

GUJARAT TECHNOLOGICAL UNIVERSITY

Diploma Engineering – SEMESTER – 6 (NEW) – EXAMINATION – Winter-2023

Subject Code: 3361604
Date: 06-12-2023
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 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. Define: Data Mining
૧. ડેટા માઇનિંગ વ્યાખ્યાયિત કરો.
 2. Write the differences between data warehousing and data mining
૨. ડેટા વેરહાઉસિંગ અને ડેટા માઇનિંગ ના તફાવત લખો.
 3. Write the full form of OLAP and OLTP
૩. OLAP અને OLTP નું પૂર્ણ સ્વરૂપ લખો.
 4. What is data mart in data warehouse?
૪. ડેટા વેરહાઉસ માં ડેટા માર્ટ શું છે?
 5. Define Noisy Data in terms of Data Mining
૫. ડેટા માઇનિંગ માં Noisy ડેટા શું છે?
 6. List the major tasks involved in data preprocessing
૬. ડેટા પ્રીપ્રોસેસિંગમાં સામેલ મુખ્ય કાર્યોની યાદી બનાવો.
 7. Define Nominal Attributes
૭. Nominal Attributes વ્યાખ્યાયિત કરો.
 8. Define Continuous Attributes in Data Mining
૮. ડેટા માઇનિંગ માં Continuous Attributes વ્યાખ્યાયિત કરો.
 9. Explain in short density-based clustering
૯. Density-based ક્લસ્ટેરિંગ ને ટૂંકમાં સમજાવો.
 10. Explain in brief Redundancy in Data Integration
૧૦. ડેટા ઇન્ટિગ્રેશન માં રીડન્ડન્સી સમજાવો.
- Q.2** (a) Describe the usage of Statistics in Data Mining **03**
પ્રશ્ન. ૨ (અ) ડેટા માઇનિંગ માં statistics ના ઉપયોગ વર્ણવો **૦૩**
- OR
- (a) Describe Mean and Mode with Example **03**
(અ) Mean અને Mode ઉદાહરણ આપી વર્ણવો. **૦૩**
- (b) Describe three challenges to data mining. **03**
(બ) ડેટા માઇનિંગ માટેના ૩ પડકારોનું વર્ણન કરો. **૦૩**
- OR
- (b) What is data preprocessing and why it is important? **03**

	(બ) ડેટા પ્રીપ્રોસેસિંગ શું છે અને તે શા માટે મહત્વપૂર્ણ છે?	૦૩
	(c) Explain ordinal and numeric attributes with example.	04
	(ક) Ordinal અને numeric attributes ઉદાહરણ સાથે સમજાવો.	૦૪
	OR	
	(c) What are the different techniques of data mining? Explain in brief.	04
	(ક) ડેટા માઇનિંગની વિવિધ તકનીકો શું છે? ટૂંકમાં સમજાવો.	૦૪
	(d) What are the major applications of data mining? Explain in brief.	04
	(ડ) ડેટા માઇનિંગના મુખ્ય એપ્લિકેશનો શું છે? ટૂંકમાં સમજાવો.	૦૪
	OR	
	(d) What is range and quartile in statistical calculations? Explain with example.	04
	(ડ) Statistical ગણતરીઓમાં range અને quartile શું છે? ઉદાહરણ સાથે સમજાવો.	૦૪
Q.3	(a) What is noise? How to handle noisy data in data mining?	03
પ્રશ્ન. ૩	(અ) Noise શું છે? ડેટા માઇનિંગ માં noisy ડેટાને કેવી રીતે હેન્ડલ કરવું?	૦૩
	OR	
	(a) What is data cleaning and why is it important?	03
	(અ) Data cleaning શું છે અને તે શા માટે મહત્વપૂર્ણ છે?	૦૩
	(b) Differentiate classification and clustering.	03
	(બ) Classification અને clustering વચ્ચે તફાવત આપો.	૦૩
	OR	
	(b) Explain variance and standard deviation as data dispersion measures.	03
	(બ) ડેટા dispersion ની માપદંડ તરીકે variance અને standard deviation સમજાવો.	૦૩
	(c) Write a brief note on data value conflict detection and resolution.	04
	(ક) ડેટા વેલ્યુ કન્ફ્લિક્ટ ડેટેક્શન અને રિઝોલ્યુશન પર સંક્ષિપ્ત નોંધ લખો.	૦૪
	OR	
	(c) Explain advantages of Data warehousing.	04
	(ક) ડેટા વેરહાઉસિંગ ના લાભો સમજાવો.	૦૪
	(d) Explain data filtering in WEKA.	04
	(ડ) Weka માં ડેટા ફિલ્ટરિંગ સમજાવો.	૦૪
	OR	
	(d) Describe data visualization in WEKA.	04
	(ડ) Weka માં ડેટા વિઝ્યુલાઇઝેશન વર્ણવો.	૦૪
Q.4	(a) Discuss the limitations of decision tree.	03
પ્રશ્ન. ૪	(અ) Decision tree ની મર્યાદાઓ વર્ણવો.	૦૩
	OR	
	(a) What is classification? How does it work?	03
	(અ) વર્ગીકરણ શું છે? તે કેવી રીતે કામ કરે છે?	૦૩
	(b) Explain Enterprise data warehouse models.	04
	(બ) Enterprise ડેટા વેરહાઉસ મોડલ સમજાવો.	૦૪
	OR	
	(b) Explain Virtual data warehouse models.	04
	(બ) Virtual ડેટા વેરહાઉસ મોડલ સમજાવો.	૦૪
	(c) Explain in detail Data Cleaning, Data Integration, and Transformation	07
	(ક) Data Cleaning, Data Integration, અને Transformation ને ડીટેલ માં સમજાવો.	૦૭

Q.5	(a)	Explain difference between operational data base system and data warehouse.	04
પ્રશ્ન. ૫	(અ)	ઓપરેશનલ ડેટાબેઝ અને ડેટા વેરહાઉસ વચ્ચે તફાવત સમજાવો.	૦૪
	(b)	What is ARFF file in WEKA? Explain use of csv file in WEKA?	04
	(બ)	Weka માં ARFF ફાઇલ શું છે? Weka માં csv ફાઇલનો ઉપયોગ સમજાવો.	૦૪
	(c)	What is KDD? Explain about data mining as a step in the process of knowledge discovery.	03
	(ક)	KDD શું છે? Knowledge discovery ની પ્રક્રિયામાં એક પગલા તરીકે ડેટા માઇનિંગ વિશે સમજાવો.	૦૩
	(d)	Explain Tuple Duplication with example.	03
	(ડ)	Tuple Duplication ઉદાહરણ સાથે સમજાવો.	૦૩
