

**GUJARAT TECHNOLOGICAL UNIVERSITY**  
Diploma Engineering – SEMESTER – 6 (OLD) – EXAMINATION – Winter-2024

**Subject Code: 3361604**

**Date: 25-11-2024**

**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
5. English version is authentic.

- Q.1** Answer any seven out of ten. દશમાંથી કોઇપણ સાતના જવાબ આપો. **14**
1. Define Noisy Data in terms of Data Mining  
૧. ડેટા માઇનિંગ ના સંદર્ભ માં Noisy ડેટા વ્યાખ્યાન કરો.
  2. Define Nominal Attributes  
૨. Nominal Attributes વ્યાખ્યાન કરો.
  3. Write down Features of WEKA tool.  
૩. WEKA ટૂલ ની નવશેષત ઓ લખો.
  4. Write full form of KDD and OLTP.  
૪. KDD અને OLTP નાં પૂર્ણ સ્વરૂપ લખો.
  5. Define Numeric attributes.  
૫. Numeric attributes વ્યાખ્યાન કરો.
  6. What is clustering in data mining?  
૬. ડેટા માઇનિંગ માં clustering શું છે?
  7. Write the differences between data warehousing and data mining  
૭. ડેટા વેરહાઉસિંગ અને ડેટા માઇનિંગ ના તફતો લખો.
  8. What is WEKA?  
૮. WEKA શું છે?
  9. Explain median with example.  
૯. median ઉદાહરણ આપી સમજાવો.
  10. What is ARFF file in WEKA? Can we use csv file in WEKA?  
૧૦. WEKA માં ARFF ફાઇલ શું છે? શું આપણે WEKA માં csv ફાઇલનો ઉપયોગ કરી શકીએ?
- Q.2** (a) Describe three challenges to data mining. **03**  
પ્રશ્ન. ૨ (અ) ડેટા માઇનિંગ માટેના ૩ પડકારોનું વર્ણન કરો. **૦૩**
- OR
- (a) What are the different techniques of data mining? Explain in brief. **03**  
(અ) ડેટા માઇનિંગની વિવિધ તકનીકો શું છે? ટૂંકમાં સમજાવો. **૦૩**
  - (b) What is standard deviation and variance in statistical calculations? Explain with example. **03**  
(બ) આંકડાકીય ગણતરીઓમાં પ્રમાણભૂત વિચલન અને વિચલન શું છે? સમજાવો ઉદાહરણ સાથે. **૦૩**

OR

- (b) List any six applications of data mining. 03  
(બ) ડેટા માઇનિંગની કોઈપણ 6 એપ્લિકેશન ની બનવો. ૦૩
- (c) Explain discrete and continues attributes with example. 04  
(ક) discrete અને continues attributes ઉદાહરણ આપી સમજાવો. ૦૪

OR

- (c) What is range and quartile in statistical calculations? Explain with example. 04  
(ક) Statistical ગણતરીઓમાં range અને quartile શું છે? ઉદાહરણ સાથે સમજાવો. ૦૪
- (d) Explain ordinal and numeric attributes with example. 04  
(ડ) ordinal અને numeric એટ્રીબ્યુટસ સમજાવો. ૦૪

OR

- (d) What is data preprocessing and why it is important? 04  
(ડ) ડેટા પ્રીપ્રોસેસિંગ શું છે અને તે શા માટે મહત્વપૂર્ણ છે? ૦૪

**Q.3**  
**પ્રશ્ન. ૩**

- (a) Explain data filtering in WEKA. 03  
(અ) WEKAમાં ડેટા ફિલ્ટરીંગ સમજાવો. ૦૩

OR

- (a) Differentiate classification and clustering. 03  
(અ) Classification અને clustering વચ્ચે તફાવત આપો. ૦૩
- (b) What is data cleaning and why is it important? 03  
(બ) Data cleaning શું છે અને તે શા માટે મહત્વપૂર્ણ છે? ૦૩

OR

- (b) Explain redundancy and correlation analysis in brief. 03  
(બ) redundancy અને correlation analysis ટૂંકમાં સમજાવો. ૦૩
- (c) What is noise? How to handle noisy data in data mining? 04  
(ક) Noise શું છે? ડેટા માઇનિંગ માં noisy ડેટાને કેવી રીતે હેન્ડલ કરવું? ૦૪

OR

- (c) Explain advantages of Data warehousing. 04  
(ક) ડેટા વેરહાઉસિંગ ના લાભો સમજાવો. ૦૪
- (d) Explain data value conflict detection and resolution. 04  
(ડ) ડેટા વેલ્યુ કોન્ફ્લિક્ટ ડિટેક્શન અને રિસોલ્યુશન સમજાવો. ૦૪

OR

- (d) Describe data visualization in WEKA. 04  
(ડ) Weka માં ડેટા વિઝ્યુલાઇઝેશન વર્ણવો. ૦૪

**Q.4**  
**પ્રશ્ન. ૪**

- (a) Explain Enterprise data warehouse models. 03  
(અ) Enterprise ડેટા વેરહાઉસ મોડલ સમજાવો. ૦૩

OR

- (a) Explain in detail Data Transformation methods in detail. 03  
(અ) Data Transformation ને ડીટેલ માં સમજાવો. ૦૩
- (b) Describe about Data discretization? 04  
(બ) ડેટા ડિસ્ક્રિટાઇઝેશન વિશે વર્ણન કરો? ૦૪

OR

- (b) Compare and contrast OLTP and OLAP 04  
(બ) OLTP અને OLAP ની તુલના કરો અને વિપરીત કરો ૦૪
- (c) Explain the Roll up operation, Drill Down operation, and Slice operation in OLAP 07

	(ક) OLAP માં રોલ અપ ઓપરેશન, ડ્રિલ ડાઉન ઓપરેશન અને સ્લાઈસ ઓપરેશન સમજાવો	૦૭
<b>Q.5</b>	(a) Explain the concept of cluster analysis.	04
<b>પ્રશ્ન. ૫</b>	(અ) ક્લસ્ટર analysis ની concept સમજાવો.	૦૪
	(b) Explain Tuple Duplication with example.	04
	(બ) Tuple Duplication ઉદાહરણ સાથે સમજાવો.	૦૪
	(c) What is decision tree? How are decision tree used for classification?	03
	(ક) decision tree શું છે? વર્ગીકરણ માટે decision tree ઉપયોગ કેવી રીતે થાય છે?	૦૩
	(d) Define following terms and give examples of each 1. Mean 2. Mode	03
	(ડ) નીચેના શબ્દો વ્યાખ્યાયિત કરો અને દરેકના ઉદાહરણો આપો 1. Mean 2. Mode	૦૩

\*\*\*\*\*