



**Avinashilingam Institute for Home Science and Higher Education for Women**

(Deemed to be University under Category 'A' by MHRD, Estd. u/s 3 of UGC Act 1956)

Re-accredited with 'A+' Grade by NAAC. Recognised by UGC under Section 12 B

Coimbatore - 641 043, Tamil Nadu, India

**Bachelor of Vocation Degree Examination – June 2021  
IV Semester**

**Class : II B.Voc**

**Major : Artificial Intelligence and Machine Learning**

**Time: 3 Hrs**

**Max. Marks: 100**

**19VAIC11 Big Data Analytics  
Part A**

**10 x 1 = 10**

**Choose the correct answer**

1. What was Hadoop named after? CO1 K1
  - a. Doug Cutting's school rock band
  - b. Doug Cutting's son's toy elephant
  - c. Doug Cutting's favourite song
  - d. Doug Cutting's son's name
  
2. The process of organizing data into informational summaries is ----- CO1 K2
  - a. Streaming
  - b. Analysing
  - c. Visualization
  - d. Reporting
  
3. XML data is \_\_\_\_\_ in nature. CO2 K2
  - a. structured data
  - b. unstructured data
  - c. semi structured data
  - d. hybrid
  
4. \_\_\_\_\_ Schedules jobs and tracks the assigns jobs to Task tracker. CO2 K1
  - a. JobTracker
  - b. PayLoad
  - c. Task Tracker
  - d. Mapper
  
5. In how many stages the MapReduce program executes?. CO3 K1
  - a. 3
  - b. 4
  - c. 2
  - d. 5
  
6. In HDFS, Namenodes run on \_\_\_\_\_ nodes. CO3 K1
  - a. deamons
  - b. master
  - c. cluster
  - d. slave
  
7. The Data model in Pig is ----- CO4 K2
  - a. flat relational
  - b. column relational
  - c. nested relational
  - d. row relational
  
8. HBase is a distributed \_\_\_\_\_ database built on top of the Hadoop file system. CO4 K1
  - a. column wise
  - b. tuple
  - c. row wise
  - d. record
  
9. \_\_\_\_\_ are distributed significantly away from the rest of the data set. CO5 K1
  - a. Slope
  - b. Raw data
  - c. Noise data
  - d. Outliers
  
10. In regression analysis, the variable that is being predicted is----- CO5 K3
  - a. independent variable
  - b. both a and b
  - c. dependent variable
  - d. none

**Part B**  
**Answer the following**  
**Answer should not exceed 400 words or two pages**

**5x6 = 30**

- |  |     |    |
|--|-----|----|
| 11.a. Write a short note on Big Data and its evolution.<br>(or)                | CO1 | K1 |
| 11.b. Differentiate Analysis and Reporting in Big Data.                        | CO1 | K2 |
| 12.a. Define Mining Data Streams. Give two examples.<br>(or)                   | CO2 | K3 |
| 12.b. Write the Decaying Window Algorithm.                                     | CO2 | K1 |
| 13.a. List the salient features of HDFS.<br>(or)                               | CO3 | K2 |
| 13.b. Write a short note on Job Scheduling.                                    | CO3 | K1 |
| 14.a. State the fundamentals of Zookeeper.<br>(or)                             | CO4 | K1 |
| 14.b. Write the contribution of IBM in Big Data Framework.                     | CO4 | K2 |
| 15.a. List the statistical interpretations of Regression Coefficients.<br>(or) | CO5 | K1 |
| 15.b. Discuss the important aspects of Visualization techniques.               | CO5 | K3 |

**Part C**  
**Answer the following**  
**Answer should not exceed 800 words or four pages**

**5x12 = 60**

- |  |     |    |
|--|-----|----|
| 16.a. Explain the challenges of Conventional systems in handling Data.<br>(or)           | CO1 | K4 |
| 16.b. Discuss about Intelligent Data Analysis and its role.                              | CO1 | K1 |
| 17.a. Explain with a diagram, the Key Components of Streaming Data Architecture.<br>(or) | CO2 | K3 |
| 17.b. Give a detail note on Filtering techniques for Data Streams.                       | CO2 | K1 |
| 18.a. Write the role of Map Reduce Programs in Big Data Analytics.<br>(or)               | CO3 | K1 |
| 18.b. With a neat sketch , explain Hadoop Environment.                                   | CO3 | K2 |
| 19.a. Write about the objectives of Pig and Hive in Hadoop.<br>(or)                      | CO4 | K1 |
| 19.b. Enumerate on HBase in detail with its salient features.                            | CO4 | K1 |
| 20.a. Explain Regression techniques and its types.<br>(or)                               | CO5 | K4 |
| 20.b. Discuss about the applications of Big Data in various domains.                     | CO5 | K2 |

\*\*\*\*\*