

GUJARAT TECHNOLOGICAL UNIVERSITY

Diploma Engineering – SEMESTER – 6 (OLD) – EXAMINATION – Summer-2024

Subject Code: 3361604

Date: 18-05-2024

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 કોઇપણ સાતના જવાબ આપો. **14**
1. What is a data mining?
૧. ડેટા માઇનિંગ શું છે?
 2. List any four applications of data mining.
૨. ડેટા માઇનિંગની કોઇપણ ચાર એપ્લિકેશનની યાદી બનાવો.
 3. Define continuous attributes in data mining.
૩. ડેટા માઇનિંગમાં સતત વિશેષતાઓને વ્યાખ્યાયિત કરો
 4. List major tasks involved in data pre-processing.
૪. ડેટા પ્રી-પ્રોસેસિંગમાં સામેલ મુખ્ય કાર્યોની યાદી બનાવો
 5. Write full form of ARFF and WEKA.
૫. ARFF અને WEKA નું સંપૂર્ણ સ્વરૂપ લખો.
 6. What is clustering in data mining?
૬. ડેટા માઇનિંગમાં ક્લસ્ટરિંગ શું છે?
 7. What is redundancy in data integration?
૭. ડેટા એકીકરણમાં રીડન્ડન્સી શું છે?
 8. What are classification in data mining?
૮. ડેટા માઇનિંગમાં વર્ગીકરણ શું છે?
 9. What is data mart?
૯. ડેટા માર્ટ શું છે?
 10. What is WEKA?
૧૦. WEKA શું છે?
- Q.2** (a) What are the major applications of data mining? Explain in brief **03**
પ્રશ્ન. ૨ (અ) ડેટા માઇનિંગના મુખ્ય એપ્લિકેશનો શું છે? સંક્ષિપ્તમાં સમજાવો **૦૩**
- OR
- (a) Briefly explain challenges of data mining. **03**
(અ) ડેટા માઇનિંગના પડકારોને સંક્ષિપ્તમાં સમજાવો. **૦૩**
- (b) Briefly explain types of data in data mining. **03**
(બ) ડેટા માઇનિંગમાં ડેટાના પ્રકારોને સંક્ષિપ્તમાં સમજાવો. **૦૩**
- OR
- (b) What are the different techniques of data mining? Explain in brief. **03**

- (બ) ડેટા માઇનિંગની વિવિધ તકનીકો શું છે? સંક્ષિપ્તમાં સમજાવો. ૦૩
- (c) Explain nominal and binary attributes with examples. ૦૪
- (ક) nominal અને binary એટ્રિબ્યુટ્સ ઉદાહરણ સાથે સમજાવો. ૦૪
- OR
- (c) Explain ordinal and numeric attributes with examples. ૦૪
- (ક) ક્રમબદ્ધ અને સંખ્યાત્મક લક્ષણોને ઉદાહરણો સાથે સમજાવો. ૦૪
- (d) What is standard deviation and variance in statistical calculations? Explain with example. ૦૪
- (S) આંકડાકીય ગણતરીઓમાં standard deviation અને variance શું છે? ઉદાહરણ સાથે સમજાવો. ૦૪
- OR
- (d) What is range and quartile in statistical calculations? Explain with an example. ૦૪
- (S) આંકડાકીય ગણતરીમાં શ્રેણી અને ચતુર્થાંશ શું છે? ઉદાહરણ સાથે સમજાવો. ૦૪
- Q.3** (a) How do we perform data visualization in WEKA? Explain in brief. ૦૩
- પ્રશ્ન. ૩** (અ) આપણે WEKA માં ડેટા વિઝ્યુલાઇઝેશન કેવી રીતે કરી શકીએ? ટૂંકમાં સમજાવો. ૦૩
- OR
- (a) How do we perform data filtering in WEKA? Explain in brief. ૦૩
- (અ) આપણે WEKA માં ડેટા ફિલ્ટરિંગ કેવી રીતે કરી શકીએ? ટૂંકમાં સમજાવો. ૦૩
- (b) What is data cleaning and why is it important? ૦૩
- (બ) ડેટા સફાઈ શું છે અને તે શા માટે મહત્વપૂર્ણ છે? ૦૩
- OR
- (b) What is correlation analysis? How is correlation analysis used to detect redundancies? ૦૩
- (બ) correlation analysis શું છે? ટ્રાન્ડેન્સી શોધવા માટે correlation analysis નો ઉપયોગ કેવી રીતે થાય છે? ૦૩
- (c) Explain the concept of probability-based solving. ૦૪
- (ક) સંભાવના-આધારિત ઉકેલની વિભાવના સમજાવો. ૦૪
- OR
- (c) What is clustering? Explain its types. ૦૪
- (ક) ક્લસ્ટરિંગ શું છે? તેના પ્રકારો સમજાવો. ૦૪
- (d) Explain missing values and noisy data in data cleaning. ૦૪
- (S) ડેટા ક્લિનિંગમાં ગુમ થયેલ મૂલ્યો અને ઘોંઘાટીયા ડેટા સમજાવો. ૦૪
- OR
- (d) Explain tuple duplication and identification problems in data integration. ૦૪
- (S) ડેટા એકીકરણમાં ટ્યુપલ ડુપ્લિકેશન અને ઓળખ સમસ્યાઓ સમજાવો. ૦૪
- Q.4** (a) What is a data warehouse? What are the benefits of a data warehouse? ૦૩
- પ્રશ્ન. ૪** (અ) ડેટા વેરહાઉસ શું છે? ડેટા વેરહાઉસના ફાયદા શું છે? ૦૩
- OR
- (a) Describe any three differences between operational database systems and data warehouse. ૦૩
- (અ) ઓપરેશનલ ડેટાબેઝ સિસ્ટમ્સ અને ડેટા વેરહાઉસ વચ્ચેના કોઈપણ ત્રણ તફાવતોનું વર્ણન કરો. ૦૩
- (b) Explain the virtual data warehouse model. ૦૪
- (બ) વર્ચ્યુઅલ ડેટા વેરહાઉસ મોડલ સમજાવો. ૦૪

OR

- (b) Explain the enterprise data warehouse model. **04**
- (બ) એન્ટરપ્રાઇઝ ડેટા વેરહાઉસ મોડલ સમજાવો. **૦૪**
- (c) What is a decision tree in data mining? Explain the ID3 decision tree algorithm with an example. **07**
- (ક) ડેટા માઇનિંગમાં decision tree શું છે? ઉદાહરણ સાથે ID3 decision tree અલ્ગોરિધમ સમજાવો. **૦૭**
- Q.5** (a) How does data warehousing and data mining work together? **04**
- પ્રશ્ન. ૫** (અ) ડેટા વેરહાઉસિંગ અને ડેટા માઇનિંગ એકસાથે કેવી રીતે કાર્ય કરે છે? **૦૪**
- (b) What is the difference between OLAP and OLTP? **04**
- (બ) OLAP અને OLTP વચ્ચે શું તફાવત છે? **૦૪**
- (c) Define the following terms. **03**
1. Mean 2. Median 3. Mode
- (ક) નીચેના શબ્દો વ્યાખ્યાયિત કરો. **૦૩**
1. Mean 2. Median 3. Mode
- (d) What is ARFF file in WEKA? Can we use csv file in WEKA? **03**
- (ડ) WEKA માં ARFF ફાઇલ શું છે? શું આપણે WEKA માં csv ફાઇલનો ઉપયોગ કરી શકીએ? **૦૩**
