PAPER ID-410530

Subject Code: KCA034

Roll No: | | | | | | | | | | | | | |

MCA
(SEM IV) THEORY EXAMINATION 2023-24
DATA ANALYTICS

TIME: 3 HRS **M.MARKS: 100**

Note: 1. Attempt all Sections. If require any missing data; then choose suitably.

SECTION A

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

a.	Explain the characteristics of data.	02
b.	What is the need of data analytics?	02
c.	Explain support vector method in data analytics.	02
d.	Explain rule induction in data analytics.	02
e.	What is Real-time Analytics Platform (RTAP)?	02
f.	What is stream computing?	02
g.	Write any two challenges of clustering for data streams.	02
h.	Explain frequent pattern-based clustering.	02
i.	What is Pig?	02
j.	What is Hive?	02

SECTION B

2. Attempt any *three* of the following: **3 x 10 = 30**

a.	How the data analytics is important for various industries? Explain with an example.	10
b.	How does principal component analysis (PCA) work in neural networks.	10
c.	Explain the characteristics of data streams.	10
d.	How can frequent item sets be mined in a stream of data?	10
e.	What is NoSQL database and how it is different from tradition relational database?	10

SECTION C

3. Attempt any *one* part of the following: **1 x 10 = 10**

a.	Explain any two data analytics tools.	10
b.	Write a short note on: i. Autoencoder neural network ii. Convolutional neural network	10

4. Attempt any *one* part of the following: **1 x 10 = 10**

a.	What are the benefits and limitations of Filtering stream technique?	10
b.	What is parallel clustering and how does it work?	10

5. Attempt any *one* part of the following: **1 x 10 = 10**

a.	What is Hadoop Distributed File System (HDFS)?	10
b.	Explain the difference between analysis and reporting in data analytics.	10

6. Attempt any *one* part of the following: **1 x 10 = 10**

a.	How to rule induction algorithm work?	10
b.	Explain the key components of Real-time Analytics Platform (RTAP).	10

7. Attempt any *one* part of the following: **1 x 10 = 10**

a.	What is CLIQUE clustering and how it is used in high dimensional data analysis?	10
b.	Explain the major use case of Hive?	10