PSGR
Krishnammal College for Women

College of Excellence, nirf 2023 - 4th Rank
Autonomous and Affiliated to Bharathiar University
Reaccredited with A++ grade by NAAC, An ISO 9001: 2015 Certified Institution
Peelamedu, Coimbatore-641004

DEPARTMENT OF DATA ANALYTICS (PG)

CHOICE BASED CREDIT SYSTEM (CBCS) &
LEARNING OUTCOME BASED CURRICULAR FRAMEWORK (LOCF)

MASTER OF DATA ANALYTICS

2023 – 2025 BATCH ONWARDS

PROGRAMME OUTCOMES

PO1: Demonstrate understanding of the core concepts of analytics to discover data-driven insights

PO2: Illustrate higher degree of technical skills that enable business decision making

PO3: Apply analytical and managerial skills to enhance employment potential

PO4: Exhibit holistic development with emphasis on values and ethics

PROGRAMME SPECIFIC OUTCOME

Graduates will be able to

PSO1: Demonstrate understanding of concepts in data science, statistical concepts and probability.

PSO2: Identify and analyze complex issues reaching substantiated conclusions using the techniques in data science.

PSO3: Design and propose innovative solutions for complex problems that meet the specified business needs.

PSO4: Ability to understand the industry requirements and to have sound knowledge about the professional skills required for data science.

PSO5: Create, select and apply appropriate techniques, tools, resources in data science for prediction and modeling of complex activities with an understanding of the limitation.

PSO6: Communicate effectively on complex tasks in profession as well as with society at large, such as, being able to comprehend and write effective reports, make effective presentations and provide as well as receive clear instructions.

PSO7: Apply ethical principles in research and commit to professional ethics and responsibilities.

PSO8: Recognize the need for lifelong learning and have the ability to engage in independent learning keeping in mind the rapid technological changes.

College of Excellence, nirf 2023 - 4th Rank
Autonomous and Affiliated to Bharathiar University
Reaccredited with A++ grade by NAAC, An ISO 9001: 2015 Certified Institution
Peelamedu, Coimbatore-641004

DEPARTMENT OF DATA ANALYTICS (PG)
CHOICE BASED CREDIT SYSTEM & LEARNING OUTCOME BASED CURRICULAR FRAMEWORK SYLLABUS&SCHEME OF EXAMINATION

MASTER OF DATA ANALYTICS – 2023-2025 BATCH ONWARDS
SEMESTER I , II & III

Semester	Subject Code	Title of paper	Instruction Hrs / week	Contact Hrs	Tutorial Hrs	Duration of Examination	Examination Marks			Credits
							CA	ESE	Total	
I	MDA2301	Paper 1: Descriptive Statistics	4	58	2	3	25	75	100	4
I	MDA2302	Paper 2: Foundations of Data Science	4	58	2	3	25	75	100	4
I	MDA2303	Paper 3: Linear Algebra	4	58	2	3	25	75	100	4
I	MDA2304	Paper 4: Data Structures	4	58	2	3	25	75	100	4
I	MDA2305	Paper 5: Information Retrieval	4	58	2	3	25	75	100	4
I	MDA23P1	Lab1: RDBMS Lab	5	75	-	3	25	75	100	3
I	MDA23P2	Lab2: Full Stack development Lab	5	75	-	3	25	75	100	3
I		Online Course	-	-	-	-	-	-	-	-
II	MDA2306	Paper 6: Probability and Distributions	5	73	2	3	25	75	100	5
II	MDA2307	Paper 7: Machine Learning	4	58	2	3	25	75	100	4
II	MDA2308	Paper 8: Big Data Framework	4	58	2	3	25	75	100	4
II / III	MDA23CE / MDA2313	Coursera: Predictive Analytics / Special Course: Business Analytics	3/3	45/43	- /2	- / 3	100/ 25	-/75	100$ /100#	3

II	MDA23E2 / MDA23E3	Elective – I: Marketing Analytics Elective – I: Text Analytics	4	58	2	3	25	75	100	4
II	MDA23P3	Lab 3: Data Analytics Lab II(Hadoop, MapReduce & R,SPSS)	5	75	-	3	25	75	100	3
II	MDA23P4	Lab 4: Machine learning Lab	5	75	-	3	25	75	100	3
II		Online Course	-	-	-	-	-	-	-	-
III	MDA2310	Paper 10:Big Data with Spark and Hive	4	58	2	3	25	75	100	4
III	MDA2311	Paper 11:Applied Statistics	3	43	2	3	25	75	100**	3
III	MDA2312	Paper 12: Data Mining	4	58	2	3	25	75	100	4
III	MDA23E4 / MDA23E5 / MDA23E6	Elective II: IoT Analytics Social Media Analytics Web Data Analytics	4	58	2	3	25	75	100	4
III	MDA23S1	Special Course: Research Methodology	2	30	-	3	-	100	100*	2
II / III	MDA23CE / MDA2313	Coursera: Predictive Analytics / Business Analytics	3/3	45/43	- /2	- / 3	100/ 25	-/75	100$/ 100#	3
III	MDA23P5	Lab 5:Data Analytics Lab III(Spark, Hive)	5	75	-	3	25	75	100	3
III	MDA23P6	Lab 6: Data Mining and Visualization Lab	5	75	-	3	25	75	100	3
III	MDA23COM	Comprehensive Exam	-	-	-	3	-	-	100	Grade
III	MDA23IST	Summer Internship	-	-	-	-	-	-	-	Grade
III	MNM22CS2	Self-Study Course: Cyber Security - II	-	-	-	-	100	-	100	Grade
III		Job Oriented course - : Microsoft Power BI Data Analyst	-	-		-	-	-	-	-
III	17MONL1	Online Course	-	-	-	-	-	-	-	-

