| Class No | Chapter Title/Reference Literature | Topics to be Covered | % of Portions Covered | |
|---|---|---|---|---|
| | | | Reference Chapter | Cumulative |
| 1 | Big Data Introduction/T1 | Big Data definition, Challenges and opportunities with Big Data | 1 | 3.6 |
| 2 | Big Data Characterisitics/T1 | Data intensive scientific discovery and the role of Big Data, History | 2 | 7.2 |
| 3 | HDFS/T1 | Map Reduce – Storage (HDFS) | 2 | 10.8 |
| 4 | Map Reduce/T1 | Map Reduce – Computation model, Map Reduce architecture, | 2,4 | 14.3 |
| 5 | Hands on – Map Reduce/T1 | Demo class: Map-Reduce – Hands on programming | 2,4 | 17.9 |
| 6 | YARN/T1 | Case Study: Google. YARN introduction. | 2 | 21.5 |
| 7 | Hadoop Ecosystem/T1 | Overview of Hadoop Ecosystem – Oozie, Ambari, Sqoop and Flume | 2 | 25 |
| 8 | Matrix Vector Multiplication/T1 | Introduction to sample Big Data Algorithms – Sparse Matrices, matrix vector multiplication with MR | 4 | 28.6 |
| 9 | Pagerank/T1 | Introduction to sample Big Data Algorithms - Pagerank computations | 9 | 32.2 |
| 10 | Relational Operators with MR/T1 | Relational operators on Map-reduce, Select, Project, Join, Grouping, HIVE | 4 | 35.8 |
| 11 | Hands On with HIVE/T1 | HIVE hands on | 4 | 39.3 |
| 12. | Hbase-Cassandra/T1 | case study: Other storage - Hbase/Cassandra architecture and columnar storage for analytics | 3 | 42.9 |
| 13 | Hadoop issues/T1 | Issues with Hadoop, Spark and Scala | 5 | 46.5 |
| 14 | PySpark/T1 | PySpark programming model | 5 | 50 |
| 15 | Spark Programming Model/T1 | Transformations and Actions, Spark SQL | 5 | 53.6 |

| | | | | |
|---|---|---|---|---|
| 16 | Spark Architecture/T1 | Spark architecture – RDD, DataFrames, Wide and Narrow dependencies, | 5 | 57.2 |
| 17 | Algorithm Complexity/T3 | Complexity of Big Data algorithms – Communication Cost complexity model. | 2 | 60.8 |
| 18 | Hands On with Spark/T1 | Spark HandsOn | 5 | 64.3 |
| 19 | Streaming Spark/T1 | Streaming analytics use cases, Streaming Spark, | 7 | 67.9 |
| 20 | Kafka/T1 | Kafka – use cases, architecture | 7 | 71.5 |
| 21 | Streaming Algorithms 1/T1 | Streaming Algorithms - Sampling, set membership | 7 | 75 |
| 22 | Kafka Hands on/T1 | Kafka with HandsOn | 7 | 78.6 |
| 23 | Streaming Algorithms 2/T1 | Streaming Algorithms - Bloom Filters, Counting<br>Counting unique elements – Flajolet Martin Algorithm. | 7 | 82.2 |
| 24 | ML Algorithms/T1 | Clustering Algorithms - kmeans and collaborative filtering | 6 | 85.8 |
| 25 | ML andBig Data/T1 | Scaling Neural Networks for Big Data, case study MLLib. | 6 | 89.3 |
| 26 | Project work | Project Work feedback | | 92.9 |
| 27 | Project work | Project Work feedback | | 96.5 |
| 28 | Project work | Project Evaluations | | 100 |