNIVERSITY**  
DIPLOMA ENGINEERING – SEMESTER – 6 (NEW) - EXAMINATION – SUMMER-2022

**Subject Code:3361604****Date :04-06-2022****Subject Name: Data Mining And Warehousing****Time:10:30 AM TO 01:00 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 programmable & Communication aids are strictly prohibited.
5. Use of non-programmable scientific calculator is permitted.
6. English version is authentic.

**Q.1**

Answer any seven out of ten. દશમાંથી કોઈપણ સાતના જવાબ આપો.

**14**

1. What is data mining?  
૧. ડેટા માઈનિંગ શું છે?
2. How is a data warehouse different from a database?  
૨. ડેટા વેરહાઉસ ડેટાબેઝથી કેવી રીતે અલગ છે?
3. Define transactional data.  
૩. transactional data વ્યાખ્યાયિત કરો.
4. Define nominal attributes.  
૪. nominal attributes વ્યાખ્યાયિત કરો.
5. Define mean and mode.  
૫. mean અને mode વ્યાખ્યાયિત કરો.
6. Why do we need to preprocess data in data mining?  
૬. શા માટે આપણે ડેટા માઈનિંગમાં ડેટાને પ્રીપ્રોસેસ કરવાની જરૂર છે?
7. What is noisy data in data mining?  
૭. ડેટા માઈનિંગમાં noisy ડેટા શું છે?
8. Write full form of KDD and OLAP.  
૮. KDD અને OLAP નું પૂર્ણ સ્વરૂપ લખો.
9. What is data mart in data warehouse?  
૯. ડેટા વેરહાઉસમાં ડેટા માર્ટ શું છે?
10. Why do we use WEKA tool?  
૧૦. શા માટે આપણે WEKA ટૂલનો ઉપયોગ કરીએ છીએ?

**Q.2****પ્રશ્ન. ૨**

- (a) Describe three challenges to data mining.  
(અ) ડેટા માઈનિંગ માટેના ત્રણ પડકારોનું વર્ણન કરો.

**03****૦૩****OR**

- (a) List any six applications of data mining.  
(અ) ડેટા માઈનિંગની કોઈપણ છ એપ્લિકેશનોની યાદી બનાવો.
- (b) Explain numeric and binary attributes with example.  
(બ) numeric અને binary attributes ઉદાહરણ સાથે સમજાવો.

**03****૦૩****03****૦૩****OR**

- (b) Explain discrete and continuous attributes with example.

**03**

- (બ) discrete અને continuous attributes ઉદાહરણ સાથે સમજાવો. ૦૩
- (c) What are the steps involved in data mining process? Explain with figure. 04
- (ક) ડેટા માઈનિંગ પ્રક્રિયામાં કયા steps સામેલ છે? આકૃતિ સાથે સમજાવો. ૦૪

OR

- (c) What are the types of data in data mining? Explain in brief. 04
- (ક) ડેટા માઈનિંગમાં ડેટાના પ્રકારો શું છે? ટૂંકમાં સમજાવો. ૦૪
- (d) What is data preprocessing and why it is important? 04
- (ડ) ડેટા પ્રીપ્રોસેસિંગ શું છે અને તે શા માટે મહત્વપૂર્ણ છે? ૦૪

OR

- (d) Briefly explain major tasks involved in data preprocessing. 04
- (ડ) ડેટા પ્રીપ્રોસેસિંગમાં સામેલ મુખ્ય કાર્યોને સંક્ષિપ્તમાં સમજાવો. ૦૪

**Q.3**  
**પ્રશ્ન. 3**

- (a) List and describe methods for handling missing values in data cleaning. 03
- (અ) ડેટા ક્લિનિંગમાં missing values ને હેન્ડલ કરવા માટેની પદ્ધતિઓની સૂચિ બનાવો અને તેનું વર્ણન કરો. ૦૩

OR

- (a) What is noise? How to handle noisy data in data mining? 03
- (અ) noise શું છે? ડેટા માઈનિંગમાં noisy ડેટાને કેવી રીતે હેન્ડલ કરવું? ૦૩
- (b) What is data integration? Where is data integration used? 03
- (બ) ડેટા એકીકરણ શું છે? ડેટા એકીકરણનો ઉપયોગ ક્યાં થાય છે? ૦૩

OR

- (b) Briefly describe issues to consider during data integration in data mining. 03
- (બ) ડેટા માઈનિંગમાં ડેટા એકીકરણ દરમિયાન ધ્યાનમાં લેવાના મુદ્દાઓનું સંક્ષિપ્તમાં વર્ણન કરો. ૦૩
- (c) Explain quartiles and interquartile range as data dispersion measures. 04
- (ક) ડેટા dispersion ના માપદંડો તરીકે quartiles અને interquartile range સમજાવો. ૦૪

OR

- (c) Explain variance and standard deviation as data dispersion measures. 04
- (ક) ડેટા dispersion ના માપદંડો તરીકે variance અને standard deviation સમજાવો. ૦૪
- (d) What is redundancy and correlation analysis? Explain in brief. 04
- (ડ) redundancy અને correlation analysis શું છે? ટૂંકમાં સમજાવો. ૦૪

OR

- (d) Write a brief note on data value conflict detection and resolution. 04
- (ડ) ડેટા વેલ્યુ કન્ફ્લિક્ટ ડિટેક્શન અને રિઝોલ્યુશન પર સંક્ષિપ્ત નોંધ લખો. ૦૪

**Q.4**  
**પ્રશ્ન. 4**

- (a) What is classification? How does it work? 03
- (અ) વર્ગીકરણ શું છે? તે કેવી રીતે કામ કરે છે? ૦૩

OR

- (a) What is decision tree? How are decision trees used for classification? 03
- (અ) decision tree શું છે? વર્ગીકરણ માટે decision tree નો ઉપયોગ કેવી રીતે થાય છે? ૦૩
- (b) Differentiate between operational database systems and data warehouses. 04
- (બ) ઓપરેશનલ ડેટાબેઝ સિસ્ટમ્સ અને ડેટા વેરહાઉસ વચ્ચે તફાવત કરો. ૦૪

OR

- (b) What is data warehouse? Explain its characteristics. 04
- (બ) ડેટા વેરહાઉસ શું છે? તેના લક્ષણો સમજાવો. ૦૪
- (c) Explain following data warehouse models: 07
1. Enterprise Warehouse 2. Data Mart 3. Virtual Warehouse.

	(ક) નીચેના ડેટા વેરહાઉસ મોડેલો સમજાવો	૦૭
	1. Enterprise Warehouse 2. Data Mart 3. Virtual Warehouse.	
<b>Q.5</b>	(a) Explain the concept of cluster analysis.	<b>04</b>
<b>પ્રશ્ન. ૫</b>	(અ) કલસ્ટર analysis નો concept સમજાવો.	૦૪
	(b) Describe data visualization in WEKA.	<b>04</b>
	(બ) WEKA માં ડેટા વિઝ્યુલાઈઝેશનનું વર્ણન કરો.	૦૪
	(c) What are the features of WEKA?	<b>03</b>
	(ક) WEKA ની વિશેષતાઓ શું છે?	૦૩
	(d) Describe data filtering in WEKA.	<b>03</b>
	(ડ) WEKA માં ડેટા ફિલ્ટરિંગનું વર્ણન કરો.	૦૩

\*\*\*\*\*