# Insights of Big Data - MID_1

## Unit-1

1. Define Bigdata, explain the evolution of Big Data and their characteristics
2. Define Data, met data and Big Data, also explain structured data, semi structured data, and unstructured data
3. Explain the techniques used in Big Data Analytics.
4. Explain the meaning of Big Data Analytics and Highlight its characteristics.
5. Explain in detail about the challenges of conventional systems.
6. Discuss big data in Healthcare, Ecommerce and Medicine.

## Unit-2

1. Explain the following a. HDFS block replication b. Rack awareness c. Name Node High Availability.
2. What are the challenges of distributed Company and How does Hadoop overcome these challenges.
3. Discuss the features of Hadoop HDFS.
4. Explain the concept of Hadoop Distributed File System. How does it manage Data storage in Hadoop cluster.
5. Compare and Contrast, RDBMS, and HADOOP in terms of architecture and suitability for handling big data.
6. Illustrate App based execution model and its functions with neat diagram.

## Unit – 3

1. Discuss the map reduce paradigm? how does it aggregate intermediate results?

2.Explain the role of Mapper in MapReduce framework. How does it process Input data?

3. Describe how MapReduce can be used to sort a large dataset. What are the challenges and benefits of sorting in a distributed environment?

## Open Book:

1. What are the benefits of Big Data, Discuss the challenges under Big Data, How Big data analytics can be useful in the development of smart cities.
2. Elaborate on the importance of big data analytics in business success and innovation.
3. What are the key features of the HDFS that making suitable for handling large datasets.
4. Describe the main Functionalities of the resource manager and Node Manager in Yarn.
5. Demonstrate the Implementation of Map Reduce …
6. With a neat sketch explain the logical data flow of Map Reduce.