***Only ESE will be conducted for100 marks** **$ Coursera Course**

**** Open Book Exam for CA**

#No ESE for Theory, only CA & Project Viva Voce will be conducted

List of Electives:

S. No	Course Code	Course Title
Semester II – Elective 1		
1	MDA23E1	Artificial Intelligence and Analytics in Business
2	MDA23E2	Marketing Analytics
3	MDA23E3	Text Analytics
Semester III – Elective 2		
4	MDA23E4	IOT Analytics
5	MDA23E5	Social Media Analytics
6	MDA23E6	Web Data Analytics

EVALUATION PATTERN

CA Question from each unit comprising of

One question with a weightage of 2 Marks :2 x 3 = 6
One question with a weightage of 5 Marks (Internal Choice at the same CLO level) :5 x 3 =15
One question with a weightage of 8 Marks (Internal Choice at the same CLO level) :8 x 3 =24

Total :45 Marks

ESE Question Paper Pattern:

Question from each unit comprising of

One question with a weightage of 2 Marks: 2 x 5=10
One question with a weightage of 5 Marks (Internal Choice at the same CLO level): 5 x 5 =25
One question with a weightage of 8 Marks (Internal Choice at the same CLO level): 8 x 5 =40

Continuous Assessment Pattern

Theory

CIA Test: 5 marks (conducted for 45 marks after 50 days)
Model Exam: 7 marks (Conducted for 75 marks after 85 days (Each Unit 15 Marks))
Seminar/Assignment/Quiz: 5 marks
Class Participation: 5 marks
Attendance: 3 marks
Total: 25 Marks

Practical Lab

Performance: 7 marks
Regularity: 5 marks
Model Exam: 10 marks
Attendance: 3 marks
Total: 25 marks

ESE Practical Pattern

The End Semester Examination will be conducted for a maximum of 75 marks respectively with a maximum 15 marks for the record and other submissions if any.

Open Book Exam

CA Test Pattern: 4 (4 out of 6) x 15 Marks = 60 Marks

Cyber Security II:

Quiz: 60 Marks

Case Study: 20 Marks

Poster: 20 Marks

PG Special Course:

Section A: 5 questions (Internal choice): 25 marks

Section B: 5 questions (Internal choice): 75 marks

Total: 100 marks

Marks allotted for attendance component in CA

91-100% attendance : 3 Marks

81-90% attendance : 2 Marks

75-80% attendance : 1 Marks

MDA2301	DESCRIPTIVE STATISTICS	Category	L	T	P	Credit
		Theory	58	2	-	4

Preamble

The course introduces the measures of central tendency and dispersion. It also provides thestudents with systematic knowledge in correlation, regression and outlier analysis in R Programming.

Course Learning Outcomes

Upon the successful completion of the course, students will be able to

CLO Number	CLO Statement	KnowledgeLevel
CLO1	Demonstrate the basic concepts of statistics	**K2**
CLO2	Apply the concepts of correlation, regression	**K3**
CLO3	Identify the methods for different measures of central tendency, dispersion	**K4**
CLO4	Evaluate the methods for representation of data	**K5**
CLO5	Construct various plots, outliers for regression, diagnostics	**K6**

Mapping with Programme Outcomes

CLOs	PO1	PO2	PO3	PO4
CLO1	S	M	S	S
CLO2	S	S	S	M
CLO3	S	S	M	S
CLO4	M	S	S	S
CLO5	S	S	M	S

S- Strong; M-Medium; L-Low

Syllabus

UNIT I **(12 Hrs)**

Introduction: **Nature and scope of Statistics, limitations of statistics** - Types of data: Concept of population and sample, primary and secondary data, quantitative and qualitative data, discrete and continuous data, cross-sectional and time series data. **Scales of measurement: Nominal, Ordinal, Ratio and Interval.** Case Studies using Statistical Methods in Excel.

UNIT II **(12 Hrs)**

Tabulation of Data- Difference between Classification and Tabulation – Parts of Tabulation –Rules for Tabulation – Types of Tables. Diagrammatic representations: - Line diagram, bar diagram, pie diagram and sub-divided bar diagram, **Frequency distribution and cumulative frequency distribution and their graphical representations, Frequency polygon, histogram, ogive, frequency curves, stem and leaf displays. Exploring R Basics – R Features – R Basic Data types – R Basic Operators – Data visualization in R.**

UNIT–III **(12 Hrs)**

Univariate data: Different measures of location, dispersion, relative dispersion, **skewness and kurtosis**, Moments, Quantiles and measures based on them – comparison with moment Measures - **Box-plot and detection of outliers. Trimmed mean and Winsorised mean** – Simple problems. Interpretation of Histogram, Box Plot in R - Outliers on Univariate Data in R

UNIT IV **(11 Hrs)**

Correlation – Scatter diagram – Karl Pearson Coefficient of Correlation -Spearman's Rank coefficient of Correlation - **Rank Correlation Coefficient**.
Interpretation of Scatter Plot in R

UNIT V **(11 Hrs)**

Regression: Introduction – Uses of regression analysis – **regression lines** – **regression equationsof X on Y and Y on X – regression equation in terms of correlation table.**
Visualization plots for Linear Regression in R – Interpretation of Logistic Regression in R.

