CONTENTS

CHAPTER NO TITLES	PAGE NO
1 BIG DATA INTRODUCTION	**1-45**
1.1 Introduction	
1.2 Big Data Characteristics	
1.3 A Big Data Architecture	
1.3.1 Challenges in Designing Big Data Architecture	
1.3.2 Benefits of Big Data Architecture	
1.4 Types of Big Data	
1.5 Features of Big Data Analytics and Requirements	
1.6 Building Blocks of Hadoop	
1.7 Big Data Technology	
1.8 Big Data vs Business intelligence	
1.9 Benefits of Using Big Data	
1.10 Disadvantages of Big Data	
2 WORKING WITH BIG DATA	**46-74**
2.1 Hadoop Introduction	
2.2 Hadoop Framework	
2.3 Hadoop Architecture	
2.3.1 Work with Hadoop	
2.3.2 Work with Hadoop HDFS	
2.4.1 Architecture of HDFS	
2.4.2 HDFS Objectives	
2.5 Hadoop File Systems	
2.5.1 Google File System	
2.5.2 Hadoop Distributed File System	
2.5.3 HDFS-Master Node Nodes	
2.5.4 Scalability Comparison Between GFS and HDFS	
2.5.5 Data Flow Input and Output	
2.6 HDFS Operations in Hadoop	
2.6.1 Activating HDFS	
2.6.2 Listing HDFS Files	
2.6.3 Data Entry into HDFS	
2.6.4 Data Retrieval from HDFS	
2.6.5 The HDFS is Being Shut Down	
2.6.6 HDFS Advantages	
2.7 Hadoop Cluster	
2.7.1 Hadoop Cluster Basic Architecture	
2.7.2 Benefits of Hadoop Clusters	

3 MAPREDUCE **75-91**
3.1 Introduction
Writable Wrappers for Java Primitives
4.6 Raw Comparator Will Speed Up Your
Hadoop Map/Reduce (Mr)
4.7 Comparators That are Made to Order

5 Apache PIG **106-131**
5.1 Introduction
5.1.1 Apache Pig Applications
5.1.2 History of the Apache Pig
5.2 Importance of the Apache Pig
5.3 Pig Characteristics
5.4 Apache Pig Vs MapReduce
5.5 Apache Pig – Architecture
5.6 Pig Latin Data Model
5.7 The Pig Latin Application Flow in Hadoop
5.8 Working Through The Abcs Of Pig Latin
5.9 Pig Latin – Basics
5.10 Evaluating Local and Distributed
Modes of Running Pig Scripts
5.11 Pig Script Interfaces in Hadoop
5.12 Pig Script Execution | Apache Pig
Running Scripts and Comments

6 HIVE **132-176**
Introduction to the Hive
6.1.1 Mapreduce Operations Can be
Carried Out in A Variety of Ways
6.1.2 Important Hive Characteristics
6.1.3 Hive Can be Interacted with Using Methods
6.1.4 The Following are Some of the
Most Critical Elements of Hive
6.1.5 Relational databases vs Hive
6.1.6 Features of Hive
6.2 Hive Architecture
6.3 Hive Job Execution Flow
6.4 Different Modes of Hive
6.5 Workflow of Hive
6.6 Hive - Data Types
6.6.1 Types of Strings
6.7 Hive - Create Database
6.8 Hive - Drop Database

6.9 Hive - Create Table
6.10 Load Data Statement in Table
6.11 Hive - Alter Table
6.12 Replace Statement
Hive - Drop Table
6.14 Partitioning in Hive
6.15 Hive - Built-in Operators
6.15.1 Relational Operators
6.15.2 Arithmetic Operators
6.15.3 Logical Operators
6.15.4 Complex Operators
6.16 Hive - Built-in Functions
6.16.1 Aggregate Functions
6.17 Hive - View and Indexes
6.18 HIVEQL - Select-Where
6.18.1 HIVEQL - Select-Group By
6.19 Joins
6.19.1 Left Outer Join
6.19.2 Right Outer Join
6.19.3 Full Outer Join
6.20 Limitations of Hive
6.21 Apache Hive's Advantages
6.22 Query Language for Hive (HQL)

7 BDA APPLICATIONS **177-223**
7.1 Introduction
7.2 Big Data Applications in Healthcare
7.3 Big Data Analytics in Education
7.4 Big Data Analytics in Industries
7.5 Big Data Analytics in Social Media
7.6 Big Data Analytics in Cyber Security
7.7 Big Data Analytics in E-Commerce
Acronyms