MANGALORE UNIVERSITY

DEPARTMENT OF COMPUTER SCIENCE

MASTER OF COMPUTER APPLICATIONS (MCA) PROGRAMME

MCAS504 : BIG DATA ANALYTICS		
Hours/Week: 4 **Credits : 4**		**I.A. Marks: 30** **Exams. Marks: 70**

Course Outcomes:

CO1: Work with big data platform and explore the big data analytics techniques business applications.
CO2: Design efficient algorithms for mining the data from large volumes.
CO3: Analyze the HADOOP and Map Reduce technologies associated with big data analytics.
CO4: Explore on Big Data applications Using Pig and Hive.
CO5: Understand the fundamentals of various big data analytics techniques.
CO6: Build a complete business data analytics solution

	UNIT-I	12 Hours

Introduction To Big Data: Introduction – distributed file system – Big Data and its importance, Four Vs, Drivers for Big data, Big data analytics, Big data applications. Algorithms using map reduce, Matrix-Vector Multiplication by Map Reduce.

	UNIT-II	12 Hours

Introduction Hadoop: Big Data – Apache Hadoop &Hadoop Eco System – Moving Data in and out of Hadoop – Understanding inputs and outputs of Map Reduce - Data Serialization. Hadoop Architecture: Hadoop Architecture, Hadoop Storage: HDFS, Common Hadoop Shell commands , Anatomy of File Write and Read., Name Node, Secondary Name Node, and Data Node, Hadoop Map Reduce paradigm, Map and Reduce tasks, Job, Task trackers - Cluster Setup – SSH &Hadoop Configuration – HDFS Administering –Monitoring & Maintenance.

	UNIT-III	12 Hours

Hadoop Ecosystem And Yarn: Hadoop ecosystem components - Schedulers - Fair and Capacity, Hadoop 2.0 New Features Name Node High Availability, HDFS Federation, MRv2, YARN, Running MRv1 in YARN.

	UNIT-IV	12 Hours

Hive And Hiveql, Hbase : Introduction to No Query Language, Hive Architecture and Installation, Comparison with Traditional Database, HiveQL - Querying Data - Sorting And Aggregating, Map Reduce Scripts, Joins & Sub queries, HBase concepts Advanced Usage, Schema Design, Advance Indexing - PIG, Zookeeper - how it helps in monitoring a cluster, Hbase uses Zookeeper and how to Build Applications with Zookeeper.

REFERENCE BOOKS

1. Boris Lublinsky, Kevin T. Smith, Alexey Yakubovich, “Professional Hadoop Solutions”, Wiley, 2015.
2. Chris Eaton, Dirk Derooset al. , “Understanding Big data ”, McGraw Hill, 2012.
3. Tom White, “HADOOP: The definitive Guide” , O Reilly, 2012.
4. VigneshPrajapati, “Big Data Analytics with R and Haoop”, Packet Publishing 2013.
5. Tom Plunkett, Brian Macdonald et al, “Oracle Big Data Handbook”, Oracle Press, 2014.
6. JyLiebowitz, “Big Data and Business analytics”,CRC Press, 2013.