Text Book

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	S.C.Gupta ,V.K.Kapoor	Fundamentals of Mathematical Statistics	Sultan Chand & Sons	2019, 4th Edition Revised
2	Price Michael	Excel 2010 In Easy Steps	Tata MC Graw-Hill	2010,1st Edition
3	R.S.N.Pillai and Bhagavathi	Statistics Theory and Practice	Sultan Chand & Sons	2019, 8th Edition
4	R.Wilcox	Basic Statistics	Oxford University Press	2009, 1st Edition
5	S.P.Gupta	Statistical Methods	Sultan Chand & Sons	2021,46th Edition
6	V.Bhuvaneswari	Data Analytics with R Programming	Scitech Publications	2018,1st Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Murray R Spiegel and Larry J Stephens	Statistics	Schaum's Outline	2008,4th Edition
2	R.S.N.Pillai	Statistics	S. Chand Publishing Company Pvt Ltd	1992, 8th Edition
	1. https://www.indiabix.com/data-interpretation/questions-and-answers/ 2. https://www.mathsisfun.com/data/pictographs.html			

Pedagogy: Lectures, Simulation Exercises, Demonstration

Course Designers

1. Dr. T. A. Albinaa
2. Dr. N. Radha

MDA2302	FOUNDATIONS OF DATA SCIENCE	Category	L	T	P	Credit
		Theory	58	2	-	4

Preamble

This course introduces the essential knowledge about foundations of data science with python. It provides various steps involved in the data science process and core python with advanced concepts like regular expressions, exception handling, multithreading. It also introduces the python libraries that are most commonly used for data analysis.

Prerequisite

- Basic understanding of open source software
- Database concepts

Course Learning Outcomes

On successful completion of the course, the students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand data classification, process of big data technology, userroles and skills in data science and its tools.	K2
CLO2	Apply the fundamental concepts, tools and techniques of data science in 360 view of Customer.	K3
CLO3	Analyze the methodologies of data science and its tools.	K4
CLO4	Evaluate the utilities and functionalities of Python	K5
CLO5	Design solutions for data analytics problems using Python.	K6

Mapping with Programme Outcomes

CLOs	PO1	PO2	PO3	PO4
CLO1	S	M	M	M
CLO2	S	S	M	M
CLO3	S	S	S	S
CLO4	S	S	S	S
CLO5	S	S	S	S

S- Strong; M-Medium; L-Low

Syllabus

UNIT I **(12 Hrs)**

Data Evolution: Data Growth- a Perspective - IT Components - Business Process – Landscape - Data to Data Science –- Understanding data: Introduction - Types of Data: Numeric - Categorical – Graphical - High Dimensional Data - Data Classification - **Sources of Data: Time Series - Transactional Data - Biological Data -Spatial Data - Social Network Data -** Introduction to Python programming – Libraries – Plugins – Problem solving.

UNIT II **(12 Hrs)**

Introduction to Data Science - **A Discipline -Data Science vs Statistics, Data Science vs Mathematics**, Data Science vs Programming Language, Data Science vs Database, Data Science vs Machine Learning. Data Analytics - Data Manipulation with NumPy - Introduction to NumPy arrays

and their operations-**Manipulating arrays: indexing, slicing, and reshaping -** Working with NumPy, Working with statistical toolbox.

UNIT III **(11 Hrs)**

Relation: Data Science, Analytics and Big Data Analytics. Data Science Components –Big data technology - Data Science user - roles and skills **- Data Science use cases.** Introduction to the Pandas library - Loading and exploring datasets - **Data indexing and selection** – Data cleaning and preprocessing - Aggregation and summarization - Data ingestion: export/import data from various file formats (CSV, Excel, etc.).

UNIT IV **(11 Hrs)**

Digital Data-an Imprint: Evolution of Big Data –What is Big Data –Sources of Big Data**. Characteristics of Big Data 6Vs –Big Data Myths** –Data Discovery-Traditional Approach. Introduction to data visualization - Plotting with Matplotlib: line plots, scatter plots, bar plots, and histograms – **Exploratory data analysis** - Data visualizations to explore and analyze datasets.

UNIT V **(12 Hrs)**

Big data Technology: Big Data Technology Process –**Big Data Exploration -Data Augmentation – Operational Analysis –Security and Intelligence**. Introduction to NLTK-Installing NLTK and downloading necessary Resources-Basic NLTK functionalities and data structures-Tokenization and Text Preprocessing: **Text cleaning and normalization techniques** - Exploring NLTK for text processing applications like sentiment analysis.

Text Book

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	V. Bhuvaneswari, T. Devi	Big Data Analytics: A Practitioner's Approach	Bharathiar University	2016, 1st Edition
2	Wesley J.Chun	Core Python Programming Application Programming	Pearson Education Publication, United States.	2016, 3 rd Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Mark Lutz	Programming Python	O'Reilly Media	2010,4th Edition
2	Wes McKinney	Python for Data Analysis: Data Wrangling with pandas,	Grayscale Indian Edition	2022,3rd Edition

		NumPy, and Jupyter		
3	Alberto Boschetti, Luca Massaron	Python Data Science Essentials	Packt Publishing	2016,2nd Edition
4	Wesley J.Chun	Core Python Programming	Prentice Hall Publication.	2006, 2nd Edition

Pedagogy: Lectures, Demonstration and Case Studies

Course Designers

1. Mrs. K. Gandhimathi
2. Mrs. G. Anitha

MDA2303	LINEAR ALGEBRA	Category	L	T	P	Credit
		Theory	58	2	-	4

Preamble

The course introduces the principles underlying linear equations and vector spaces. It also provides the concepts of Eigen values, Eigen vectors and Positive Definite Matrices in Scilab.

