



Syllabus

Big Data Analytics and its Applications

Duration: 30 hours

Module I

What is big data, why big data, unstructured data, industry examples of big data, web analytics, big data and marketing, fraud and big data, risk and big data, big data and algorithmic in transport, big data and healthcare, big data in medicine, advertising and big data, big data technologies, introduction to Hadoop, open source technologies, cloud and big data, mobile business intelligence, Crowdsourcing analytics.

Module II

Databases and the relational algebra, Parallel databases, parallel query processing, in-database analytics, Map Reduce, Hadoop, relationship to databases, algorithms, extensions, languages, Key-value stores and NoSQL; tradeoffs of SQL and NoSQL

Module III

Topics in statistical modeling: basic concepts, experiment design, pitfalls, Topics in machine learning: supervised learning rules, trees, forests, nearest neighbor, regression), optimization (gradient descent and variants), unsupervised learning.

Text Book:

- **Intelligent Data Analysis**, Michael Berthold, David J. Hand, 2/e, Springer, 2007.
- **Taming the Big Data Tidal Wave : Finding Opportunities in Huge Data Streams with Advanced Analytics, Bill Franks**”, 1/e, John Wiley & sons, 2012.

Reference Book:

- **Making Sense of Data I**, Glenn J. Myatt, 2/e, John Wiley & Sons, 2014
- **Data Mining Concepts and Techniques**, Jiawei Han, Micheline Kamber, 2/e, Elsevier, Reprinted 2008.



Schedule

Duration: 30 hours			
Session	Content	Time	Date
S 1	Introduction of big data,	03:00-4:00 PM	16-Aug-2021
S 2	Importance of big data,	03:00-4:00 PM	20-Aug-2021
S 3	Types of big data	03:00-4:00 PM	23-Aug-2021
S 4	Industry examples of big data,	03:00-4:00 PM	27-Aug-2021
S 5	Web analytics,	03:00-4:00 PM	3-Sep-2021
S 6	Big data and Marketing,	03:00-4:00 PM	6-Sep-2021
S 7	Risk in big data,	03:00-4:00 PM	10-Sep-2021
S 8	Big data Applications,	03:00-4:00 PM	13-Sep-2021
S 9	Big data technologies,	03:00-4:00 PM	17-Sep-2021
S 10	Revision of Module 1	03:00-4:00 PM	20-Sep-2021
S 11	Introduction to Hadoop,	03:00-4:00 PM	24-Sep-2021
S 12	Concept of open source technologies,	03:00-4:00 PM	27-Sep-2021
S 13	Introduction to cloud and big data,	03:00-4:00 PM	1-Oct-2021
S 14	Concepts of mobile business intelligence,	03:00-4:00 PM	4-Oct-2021
S 15	Analysis in Big Data and Crowd sourcing analytics	03:00-4:00 PM	8-Oct-2021
S 15	Basics of databases and the relational algebra	03:00-4:00 PM	11-Oct-2021
S 16	Parallel databases, parallel query processing,	03:00-4:00 PM	18-Oct-2021



S 17	Database analytics	03:00-4:00 PM	22-Oct-2021
S 18	MapReduce, Hbase,	03:00-4:00 PM	25-Oct-2021
S 19	Relationship to databases, algorithms, extensions, languages	03:00-4:00 PM	29-Oct-2021
S 20	Introduction of statistical modeling: basic concepts,	03:00-4:00 PM	1-Nov-2021
S 21	Experiment design, pitfalls,	03:00-4:00 PM	8-Nov-2021
S 22	Introduction of machine learning	03:00-4:00 PM	12-Nov-2021
S 23	Types of Machine Learning	03:00-4:00 PM	15-Nov-2021
S 24	Supervised learning rules,	03:00-4:00 PM	22-Nov-2021
S 25	Introduction of trees, forests, nearest neighbor	03:00-4:00 PM	29-Nov-2021
S26	Statistical Regression Analysis	03:00-4:00 PM	3-Dec-2021
S 27	Optimization Techniques in Big Data	03:00-4:00 PM	6-Dec-2021
S 28	Unsupervised learning.	03:00-4:00 PM	10-Dec-2021
S 29	Key-value stores and NoSQL;	03:00-4:00 PM	13-Dec-2021
S 30	tradeoffs of SQL and NoSQL	03:00-4:00 PM	17 Dec-2021



