

Big Data Analytics

Index

Chapter 1: Introduction to Big Data

- 1.1 Big Data - Concepts and Terminologies
- 1.2 Types of Big Data
- 1.3 Big Data Adoption and Planning Considerations
- 1.4 Big Data Business Architecture
- 1.5 Business Process Management Information
- 1.6 Big Data Analytics Lifecycle
- 1.7 Big Data Online Analytical Processing
- 1.8 Extract, Transform, Load (ETL)
- 1.9 Traditional BI
- 1.10 Big Data BI

Chapter 2: Hadoop and Pig Framework

- 2.1 Hadoop Architecture
- 2.2 Hadoop Distributed File System (HDFS)
- 2.3 YARN
- 2.4 Hadoop I/O
- 2.5 Developing a MapReduce Application
- 2.6 MapReduce Working Procedure
- 2.7 Types and Formats of MapReduce
- 2.8 Features of MapReduce: Sorting and Joins
- 2.9 Data Pipelining
- 2.10 Introduction to Pig

Chapter 3: Data Modelling and Optimization Problems

- 3.1 Exploratory Data Analysis
- 3.2 Data Analytics Methods
- 3.3 Analytics for Unstructured Data
- 3.4 HBase
- 3.5 NoSQL
- 3.6 Data Visualization Techniques

Chapter 4: Streaming Data Analysis

4.1 Text Mining

4.2 Events and Trends in Text Streams

4.3 Embedding Semantics in LDA Model

4.4 Social Media Analytics

4.5 Twitter API

4.6 Facebook API

4.7 YouTube API

4.8 Social Network Analysis

4.9 Real-time Analytics with Spark Streaming and Structured Streaming

4.10 Introducing Real-time Processing

4.11 Architecture of Spark Streaming

4.12 Structured Streaming

Chapter 5: Structured Streaming

5.1 Introduction

5.2 Hive QL - Data Definition and Data Manipulation

5.3 Hive QL Queries

5.4 Aggregate Functions

5.5 Overview of Spark