

GUJARAT TECHNOLOGICAL UNIVERSITY
BE - SEMESTER-V (NEW) EXAMINATION – WINTER 2023

Subject Code:3151608**Date:05-12-2023****Subject Name:Data Science****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. Simple and non-programmable scientific calculators are allowed.

- Q.1** (a) Define continuous random variable and discrete random variable with example. **03**
 (b) Differentiate Stratified Sampling and Cluster Sampling. **04**
 (c) Explain the Framework for Data-Driven Decision Making process. **07**
- Q.2** (a) Define Following Terms: **03**
 1. Entropy 2. Information Gain 3. Population
 (b) Explain the measures of shape- Skewness and Kurtosis. **04**
 (c) Explain Probability Density Function (PDF) and Cumulative Distribution Function (CDF) of a Continuous Random Variable with suitable example. **07**
- OR**
- (c) Discuss the various measures used to validate the simple linear regression. **07**
- Q.3** (a) Discuss pyramid of analytics with diagram. **03**
 (b) Explain Central Limit Theorem. **04**
 (c) How the Chi-square distribution is differ from student's t-distribution explain with example. **07**
- OR**
- Q.3** (a) Define following terms: **03**
 1. Standard Error 2. Sample Mean 3. Degrees of Freedom
 (b) List Machine Learning Algorithms and explain any one in brief. **04**
 (c) Explain classification of various Sampling methods. **07**
- Q.4** (a) Give the difference between Probabilistic Sampling and Non-Probability Sampling. **03**
 (b) Explain why Central Limit Theorem is called as a heart of the Data Science. **04**
 (c) Explain Simple Linear Regression model with example. **07**
- OR**
- Q.4** (a) What is Weight and Bias Tradeoff in Linear Regression ? **03**
 (b) How do you calculate maximum likelihood estimation? **04**
 (c) Explain Decision tree algorithm with suitable example.
- Q.5** (a) What is Outlier Analysis explain in detail. **03**
 (b) Explain the term: Method of Moments. **04**
 (c) Explain Random forest algorithm with suitable example. **07**
- OR**
- Q.5** (a) Explain significance of GINI impurities in splitting dataset. **03**
 (b) Explain various Attribute Selection Measures. **04**
 (c) Explain in detail: Maximum Likelihood Estimation (MLE). **07**
