|  |  |
| --- | --- |
| LOGO.jpg | **GEETHANJALI INSTITUTE OF SCIENCE & TECHNOLOGY**  (**AN AUTONOMOUS INSTITUTION**)  **(Approved by AICTE, New Delhi & Affiliated to JNTUA, Ananthapuramu)**  **(Accredited by NAAC with “A” Grade, NBA (EEE,ECE &ME) & ISO 9001:2008Certified Institution)** |
| **QUESTIONBANK(DESCRIPTIVE)**  **Subject Name with Code: Big Data Analytics (22A0534c)**  **Course & Branch: B.Tech CSE(DS) Year& Semester: III-II Regulation:RG22** | |

**UNIT - I**

|  |  |  |
| --- | --- | --- |
| **S.No.** | **Question** | **[BT Level] [CO][ Marks]** |
| **2 Marks Questions (Short)** | | |
| 1. | What is Big Data? | L1/CO1/2M |
| 2. | List the characteristics of the bigdata? | L1/CO1/2M |
| 3. | List the applications of the bigdata? | L1/CO1/2M |
| 4. | What is structured and unstructured data? | L1/CO1/2M |
| 5. | What is big data analytics? List the different types of big data analytics? | L1/CO1/2M |
| 6. | What is HDFS? | L1/CO1/2M |
| 7. | What is MapReduce? | L1/CO1/2M |
| 8. | List the components that collectively form a Hadoop ecosystem? | L2/CO1/2M |
| 9. | What are descriptive, prescriptive and predictive analytics? | L1/CO1/2M |
| 10. | List the Uses of Big Data? | L1/CO1/2M |
| **Descriptive Questions (Long)** | | |
| **1** | What is meant by Big Data? Why Big Data is important? Discuss in detail. | L2/CO1/12M |
| **2** | Explain Structuring Big Data? | L2/CO1/12M |
| **3** | Explain Big Data Analytics? | L2/CO1/12M |
| **4** | Explain i) Data ii) Data Storage and analysis | L2/CO1/12M |
| **5** | Elaborate the history of Apache Hadoop. | L2/CO1/12M |
| **6** | Draw and Explain about the Hadoop Ecosystem. | L2/CO1/12M |
| **7** | List and explain the steps in installation of Hadoop | L2/CO1/12M |
| **8** | How can you analyze the data with hadoop? Discuss with an example?. | L2/CO1/12M |
| **9** | Discuss in detail about Scaling out in Hadoop. | L2/CO1/12M |

**UNIT - II**

|  |  |  |
| --- | --- | --- |
| **S.No.** | **Question** | **[BT Level] [CO][ Marks]** |
| **2 Marks Questions (Short)** | | |
| **1** | What is HDFS? Where to use HDFS? | L1/CO2/2M |
| **2** | What is HDFS? Where not to use HDFS? | L1/CO2/2M |
| **3** | List and explain the concepts of HDFS? | L1/CO2/2M |
| **4.** | What is data node? | L1/CO2/2M |
| **5.** | What is name node? | L1/CO2/2M |
| **6.** | What is Yarn? | L1/CO2/2M |
| **7.** | What are the Differences between Mapreduce and yarn? | L1/C02/2M |
| **8.** | What is Secondary name node? | L1/CO2/2M |
| **Descriptive Questions (Long)** | | |
| **1** | Discuss about the concepts of HDFS. | L2/CO2/12M |

|  |  |  |
| --- | --- | --- |
| **2** | Draw and Explain Architecture of HDFS | L2/CO2/12M |
| **3** | Explain the concept of Command line interface. | L2/CO2/12M |
| **4** | Explain about the Anatomy of a file read and Anatomy of a file write | L2/CO2/12M |
| **5.** | How would you configure an API in a MapReduce application? | L3/CO2/12M |
| **6.** | How would you set up a development environment for MapReduce on your system? | L3/CO2/12M |
| **7.** | Can you demonstrate how to run a MapReduce job locally on test data? | L3/CO2/12M |
| **8.** | Explain the concept of running on a cluster in Map Reduce. | L2/CO2/12M |