	Report
Name of Activity	Big Data Analytics and its Applications (BCA-SP-502)
Date	16 th August 2021 to 17 th December 2021
Venue	BCA Classroom
Organized by	Computer Application Department
Resource Person	Mr. Abhay N. Tripathi (HoD, BCA)
Beneficiary	BCA 5th Semester (13 students)
Coordinator	Mr. Suraj Shukla ,Assistant Professor, MIMT
Objective	<p>This course on Big Data Analytics:</p> <ul style="list-style-type: none">• Students will be capable to quickly adapt to new technology in the field of Big Data, assimilate new information, and solve real world problems.• It consists of coherent body of ideas and methods to acquaint the student with the basic programs in the computational and human Intelligence field and their underlying theory.
Content	<p>With the initiative of IQAC, Mangalmai Institute of Management and Technology organized a add on certification course on “Big Data Analytics and its Applications”.</p> <p>Day1: The session started with the introduction of big data.</p> <p>Day 2: In this session, The importance of big data was discussed.</p> <p>Day 3: In this interactive session, the resource person introduce the types of big data.</p> <p>Day 4: The resource persons started the session with the example of big data.</p> <p>Day 5: In this session, the student came to know about web analytics.</p> <p>Day 6: In this session the resource person taught the big data and Marketing.</p> <p>Day 7: This session was focused on the risk in big data.</p> <p>Day 8: In this session, the resource persons discussed the application</p>



	<p>of big data.</p> <p>Day 9: Student learnt Big data technologies.</p> <p>Day 10: Revision of module 1 takes place.</p> <p>Day 11: Introduction to Hadoop was delivered by the resource person.</p> <p>Day 12: The session was on the concept of open source technologies.</p> <p>Day 13: The session starts with the introduction to cloud and big data.</p> <p>Day 14: In this session, concepts of mobile business intelligence was discussed.</p> <p>Day 15: Resource person discussed about the analysis in big data and crowd sourcing analytics and basics of databases and the relational algebra.</p> <p>Day 16: Parallel databases, parallel query processing was taught in the classroom.</p> <p>Day 17: In this session, the database analytics was done.</p> <p>Day 18: In this session, the resource person discusses the MapReduce, HBase</p> <p>Day 19: Relationship to databases, algorithm, extensions, languages was discussed..</p> <p>Day 20: In this session, The basic concept of statistical modelling was introduced..</p> <p>Day 21: The experiment design and pitfalls was taught.</p> <p>Day 22: This session was on the introduction of machine learning.</p>
--	---



	<p>Day 23: In this session, types of machine learning was taught.</p> <p>Day 24: In this session, the resource person delivered the supervised learning rules.</p> <p>Day 25: This session was on the introduction of trees, forests and nearest neighbor</p> <p>Day 26: The discussion was on the Statistical Regression Analysis.</p> <p>Day 27: The session was about Optimization techniques in Big Data.</p> <p>Day 28: The session was on the Unsupervised learning.</p> <p>Day 29: Key-value stores and NoSQL was delivered by the resource person.</p> <p>Day 30: The session ends with tradeoffs of SQL and NoSQL.</p>
Outcome of Activity	<p>On completion of the programme :</p> <ul style="list-style-type: none">• The candidates will be able to apply the knowledge of computing tools and techniques in the field of Big Data for solving real world problems encountered in the Software Industries.• Candidate will be able to analyze the various technologies & tools associated with Big Data and will be able to identify the challenges in Big Data with respect to IT Industry and pursue quality research in this field with social relevance.



List of Beneficiaries
BCA- V Sem

Sr. No.	Roll No.	Student Name
1	R190992106006	Arti Sinha
2	R190992106007	ASHUTOSH KUMAR
3	R190992106008	ASHWIN KUMAR SINGH
4	R190992106012	DEV RISHI
5	R190992106023	KARAN
6	R190992106024	KARTIKE KUMAR
7	R190992106029	LALIT CHAUHAN
8	R190992106034	MOHD ISRAR
9	R190992106037	MUKESH SINGH
10	R190992106046	RAMAVTAR
11	R190992106054	SHANI RAI
12	R190992106061	SUBHAM
13	R190992106066	VAQUAAR



Resource Person Profile

Name : Mr. Abhay N. Tripathi ,HOD, BCA

Core Skills: C, C++, Python, Hadoop

Qualification: B.Sc (PCM), MCA, PhD(P)

Experience: 20 years

Research Area: Data Science, Artificial Intelligence, Machine Learning and Big data Analytics.



Certificate Template:

 **MANGALMAI**
INSTITUTE OF MANAGEMENT & TECHNOLOGY

Certificate No. BCA/21-22/SP502/012

CERTIFICATE OF COURSE COMPLETION

This is to certify that
Shari Rai

Student of BCA, Batch (2019-22) has successfully completed 30 Hours
Specialization Course on Big Data Analytics
from Aug 2021 to Dec 2021 with Grade B+


Course Coordinator


Head of the Department


Chairman

 