

BIG DATA ANALYTICS

CREDITS:	3
-----------------	----------

OBJECTIVE:

- To understand the computational approaches to big data analytics
- To understand the various search methods and visualization techniques
- To learn to use various techniques for mining data stream
- To understand the applications using Map Reduce Concepts.

UNIT I - INTRODUCTION TO BIG DATA

9

Introduction to Big Data Platform – Challenges of Conventional Systems - Intelligent data analysis – Nature of Data - Analytic Processes and Tools - Analysis vs Reporting.

UNIT II - MINING DATA STREAMS

9

Introduction To Streams Concepts – Stream Data Model and Architecture - Stream Computing - Sampling Data in a Stream – Filtering Streams – Counting Distinct Elements in a Stream – Estimating Moments – Counting Oneness in a Window – Decaying Window - Real time Analytics Platform(RTAP) Applications - Case Studies - Real Time Sentiment Analysis- Stock Market Predictions.

UNIT III - HADOOP

9

History of Hadoop- the Hadoop Distributed File System – Components of Hadoop Analysing the Data with Hadoop- Scaling Out- Hadoop Streaming- Design of HDFS- Java interfaces to HDFS Basics- Developing a Map Reduce Application-How Map Reduce Works-Anatomy of a Map Reduce Job run-Failures-Job Scheduling-Shuffle and Sort – Task execution - Map Reduce Types and Formats- Map Reduce FeaturesHadoop environment.

UNIT IV - FRAMEWORKS

9

Applications on Big Data Using Pig and Hive – Data processing operators in Pig – Hive services – HiveQL – Querying Data in Hive - fundamentals of HBase and ZooKeeper - IBM InfoSphereBigInsights and Streams.

UNIT V - VISUALIZATION TECHNIQUES

9

Predictive Analytics- Simple linear regression- Multiple linear regression- Interpretation of regression coefficients. Visualizations - Visual data analysis techniques- interaction techniques - Systems and applications.

TOTAL: 45 PERIODS

OUTCOME:

- Appreciate the computational softwares and techniques for handling big data in business applications

REFERENCES:

1. Frank J Ohlhorst, “Big Data Analytics: Turning Big Data into Big Money”, Wiley and SAS Business Series, 2013.
2. Colleen Mccue, “Data Mining and Predictive Analysis: Intelligence Gathering and Crime Analysis”, Elsevier, Second Edition, 2015.

3. Michael Berthold, David J. Hand, "Intelligent Data Analysis", Springer, Second Edition, 2007.
4. AnandRajaraman and Jeffrey David Ullman, "Mining of Massive Datasets", Cambridge University Press, 2014.
5. Bill Franks, "Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics", Wiley and SAS Business Series, 2012.
6. Paul Zikopoulos, Chris Eaton "Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data", McGraw Hill, 2012.
7. Paul Zikopoulos, Dirk de Roos, Krishnan Parasuraman, Thomas Deutsch , James Giles, David Corrigan, "Harness the Power of Big data - The big data platform", McGraw Hill, McGraw-Hill Osborne Media, 2012.
8. Glenn J. Myatt, "Making Sense of Data I: A Practical Guide to Exploratory Data Analysis and Data Mining", John Wiley & Sons, Second Edition, 2014.
9. Pete Warden, "Big Data Glossary", O'Reilly, 2011.
10. Jiawei Han, Micheline Kamber "Data Mining Concepts and Techniques", Elsevier, Third Edition, 2011.
11. Tom White "Hadoop: The Definitive Guide" Third Edition, O'reilly Media, 2012
12. William Stallings, "Cryptography and Network security: Principles and Practices", Pearson/PHI, 5th Edition, 2010.
13. Mark Talabis, Robert McPherson, I Miyamoto and Jason Martin, "Information Security Analytics: Finding Security Insights, Patterns, and Anomalies in Big Data", Syngress Media, U.S., 2014.