Course Learning Outcomes

Upon the successful completion of the course, students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Demonstrate competence with the basic ideas of linear Algebra including the concepts of vector spaces, Determinants, Eigen values and Eigenvectors and positive definite matrices	K2
CLO2	The ability to understand the principles of Linear Algebra	K3
CLO3	Apply properties of linear spaces to specific mathematical structures	K4
CLO4	Compose clear and accurate proofs using the concepts of linear Algebra	K5
CLO5	Appreciate the significance of vector spaces and positive definite matrices	K6

Mapping with Programme Outcomes

CLOs	PO1	PO2	PO3	PO4
CLO1	S	S	S	S
CLO2	S	S	S	M
CLO3	S	S	S	S
CLO4	S	S	M	S
CLO5	S	S	S	S

S- Strong; M-Medium; L-Low

Syllabus

UNIT I **(12 Hrs)**

The Geometry of Linear Equations- An Example of Gaussian Elimination- Matrix Notation and Matrix Multiplication - **Triangular Factors and Row Exchanges**- Inverses and Transposes – Implementation of basic matrix operations(Sum, Multiplication, Square root, Cube, Inverse, Transpose, Rank, Fibonacci) in scilab

UNIT II **(11 Hrs)**

Vector Spaces: **Vector Spaces and Subspaces** – Solving Ax=0 and Ax=b - Linear Independence, Basis, and Dimension- The Four Fundamental Subspaces- **Graphs and Networks**- Linear Transformations, Implementation of vector space and subspace, Linear independency, to find the reduced row echelon formof a matrix in Scilab.

UNIT III **(11 Hrs)**

Determinants: Introduction- **Properties of the Determinant**- Formulas for the Determinant- Applications of Determinants – Implementation of determinants, Properties - associative, commutative

and distributive property in a matrix using Scilab

UNIT IV **(12 Hrs)**

Eigenvalues and Eigenvectors: Introduction- Diagonalization of a Matrix. - Difference Equations and Powers A^k- Differential Equations and e^{At} - **Complex Matrices- Similarity Transformations –** Implementation of Eigenvalues and eigenvectors, solving equations by Gauss elimination, Gauss Jordan Method and Gauss Siedel in Scilab

UNIT V **(12 Hrs)**

Positive Definite Matrices: Minima, Maxima, and Saddle Points - Tests for Positive Definiteness- **Singular Value Decomposition, Minimum Principles**, The Finite Element Method, Implementation of Minima, Maxima, Saddle Points and plot the function of the derivatives in Scilab

Text Book

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Gilbert Strang	Linear Algebra and Its Application	Academic Press	2020, 4th Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	David C. Lay, Steven R. Lay, Judi J. McDonald	Linear Algebra and Its Applications	Education	2016,5th Edition
2	Peter D. Lax	Linear Algebra and Its Applications	Wiley Publication	2014,2nd Edition

Pedagogy: Lectures, Demonstration and Case Studies

Course Designers

1. Dr. T.A.Albinaa
2. Dr. T.Brindha

MDA2304	DATA STRUCTURES	Category	L	T	P	Credit
		Theory	58	2	-	4

Preamble

This course covers the various data structures, including arrays, structures, stacks and queues. It includessorting and searching techniques and effective search methods in Binary trees. This course also deals withgraph data structures.

Prerequisite

Discrete mathematics

Course Learning Outcomes

On successful completion of the course, the students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand the concepts of arrays, strings and algorithms for basic operations.	**K2**
CLO2	Apply concept of stacks, queues, linked list and algorithms for basic operations.	**K3**
CLO3	Identify the familiarity with major algorithms and data structures	**K4**
CLO4	Analyze appropriate algorithms and data structures for various applications	**K5**
CLO5	Formulate the computational complexity of various algorithms	**K6**

Mapping with Programme Outcomes

CLOs	PO1	PO2	PO3	PO4
CLO1	S	M	M	M
CLO2	S	M	M	M
CLO3	S	S	M	M
CLO4	S	S	S	M
CLO5	S	S	S	M

S- Strong; M-Medium; L-Low

Syllabus

UNIT I **(12 Hrs)**

Basics: Algorithm Specification – Data Abstraction – Performance Analysis: Time, Space Complexity – Arrays – Dynamic Allocated Arrays – Structures and Unions – Polynomials - **Sparse Matrices- Representation of Multidimensional Arrays – Strings.**

UNIT II **(12 Hrs)**

Stacks and Queues: Stacks – Stacks Using Dynamic Arrays - Queues - Circular Queues Using Dynamic Arrays - Evaluation of Expressions - **Multiple Stacks And Queues - Recursion. Linked Lists:** Singly Linked List And Chains – Representing Chains in C – Linked Stack And Queues –**Polynomials - Additional List Operations - Sparse Matrices – Doubly Linked List.**

UNIT III **(11 Hrs)**

Searching: Introduction - Sequential Search - **Binary Search** – **Analysis. Sorting:** Introduction - InsertionSort - Selection Sort - Merge Sort - **Quick Sort - Heaps and Heap Sort - Analysis.**

UNIT IV **(11 Hrs)**

Efficient Binary Search Trees: Binary Tree – Traversals - Optimal Binary Search Trees – AVL Trees- **KDTrees.**

