**Printed Pages:2** **Sub Code:KIT-601**

**Paper Id:** 236803 **Roll No.** | | | | | | | | | | |

**B. TECH**
**(SEM VI) THEORY EXAMINATION 2022-23**
**DATA ANALYTICS**

***Time: 3 Hours*** ***Total Marks: 100***

**Note:** Attempt all Sections. If require any missing data; then choose suitably.

**SECTION A**

**1. Attempt *all* questions in brief.** **2 x 10 = 20**

(a) What are the main characteristics of Big Data?
(b) Generalize the role of analytical tools in Big data?
(c) What are the purposes of regression analysis?
(d) What do you mean by fuzzy qualitative model?
(e) Define association rule.
(f) State the benefits of analytic sandbox.
(g) What do you mean by data stream management system?
(h) What do you mean by response modeling?
(i) What are the benefits of visual data exploration?
(j) Mention some main goals of Hadoop.

**SECTION B**

**2. Attempt any *three* of the following:** **10x3=30**

(a) Compare and contrast analysis and reporting in data analytics with suitable example.
(b) What is a neural network? How can it be used in analytics?
(c) Explain Apriori association rule mining algorithm.
(d) List the advantages and disadvantages of K-Means clustering.
(e) What is HDFS? How does it handle Big Data?

**SECTION C**

**3. Attempt any *one* part of the following:** **10x1=10**

(a) What are the various stages in big data analytics life cycle? Illustrate with a figure, explaining each of them.
(b) What is the difference between regression modelling and Bayesian modeling? Explain in brief.

**4. Attempt any *one* part of the following:** **10x1=10**

(a) Explain the role of principal component analysis in neural networks.
(b) What are the parameters used to characterize any fuzzy membership function?

**5. Attempt any *one* part of the following: 10x1=10**

(a) Discriminate the concept of sampling data in a stream.
(b) Illustrate various Real Time Analytics Platforms (RTAPs) with examples.

**6. Attempt any *one* part of the following: 10x1=10**

(a) Explain the working of CLIQUE algorithm in brief.
(b) Identify the major issues in data stream query processing.

**7. Attempt any *one* part of the following: 10x1=10**

(a) Illustrate and explain the concept of Map Reduce framework in brief.
(b) Write R function to check whether the given number is prime or not?