

ANURAG Engineering College**(An Autonomous Institution)****IV B.Tech I Semester Supplementary Examinations, April – 2024****BIG DATA ANALYTICS****(COMPUTER SCIENCE AND ENGINEERING)****Time: 3 Hours****Max.Marks:75****Section – A (Short Answer Type Questions)****(25 Marks)****Answer ALL the questions.**

- | | |
|--|-------|
| 1. What is distributed file system? | [2 M] |
| 2. What are velocity, volume & variety in Big Data? | [3 M] |
| 3. What is Data Serialization. | [2 M] |
| 4. What is Hadoop Eco System? | [3 M] |
| 5. What is the use of PIG? | [2 M] |
| 6. Describe column oriented database. Give an example? | [3 M] |
| 7. What is the need of spark in Hadoop? | [2 M] |
| 8. What are Spark components? | [3 M] |
| 9. What are the features of MongoDB. | [2 M] |
| 10. How to create index in MongoDB | [3 M] |

Section – B (Essay Answer Type Questions)**Answer ALL the questions. All carry Equal marks.****5 X 10M = 50M**

11. A) Explain word count problem using Map Reduce Algorithm.

OR

B) What are the drivers of Big Data Analytics? Explain its Applications in Defense, Space, Agriculture, domains.

12. A) How to Read & Write data in Hadoop Eco System using HDFS?

OR

B) Explain Hadoop eco system with neat diagram of its components.

13. A) How to do Joins & Sub queries in Hive

OR

B) How to build applications using Hbase in Zookeeper.

14. A) Explain Machine Learning MLlib.

OR

B) Explain about any four NoSQL Databases with example.

15. A) Explain about document oriented MongoDB.

OR

B) Discuss all the keywords in MongoDB. Explain about join in MangoDB?