UNIT V **(12 Hrs)**

Graphs: The graph Abstract Data Type- Elementary graph operations- Minimum cost spanning trees-shortest paths and transitive closure- **AOV networks** –**AOE networks.**

Case Study: Location Identification, Game Development, Google Knowledge Graph, Path Optimization

Text Book

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Ellis Horowitz, Sartaj Sahni and Anderson Freed	Fundamentals of data structures in C	UniversityPress	2009,2nd Edition
2	Ellis Horowitz, Sartaj Shani, Sanguthevar Rajasekaran	Fundamentals of computer algorithms	Galgotias Publications private limited	2013,2nd Edition
3	Robert L Kruse	Data Structures & Program Design	Pearson India	2008,2nd Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Mark Allen Weiss	Data Structures and Algorithm Analysis in C++	Pearson Education	2012, 4th Edition
2	Adam Drozdek	Data Structures and Algorithm in C++	Course Technology Ptr	2013,4th Edition
3	Ch. Rajaramesh C.V. Sastry, Rakesh Nayak	Data Structures and Algorithms	Wiley India	2019,1st Edition
4	Akhilesh Kumar Srivastava	A Practical Approach to Data Structure and Algorithm with Programming in C.	Arcler Education Inc	2019, 1st Edition

Pedagogy: Lectures, Group Discussion, case study

Course Designers

1. Dr. N. Radha
2. Mrs. G. Anitha

MDA2305	INFORMATION RETRIEVAL	Category	L	T	P	Credit
		Theory	58	2	-	4

Preamble This vector space model, text classification and evaluation methods of IR systems.

Prerequisite

- Statistics
- Basic Programming

Course Learning Outcomes

On the successful completion of the course, students will be able to

CO Number	CLO Statement	Knowledge Level
CLO1	Demonstrate the concepts of document representation, document indexing, digital information storage, retrieval and distribution	K2
CLO2	Apply the concepts of vector spaces and classifiers to perform document classification.	K3
CLO3	Examine the advantages and disadvantages of different information-retrieval models	K4
CLO4	Determine the effective search strategies for IR systems	K5
CLO5	Able to develop information retrieval system for specific use cases.	K6

Mapping with Programme Outcomes

CLOs	PO1	PO2	PO3	PO4
CLO1.	S	M	M	M
CLO2.	S	M	M	M
CLO3.	S	S	M	M
CLO4.	S	M	S	M
CLO5.	S	S	S	M

S- Strong; M-Medium; L-Low

Syllabus

UNIT I (12 Hrs)

Boolean retrieval: Information retrieval problem - Processing Boolean queries - Boolean model versus ranked retrieval. **The term vocabulary** and postings lists: Document delineation and character sequence decoding - Determining the vocabulary of terms - Faster postings list intersection via **skip pointers** - **Positional postings and phrase queries.**

UNIT II (12 Hrs)

Dictionaries and tolerant retrieval: Search structures for dictionaries - **Wildcard queries** - Spelling correction - Phonetic correction. Index construction: Hardware basics - **Blocked sort-based indexing - Single-pass in-memory indexing** - Distributed indexing - Dynamic indexing- Other types of indexes.

UNIT III **(12 Hrs)**

Scoring, term weighting and the **vector space model**: Parametric and zone indexes - Term frequency and weighting - The vector space model for scoring. **Evaluation in information retrieval**: Information Retrieval system – Standard test Collection – Evaluation of unrankedretrieval sets – **Evaluation of ranked retrieval results** – Assessing relevance.

UNIT IV **(11 Hrs)**

XML retrieval: Basic XML concepts – A vector space model for XML retrieval – Evaluation ofXML retrieval - Text-centric vs. data-centric XML retrieval. **Text classification and Naive Bayes**:The text classification problem - Naive Bayes text classification - Properties of Naive Bayes -Feature selection - **Evaluation of text classification.**

UNIT V **(11 Hrs)**

Vector space classification: Document representations and measures of relatedness in vector spaces – **Rocchio classification** - Flat clustering: Clustering in information retrieval - **Evaluation of clustering - K-means** – Web search basics - Web characteristics - Advertising as the economicmodel – Search user experience – Basic Page Rank.

Text Book

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Christopher D. Manning, Prabhakar Raghavan, Henrich Schutze	Introduction to Information Retrieval	Cambridge University Press	2008,1st Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Stefan Buttcher et.al	Information Retrieval - Implementing and Evaluating	MIT Press	2016,1st Edition
2	Dr Ricardo Baeza-Yates et.al	Modern Information Retrieval: The Concepts andTechnology	Addison Wesley	2011, 2nd Edition
3	David A. Grossman and Ophir Frieder	Information Retrieval	Universities Press	2010, 2nd Edition

Pedagogy: Lectures, Demonstrations, Guest Lecture, Video Lectures

Course Designers

1. Dr. N.Radha
2. Dr. V. Bharathi

MDA23P1	RDBMS LAB	CATEGORY	L	T	P	CREDITS
		III	-	-	75	3

Preamble

This course provides sound introduction to implement the relational database management systems concepts in SQL. This course also provides various exercises to implement the schema and Table creation, ER diagram, integrity constraints, Joins, Nested queries, functions, procedures, cursors, triggers, exception handling and to create Simple real time applications.

