



Paper id: 250616

Roll No:

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

BTECH
(SEM VI) THEORY EXAMINATION 2024-25
DATA ANALYTICS

TIME: 3 HRS

M.MARKS: 70

Note: Attempt all Sections. In case of any missing data; choose suitably.

SECTION A

1. Attempt all questions in brief.

02 x 7 = 14

Q no.	Question
a.	Describe Big Data. List real world Big Data examples.
b.	Differentiate between Analysis and Reporting.
c.	Appraise the need of Outlier removal in the Regression Analysis.
d.	Explain multivariate analysis.
e.	Enumerate sampling data in a stream.
f.	Correlate CLIQUE with high dimensional clustering.
g.	Discuss some most popular tools for data visualization.

SECTION B

2. Attempt any three of the following:

07 x 3 = 21

a.	Illustrate various stages in Big Data lifecycle.
b.	Summarize the utility of Neural Network in Analytics.
c.	Briefly discuss the steps in query processing.
d.	Describe the process of mining frequent item sets from an FP-Tree.
e.	Compare and contrast the No SQL with RDBMS databases.

SECTION C

3. Attempt any one part of the following:

07 x 1 = 07

a.	Discuss the data preparation stage of the data analytics life cycle. Why is this stage often considered the most time-consuming?
b.	Explain the importance of data analytics in today's business environment? Discuss at least three key benefits organizations can derive from effectively leveraging data analytics.

4. Attempt any one part of the following:

07 x 1 = 07

a.	Argue the effectiveness of Fuzzy Logic in the presence of uncertainty in the data. Differentiate it with the crisp sets with an example.
b.	Determine how the PCA can help in improving the performance of SVM classifier. Do you agree with statement "The use of PCA on the given dataset can effectively reduce its dimension while preserving its key characteristics", defend your opinion.

5. Attempt any one part of the following:

07 x 1 = 07

a.	Simulate sentiment analysis application with real world example.
b.	Justify how the Real-time Analytics Platforms (RTAP) can aid in analytics and real time decision making.

6. Attempt any one part of the following:

07 x 1 = 07

a.	Illustrate the working of k-means clustering algorithm. How it is different than the K-Medoids method?
b.	Appraise the Apriori algorithm in terms of the associated advantages limitations.



Paper id: 250616

Printed Page: 2 of 2
Subject Code: BADS601

Roll No:

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

BTECH
(SEM VI) THEORY EXAMINATION 2024-25
DATA ANALYTICS

TIME: 3 HRS

M.MARKS: 70

7. Attempt any *one* part of the following:

07 x 1 = 07

a.	Explain Apache Hive emphasizing its application in querying and managing large datasets stored in Hadoop.
b.	Justify the use of R programming for data visualization. Illustrate data exporting through R programming language.

QP25EP1_290
| 02-Jun-2025 1:54:39 PM | 117.55.242.134