**UNIT - III**

|  |  |  |
| --- | --- | --- |
| **S.No.** | **Question** | **[BT Level] [CO][ Marks]** |
| **2 Marks Questions (Short)** | | |
| 1. | Explain Anatomy of a mapreduce job? | L2/CO3/2M |
| 2. | What is the task failure? | L1/CO3/2M |
| 3. | What is the application master failure? | L1/CO3/2M |
| 4. | What is shuffle and sort? | L1/CO3/2M |
| 5. | What is Hadoop? | L1/CO3/2M |
| 6. | What is mapper and reducer? | L1/CO3/2M |
| 7. | List the hadoop installation steps? | L1/CO3/2M |
| 8. | List the configuration files in hadoop? | L1/CO3/2M |
| **Descriptive Questions (Long)** | | |
| **1** | Discuss about the Anatomy of a Map Reduce job run. | L2/CO3/12M |
| **2** | Explain in detail about Failures in Map Reduce. | L2/CO3/12M |
| **3** | How can the shuffle and sort performed in map side? Explain. | L2/CO3/12M |
| **4** | How can the shuffle and sort performed in reduce side? Explain. | L2/CO3/12M |
| **5** | Explain about setting up of the Hadoop cluster. | L2/CO3/12M |
| **6** | Discuss about Cluster specification in detail. | L2/CO3/12M |
| **7** | Explain about cluster setup and installation. | L2/CO3/12M |
| **8** | Explain about Hadoop configuration. | L2/CO3/12M |

**UNIT - IV**

|  |  |  |
| --- | --- | --- |
| **S.No.** | **Question** | **[BT Level] [CO][ Marks]** |
| **2 Marks Questions (Short)** | | |
| **1** | What is Apache pig? | L1/CO4/2M |
| **2** | Why do we need Apache pig? | L1/CO4/2M |
| **3** | List the features of Apache pig? | L1/CO4/2M |
| **4.** | What is Pig Latin? | L1/CO4/2M |
| **5.** | List the components of Apache pig? | L1/CO4/2M |
| **6.** | What are the Differences between pig and SQL? | L1/CO4/2M |
| **7.** | List the user-defined functions in pig? | L1/CO4/2M |
| **8.** | List the data processing operators in pig? | L1/CO4/2M |
| **9.** | Explain the filter operator in pig? | L2/CO4/2M |
| **10.** | Explain the joins in pig? | L2/CO4/2M |
| **Descriptive Questions (Long)** | | |

|  |  |  |
| --- | --- | --- |
| **1** | Discuss about the concepts of Pig. | L2/CO4/12M |
| **2** | Draw and Explain Architecture of Pig | L2/CO4/12M |
| **3** | How would you install and run the Pig tool on your local environment or a server? | L3/CO4/12M |
| **4** | Distinguish between the Pig with Databases | L2/CO4/12M |
| **5** | Explain about Pig Latin in brief. | L2/CO4/12M |
| **6** | Discuss user defined functions in Pig Latin. | L2/CO4/12M |
| **7** | Can you demonstrate the use of data processing operators in a practical data processing scenario? | L3/CO4/12M |

**UNIT - V**

|  |  |  |
| --- | --- | --- |
| **S.No.** | **Question** | **[BT Level] [CO][ Marks]** |
| **2 Marks Questions (Short)** | | |
| **1** | What is HIVE? | L1/CO5/2M |
| **2** | List the features of HIVE? | L1/CO5/2M |
| **3** | Explain hive clients? | L1/CO5/2M |
| **4.** | What are the differences between Hive vs RDBMS? | L2/CO5/2M |
| **5.** | What is HiveQL? | L1/CO5/2M |
| **6.** | What is spark? | L1/CO6/2M |
| **7.** | List the features of Apache spark? | L1/CO6/2M |
| **8.** | What is HBase and List the features of HBase? | L1/CO5/2M |
| **9.** | List the components of HBase? | L1/CO5/2M |
| **Descriptive Questions (Long)** | | |

|  |  |  |
| --- | --- | --- |
| **1** | What is Hive? Draw and Explain Architecture of Hive | L2/CO5/12M |
| **2** | Can you demonstrate the process of installing Hive and explain how to running it on a local machine? | L3/CO5/12M |
| **4** | Explain in detail about HiveQL | L2/CO5/12M |
| **5.** | How would you create and modify a table in HiveQL? Provide examples. | L3/CO5/12M |
| **6.** | How would you write a Hive query to retrieve specific data from a Hive | L3/CO5/12M |
| **7.** | What is HBase? Draw and Explain Architecture of HBase | L2/CO5/12M |
| **8.** | How would you apply the HBase installation procedure in a real-world project? | L3/CO5/12M |
| **9** | What is Spark? Draw and Explain Architecture of Spark | L2/CO6/12M |
| **10** | Discuss in detail about Anatomy of Job run in Spark | L2/CO6/12M |