Prerequisite

- Database concepts
- Programming concepts

Course Outcomes

On the successful completion of the course, students will be able to

CO Number	CO Statement	Knowledge Level
CO1.	Implement the databases concepts and SQL queries as per user requirement	**K2**
CO2.	Apply specific SQL commands on relational tables for different situations	**K3**
CO3.	Analyse use cases and create constraints suitable for the given situation.	**K4**
CO4.	Create and analyse a database using SQL DML/DDL commands	**K6**
CO5.	Design and build a GUI application	**K6**

Mapping with Programme Outcomes

Cos	PO1	PO2	PO3	PO4
CO1.	S	S	S	S
CO2.	S	S	S	S
CO3.	S	S	S	S
CO4.	S	S	S	S
CO5.	S	S	S	S

S- Strong; M-Medium; L-Low

Syllabus

- Exercises to create schema, Table, ER Diagram using online tool
- Exercises to implement the concepts of null constraint, unique constraint, integrity constraints, check constraints, default constraints.
- Exercises to implement nested queries.
- Exercises to implement the concepts of partitioning queries
- Exercises to create a view from the tables
- Exercises to create functions and procedures
- Exercise to create triggers and queries
- Exercises to create cursors and exceptions

- List of Projects using frontend and backend (Topics not limited to)
    1. Student Admission Management
    2. Library Management System
    3. Payroll Management System
    4. Telecom
    5. Online Retail Application Database
    6. Inventory Control Management
    7. Hospital Management System
    8. Banking system management
    9. Railway ticket reservation management system
    10. An SMS-based Remote Server Monitoring System
    11. Blood Donation Database
    12. Cooking Recipe Portal
    13. Electricity Bill payment system
    14. Online tax management system
    15. Training and placement system

Course Designers:

1. Dr.N.Radha
2. Dr.S. Meera

MDA23P2	FULL STACK DEVELOPMENT LAB	CATEGORY	L	T	P	CREDIT
		III	-	-	75	3

Preamble

This lab course is intended to explore concepts into full stack development through HTML, CSS, javascript and Node.js. It makes the students to learn various designing approaches and techniques to provide a robust UI/UX experience using various software tools to implement web pages, style sheets, forms, events, server side applications.

Course Learning Outcomes

On the successful completion of the course, students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand the web design elements, functions, files, data connections, forms, events.	**K2**
CLO2	Apply web design methods to solve problems	**K3**
CLO3	Analyze the web page design requirements and design web pages.	**K4**
CLO4	Create a web application using HTML,CSS, javascript and Node.js.	**K6**
CLO5	Design and build web application using HTML,CSS, Javascript and Node.js	**K6**

Mapping with Programme Outcomes

CLOs	PO1	PO2	PO3	PO4
CLO1	S	S	S	S
CLO2	S	S	S	S
CLO3	S	S	S	S
CLO4	S	S	S	S
CLO5	S	S	S	S

S- Strong; M-Medium; L-Low

LIST OF PROGRAMS

- Exercise to design a website with forms, frames, links and table using HTML tag.
- Exercise to build a countdown timer using HTML, CSS and JavaScript to display the remaining time until a specific event.
- Exercise to design a navigation menu with HTML unordered lists and CSS styles to create a dropdown or responsive menu.
- Exercise to design responsive web page that adjusts its layout and design based on the screen size using CSS media queries.

- Exercise to design a platform for embedding video from you tube using iframe.
- Exercise to design web server using built-in module in Node.js.
- Exercise to develop a HTML and CSS program to find the average and grade of students marks.
- Exercise to implement an image gallery using a grid layout and CSS to showcase a collection of image.
- Exercise to develop and demonstrate JavaScript with POP-UP boxes and functions.
- Exercise to build a CRUD operation for a product using laravel framework.

List of Projects using web technologies (Topics not limited to):

1. E-Commerce website
2. Social Media Platform
3. Online Learning Platform
4. Task Management Application
5. Weather Application
6. Event Booking Platform
7. Blogging Platform
8. Real-time Chat Application
9. Student feedback system
10. Online Resume Builder

Course Designers:

1. Mrs.K. Gandhimathi
2. Mrs. G. Anitha

MDA2306	PROBABILITY AND DISTRIBUTIONS	CATEGORY	L	T	P	CREDIT
		Theory	73	2	-	5

Preamble

This course introduces the fundamental concepts of probability and random variables. It also provides knowledge in discrete and continuous distributions. It deals with various sampling distributions like t, F, chi-square distributions etc. It also provides the students with systematic knowledge in R Programming.

Course Learning Outcomes

On the successful completion of the course, students will be able to

CLO Number	CO Statement	Knowledge Level
CLO1	Understand the concepts of probability and its distributions.	K2
CLO2	Solve problems involving random variables and their functions	K3
CLO3	Examine the sampling distributions to contribute to the process of making rational decisions in analytical problems.	K4
CLO4	Compare the various sampling tests and choose test suitable for theproblem	K5
CLO5	Formulate the hypothesis test for various complex problems	K6

Mapping with Programme Outcomes

CLOs	PO1	PO2	PO3	PO4
CLO1	S	M	M	M
CLO2	S	S	M	M
CLO3	S	S	S	M
CLO4	S	M	S	M
CLO5	S	S	S	M

S- Strong; M-Medium; L-Low

Syllabus

UNIT I **(14 Hrs)**

Random Experiment: **Sample space, Different types of events**. Approaches of measuring probability:Mathematical, Statistical and Axiomatic probability, Laws of probability-conditional probability – Baye'stheorem.Interpretation of different types of events in R -Equally Likely Model in R - Baye's Principle in R.

