UNIT-1

* Introduction to big data, IBM 4’v,
* Hadoop ecosystem,
* Purpose of Apache Ambari in the HDP stack,
* Hadoop Distributed File System (HDFS),
* Function of the NameNode and DataNodes,
* Explaining how files are stored and blocks ("splits") are replicated,
* MapReduce model v1, limitations of Hadoop 1 and MapReduce 1,
* Mapper class and Reducer class,
* Describing the YARN model,
* Comparing Hadoop 2/YARN with Hadoop 1,
* Functions of the main components of Ambari

**UNIT–2**

* Architecture and components of the Spark unified stack,
* Understanding the nature and purpose of Apache Spark in the Hadoop ecosystem,