CBCS SCHEME

USN 18CS7

Seventh Semester B.E. Degree Examination, June/July 2023 Big Data and Analytics

Time: 3 hrs. Max. Marks: 100

Note: Answer any FIVE full questions, choosing ONE full question from each module.

Module-1

1 a. Define big data, and explain its characteristics.
b. List and explain different data sources.
c. Explain Big data designing architecture.
(06 Marks)
(10 Marks)

OR

a. Explair the functions of each of the Big Query layers in big data architecture design using Big Query Cloud Service at Google Cloud Platform. (10 Marks)
b. (i) I splain big data analytics applications. (05 Marks)
(ii) Write a short note on data storage and analysis. (05 Marks)

Module-2

- 3 a. With a reat diagram, explain the components of HDFS (Hadoop Distribution File System).
 - How does the Hadoop MapReduce data flow work for a word count program? Give an example.

OR

- 4 a. What i APACHE Flume? Describe the feature components and working of apache flume.
 - b. Explain the features and benefits of apache Hive in hadoop. (10 Marks)
 (10 Marks)

Module-3

5 a. Discus NOSQL data architecture pattern, with an example.
b. Explair four different ways for handling big data problems. (10 Marks)

OR

6 a. (i) Explain different components of Cassandra. (05 Marks)
(ii) Explain different data types built into Cassandra. (05 Marks)
b. (i) Describe different CQL commands and their functionalities. (05 Marks)
(ii) Write a short note on NOSQL to Manage Big Data. (05 Marks)

Module-4

a. Descrit the significance of apache pig in hadoop.
 b. With a near diagram, explain MapReduce Programming model. How does MapReduce enable query processing quickly in Big Data Problems?

OR

8 a Explai: Hive architecture with a neat diagram. (10 Marks)
b. (i) 1) Herentiate between Pig and MapReduce. (05 Marks)
(ii) Write a short note on Pig architecture design layers. (05 Marks)

and Note: 1. On completing your answers, computating draw diagonal cross lines on the remaining blank pages.

2. Any revealing of identification: appear to evaluator and for equations written eg. 42 - 8 - 50, will be treated as malpraetice

DOWNLOAD THIS FREE AT

www.vturesource.com

18CS72

Module-5

- a. What are outliers? Describe the reasons for the presence of outliers in a relationship.
 - b. How can a university student's GPA be predicted from his/her high school percentage (HSP) of marks? (Assume linear regression) Plot a graph for the same. (10 Marks)

OR

- 10 a. Discuss the different phases of text mining process.
 - b. Write a short note on text mining and web mining.
 - Discuss three phases for web usage mining.

(40 Marks

(05 Marks

2 of 2