UNIT II **(14 Hrs)**

Random variable - **Discrete and Continuous Random variables** - Distribution function and its properties– Expectation and Moment Generating function: Raw and Central moments-relationship between central and raw moments-moments about an arbitrary value-moment generating function-**properties of moment generating function**-characteristic function.Interpretation of Mean and Variance of a Random variable in R -Distribution function in R –Exercises in R.

UNIT III **(15 Hrs)**

Discrete Distributions: Binomial, Poisson, Geometric, Uniform distributions and their properties. **Continuous Distributions: Rectangular, Exponential, Normal distributions and their properties**- Simple problems Interpretation of Binomial Distributions in R - Poisson Distributions in R - Continuous Distributions in R.

UNIT IV **(15 Hrs)**

Central limit theorem, Liapounoff Form, Lindberg Levy Form. Sampling Distributions –Distribution of Sample Mean, t-Distribution, Chi-square Distribution, Distribution of Sample Variance, F-Distribution, Degrees of freedom, Interval Estimation, Confidence interval for one mean and difference of two means. **Testing of Hypothesis : Relation between confidence interval and testing of hypothesis-level of significance and p-value**. Classification of hypothesis tests.Interpretation of t-Distribution in R - Chi-square Distribution in R - F Distribution in R.

UNIT V **(15 Hrs)**

Large sample tests: Single mean, Difference of two means, Single proportion. **Small sample tests: t-test for single mean, difference between two means-F-test for equality of two population variances-Chi- square test for single Variance**, Chi-square test for goodness of fit- Chi-square test for independence of attributes and homogeneity and equality of proportions.

Interpretation of Large Sample and Small sample test in R. Simple Problems in R

Text Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Ravichandran. J	Probability and Statistics for Engineers	Wiley	2019, 2nd Edition
2	V. Bhuvaneswari	Data Analytics with R Programming	Scitech Publications Pvt Ltd	2018,1st Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Rohatgi V.K	An Introduction to Probability Theory and Mathematical Statistics	John Wiley	2015, 3rd Edition
2	Wiley.Scymour Lipschuts	Probability	Schaum's Outline	2011, 1st Edition
3	Goon A.M., Gupta M.K. & Dasgupta B	An Outline of Statistical Theory	World Press	2016,6 th Edition
4	V.Bhuvaneswari	Data Analytics with R Programming	Scitech Publications Pvt Ltd	2018, 1st Edition

Pedagogy : Lectures, Case studies, Demonstrations

Course Designers:

1. Dr.T.A.Alinaa
2. Dr.C.R.Parvathy

MDA2307	MACHINE LEARNING	Category	L	T	P	Credit
		Theory	58	2	-	4

Preamble

This course introduces the fundamentals of Machine Learning and algorithms. It also covers various supervised and unsupervised learning algorithms used for classification, prediction and clustering.

Prerequisite

- Statistics and probability
- Linear Algebra

Course Learning Outcomes

On the successful completion of the course, students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand the concepts of machine learning.	K2
CLO2	Apply supervised and unsupervised learning algorithms for classification, prediction and clustering.	K3
CLO3	Analyze the logic behind the execution of various classifiers.	K4
CLO4	Evaluate the performance of different algorithms available for mining data.	K5
CLO5	Predict solution for real world problems.	K6

Mapping with Programme Outcomes

CLOs	PO1	PO2	PO3	PO4
CLO1	S	M	M	M
CLO2	S	M	M	M
CLO3	S	S	S	M
CLO4	S	S	S	M
CLO5	S	S	S	M

S- Strong; M-Medium; L-Low

Syllabus

UNIT I **(11 Hrs)**

Introduction: Machine learning – Examples and Applications - Perspectives and Issues in Machine learning **- Input: Concepts, Instances, and Attributes - Output: Knowledge Representation**- Credibility: Evaluating What's Been Learned: Training and Testing - **Predicting Performance - CrossValidation - Other Estimates - Counting the cost.**

UNIT II **(11 Hrs)**

Decision Tree Learning: Decision tree representation – Decision tree learning –Issues in decision tree learning- Random forest -**Bayesian Learning: Naïve Bayes classifier** - Instance Based

Learning: Introduction – k- nearest neighbor Learning - **Radial Basis Function, Case based reasoning.**

UNIT III **(12 Hrs)**

Artificial Neural Network - Introduction – Neural Network Representation - Perceptrons – GradientDescent and Delta Rule-**Multilayer Networks and Back propagation Algorithm**–**Logistic Regression** - Linear models for Regression- Linear Discriminant Analysis.
Use Cases: Spam Filters, Price Prediction, Demand Forecasting, Image Recognition

UNIT IV **(12 Hrs)**

SVM : Introduction – Kernel methods - formulation and computation- SVM Linear classifier – SVMwith two variables –non-linear classifier-Polynomial kernels- Radial Basis Function Kernels

UNIT V **(12 Hrs)**

Clustering Methods.- Introduction – K- Means- Expectation-Maximization Algorithm- Hierarchical Clustering - Choosing the Number of Clusters – **Gaussian Mixture Models** – **Frequent Pattern (FP) Growth – PCA – Kernal PCA**.
Use cases: Customer Segmentation, Anomaly detection, Recommender System.

