

Seat No.: \_\_\_\_\_

Enrolment No. \_\_\_\_\_

**GUJARAT TECHNOLOGICAL UNIVERSITY**  
**DIPLOMA ENGINEERING - D.VOC –SEMESTER- 6 - EXAMINATION –SUMMER-2022**

**Subject Code:1260202**

**Date:02-06-2022**

**Subject Name:Concepts of Data Mining**

**Time:10:30 AM TO 12:30 PM**

**Total Marks:50**

**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. English version is authentic.
6. Use of non-programmable scientific calculator is permitted.

- Q.1(a)** Define the following terms: **05**
- i. Data Mining
  - ii. ROLAP
  - iii. KDD
  - iv. Data Mart
  - v. Decision Tree

નીચેના શબ્દો વ્યાખ્યાયિત કરો:

- i ડેટા માઇનિંગ
- ii. ROLAP
- iii KDD
- iv ડેટા માર્ટ
- v. નિર્ણય વૃક્ષ

- Q.1(b)** Explain the need for developing Data Warehouse. **05**  
ડેટા વેરહાઉસ વિકસાવવાની જરૂરિયાત સમજાવો.

- Q.2(a)** Briefly explain the steps involved in Data Mining. **05**  
સંક્ષિપ્તમાં ડેટા માઇનિંગમાં સામેલ પગલાંઓ સમજાવો.

- Q.2(b)** Discuss star schema with example. **05**  
ઉદાહરણ સાથે સ્ટાર સ્કીમાની ચર્ચા કરો.

**OR**

- Q.2(b)** Discuss Snow flake schema with example. **05**  
ઉદાહરણ સાથે સ્નો ફ્લેક સ્કીમાની ચર્ચા કરો.

- Q.3(a)** Discuss any three Data transformation strategies. **05**  
કોઈપણ ત્રણ ડેટા ટ્રાન્સફોર્મેશન વ્યૂહરચનાઓની ચર્ચા કરો.

- Q.3(b)** Discuss the concept of Association rule mining. **05**  
એસોસિયેશન નિયમ માઇનિંગની વિભાવનાની ચર્ચા કરો.

**OR**

- Q.3(a) Discuss Data smoothing methods in detail.  
ડેટા સ્મૂથિંગ પદ્ધતિઓની વિગતવાર ચર્ચા કરો. **05**
- Q.3(b) Give guidelines for successful data mining.  
સફળ ડેટા માઇનિંગ માટે માર્ગદર્શિકા આપો. **05**
- Q.4(a) Give the steps to evaluate the performance of algorithms in context of Data Mining.  
ડેટા માઇનિંગના સંદર્ભમાં અલ્ગોરિથમ્સના પ્રદર્શનનું મૂલ્યાંકન કરવા માટેનાં પગલાં આપો. **05**
- Q.4(b) Explain the Naïve Bayes algorithm.  
નેવ બેયસ અલ્ગોરિથમ સમજાવો. **05**
- OR**
- Q.4(a) Write a brief account on Direct hashing and pruning.  
ડાયરેક્ટ હેશિંગ અને પ્રુનિંગ પર સંક્ષિપ્ત એકાઉન્ટ લખો. **05**
- Q.4(b) Explain FP-Growth algorithm giving example.  
ઉદાહરણ આપતા FP-ગ્રોથ અલ્ગોરિથમ સમજાવો. **05**
- Q.5(a) What is the need of Decision tree? Explain the split algorithm on the basis of Gini Index.  
નિર્ણય વૃક્ષની શું જરૂર છે? જીની ઇન્ડેક્સના આધારે વિભાજન અલ્ગોરિથમ સમજાવો. **05**
- Q.5(b) What do you understand by overfitting? How do you overcome it?  
ઓવરફિટિંગ દ્વારા તમે શું સમજો છો? તમે તેને કેવી રીતે દૂર કરશો? **05**
- OR**
- Q.5(a) Differentiate between classification and regression.  
વર્ગીકરણ અને રીગ્રેસન વચ્ચે તફાવત કરો. **05**
- Q.5(b) Explain the desired features of cluster analysis.  
ક્લસ્ટર વિશ્લેષણની ઇચ્છિત વિશેષતાઓ સમજાવો. **05**