M.G.M's COLLEGE OF ENGINEERING, NANDED

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

ASSIGNMENT NO.: 1

SUBJECT: Big Data Analytics **CLASS:BE(CSE) - I & II**

1. Write a Short note on Big Data and Distributed File System
2. Define Big data and explain its importance.
3. Enlist and Explain Big Data Applications.
4. Explain the Four V's of Big data in brief.
5. Write a short note on Big data analytics
6. Explain MapReduce with Matrix vector multiplication example.
7. Explain Hadoop ecosystem and its components.
8. Explain moving data in and Out of Hadoop in detail.
9. Write a short note Input and outputs to mapreduce- Data serialization.
10. Draw and Explain Hadoop Architecture in detail.
11. Explain the different modes of Hadoop cluster Configuration (Local,Pseudo-distributed mod e, Fully Distributed mode.
12. Define HDFS and Explain The Design of HDFS
13. Explain HDFS Architecture in detail.
14. Compare HDFS with Traditional Database system.
15. Write a Short note on Blocks and Why Block in HDFS is So Large?
16. Explain the role of name node and data node in HDFS.
17. Write a short note on Block Caching, HDFS Federatiion, Hdfs High availability.
18. Describe Basic File System Operation in HDFS using CLI.
19. Enlist and explain the different Hadoop Filesystems.
20. Explain the Anatomy of a File Read.
21. Explain the Anatomy of a File Write HDFS.
22. Explain Replica Placement policy in HDFS.

Subject In-Charge

Mr. Juned Khan