MCA 415 Big Data Analysis

Teaching Scheme	Examination Scheme
Lectures: 3 hrs/Week Tutorials: 1 hr/Week Credits: 4	Class Test -12Marks Teachers Assessment - 6Marks Attendance – 12 Marks End Semester Exam – 70 marks

Prerequisite: - Database Management System, Data Mining and Warehousing.

Course Objectives:

1. To describe the concept of Big data and its features.
2. To understand the importance Big Data Analytics with various challenges.
3. To know about the architecture of Hadoop with its components.
4. To perform analysis on the data using R programming language.
5. To identify the role of cloud computing in Big Data.
6. To generate data and manipulating it using R.

Detailed Syllabus

UNIT I (6 Hours)
Introduction to Big Data Classification of Digital Data, Big Data and its importance, Four Vs, Drivers for Big data, Big data analytics, Classification of Analytics, Top Challenges Facing Big Data, Responsibilities of data scientists, Big data applications in healthcare, medicine, advertising.

UNIT II (6 Hours)
Hadoop Architecture Hadoop Architecture, Hadoop Storage: HDFS, Hadoop MapReduce paradigm, Introduction to Hive, Introduction to Pig.

UNIT III (6 Hours)
Introduction to NoSQL & Hadoop Introduction to NoSQL Advantages of NoSQL, SQL versus No SQL Introduction to Hadoop, Features of Hadoop, Hadoop Versions, Hadoop Versus SQL.

UNIT-IV (8 Hours)
Types of Analytics & Techniques Open source technology for Big Data Analytics – cloud and Big Dat – Mobile Business Intelligence and Big Data.

Head
Department of Computer Applications
Faculty of Computer Applications
Invertis University, Bareilly (UP)

Registrar
Invertis University
Bareilly

Dean Academics
Faculty of Computer App
Invertis University, Bare

UNIT V (8 Hours)
Predictive Analysis Predictive Analytics, Supervised, Unsupervised learning, Clustering Techniques.

UNIT VI (6 Hours)
Basics of R, Working of R - Creating, listing and deleting the objects in memory - The on-line help Data with R Objects, R data Frames and Matrices, Reading data in a file, Saving data, Generating data, Manipulating data using R

Text and Reference Books

1. 1 An Introduction to Statistical Learning: With Applications in R: Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani.
2. BIG Data and Analytics, Sima Acharya, Subhashini Chhellappan, Willey
3. VigneshPrajapati, "Big Data Analytics with R and Hadoop", Packet Publishing 2013.
4. The Culture of Big Data, Mike Barlow, by Oreilly
5. Big Data Analytics; Frank J. Ohlhorst, by Wiley

Course Outcomes:

After completing the course, students will be able to:

1. Understand the role and importance of Big Data and Big Data Analytics.
2. Understand the architecture of Hadoop.
3. Know the role of Pig and Hive.
4. Understand the concept of various types of Analysis.
5. Work on the provided data using R programming.