Department of Computer Science & Engineering
(Accredited by NBA 2019-2022)
18CS72- BIG DATA ANALYTICS

Question Bank

(Module 1 & Module 2)

Sl.No	Question
1.	Write Short notes on the following with examples. a)Big Data, b) Structured data, c) Semi structured data , d) Unstructured data
2.	Explain the big data processing architecture with neat diagram and functionalities.
3.	Explain analytics scalability in detail
4.	Describe the factors affecting the data quality
5.	Explain the various data pre-processing techniques
6.	Explain data stores with structured and semi structured data.
7.	Describe Big data storage techniques
8.	Describe the different big data platforms.
9.	Explain Berkerly Data Analytics Stack with neat diagram
10.	Case studies on different domains of applications
11.	Explain the features of Hadoop
12.	Describe the Map Reduce frame work and programming model in Hadoop
13.	With a neat diagram explain YARN based Hadoop execution model
14.	Explain the various components of HDFS with a neat diagram
15.	Discuss the need for HDFS and also list the commands to perform the following actions i)List files on HDFS ; ii) Copy files to HDFS ;iii) Delete a file within HDFS iv) Make a directory in HDFS ;v) Copy files from HDFS ;vi) Delete a directory in HDFS
16.	Write short notes on HDFS Block Replication , HDFS safe mode.
17.	Explain the following Secondary Name Node ,Journal node , HDFS Snap Shot , HDFS name space federation
18.	List all essential Hadoop tools in Hadoop eco system and their functions with neat diagram .