Text Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Ian Witten	Data mining: Practical Machine Learning Tools and Techniques	Morgan Kaufmann Publishers	2016,4th Edition
2	Tom M. Mitchell	Machine Learning	Tata McGraw-Hill, New Delhi	2013,1st Edition
3	K.P. Soman	Machine Learning with SVM and Other Kernel Methods	PHI Publishing	2011,2 nd Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Dr.Pablo Rivas	Deep Learning for Beginners	Packet Publishing	2020,1 st Edition
2	Aston Zhang, Zahery C.Lipton, Muli and Alexander J Smola	Drive into Deep Learning	Cambridge University Press	2022,1 st Edition
3	Saikat Dutt, Subramanian Chandramouli, Amit Kumar Das	MachineLearning	Pearson Publisher	2019,1 st Edition

Pedagogy : Lectures, Case studies, Demonstrations

Course Designers:

1. Mrs.G.Anitha
2. Dr.N.Radha

MDA2308	BIG DATA FRAMEWORK	Category	L	T	P	Credits
		Theory	58	2	-	4

Preamble

This course introduces big data framework, technologies, storage and Hadoop ecosystem. It also deals with the concepts of MapReduce, Hbase ,Scala and MongoDB.

Prerequisite

- Database Management systems
- Programming concepts

Course Learning Outcomes

On successful completion of the course, students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand the characteristics of big data, concepts of Hadoop ecosystem and Scala programming.	K2
CLO2	Apply Mapreduce programming model to process big data.	K3
CLO3	Analyze Hbase and Scala and its uses in big data processing.	K4
CLO4	Evaluate techniques used for handling NoSQL databases MongoDB and Hbase.	K5
CLO5	Design programs for big data applications using Hadoop components.	K6

Mapping with Program Outcomes

CLOs	PO1	PO2	PO3	PO4
CLO1	S	S	S	M
CLO2	S	S	S	M
CLO3	S	S	S	M
CLO4	S	S	S	M
CLO5	S	S	S	M

S- Strong; M-Medium; L-Low

Syllabus

UNIT I **(11 Hrs)**

Introduction to Big data: Introduction – Big Data- Characteristics of Big Data – Big data management architecture- Examining Big Data Types – Big Data Technology Components -– Bigdata analytics –**Big data analytics examples - Web Data Overview – Web Data in Action**.

UNIT II **(11 Hrs)**

Hadoop : Introduction – History of Hadoop **- Hadoop Ecosystem**- Analyzing data with Hadoop – Hadoop Distributed File System- Design **- HDFS concepts - Hadoop filesystem** –Data flow – Hadoop I / O - Data integrity – Serialization - Setting up a Hadoop cluster - Cluster specification - cluster setup and installation – **YARN**

UNIT III **(12 Hrs)**

Map Reduce: Introduction – **Understanding Map, Reduce functions** - Scaling out - **Anatomy of aMap Reduce Job Run** - Failures – **Shuffle and sort** – Map reduce types and formats - features – counters - sorting - Unit test with MR unit- local test

UNIT IV **(12 Hrs)**

Hbase: Concepts - Hbase Vs RDBMS - **Creating records- Accessing data – Updating and deleting data – Modifying data**- exporting and importing data.
SCALA: Introduction - Classes and objects- Basic types and operators **- built-in control structures** -functions and closures – **inheritance**

UNIT V **(12 Hrs)**

NoSQL Databases: Introduction to NoSQL**- MongoDB**: Introduction – Data types – **Creating, Updating and deleing documents -Querying** – Introduction to indexing – Capped collections.
Case Studies: Hadoop at Facebook, Nutch Search Engine.

Text Book

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Tom White	Hadoop: The Definitive Guide	O'Reilly Media	2015,4th Edition
2	Judith Hurwitz	Big Data for Dummies	John Wiley & Sons	2013,1 st Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Bill Franks	Taming the Big Data Tidal wave	John Wiley & Sons	2012,4th Edition
2	Martin Odersky,Lex Spoon, Bill Venners	Programming in Scala	Artima Press, California.	2019, 4th Edition
3	Shashank Tiwari	Professional NoSQL	John Wiley & Sons	2013, 1 st Edition
4	Shannon Bradshaw,Eoin Brazil & Krishna Chodorow	MongoDB.The Definitive Guide-Powerful & Scalable Data Storage	O'Reilly Media	2020, 3rd Edition
5	Boris lublinsky, Kevin t. Smith, Alexey, Yakubovich	Professional Hadoop Solutions	Wiley	2015, 1st Edition
6	Achintya Kumar	Big Data Black Book: Hadoop 2, MapReduce, Hive, YARN, Pig, R and Data Visualization	DreamTech Press	2019,1 st Edition

Pedagogy: Lectures, Group Discussions, Case studies

Course Designers

1. Dr. S. Meera
2. Mrs. G.Anitha

MDA23CE	COURSERA - PREDICTIVE ANALYTICS	Category	L	T	P	Credits
		Theory	45	-	-	3

Course contents **(45 Hrs)**

S. NO.	COURSE NAME	LINK
1	Predictive Modelling and Analytics(10 Hrs)	https://www.coursera.org/learn/predictive-modeling-analytics
2	Python Data Products for PredictiveAnalytics Specialization (40 Hrs)	https://www.coursera.org/specializations/python-data-products-for-predictive-analytics#courses
3	Practical Predictive Analytics: Models and Methods (6 Hrs)	https://www.coursera.org/learn/predictive-analytics

MDA23P3	DATA ANALYTICS LAB II	Category	L	T	P	Credit
		Practical	-	-	75	3

Preamble

This course provides implementation of statistics concepts in R and SPSS. It also provides sound introduction to implement the Hadoop framework. This course also provides various exercises to implement in the distributed environment through map reduce programming.

Prerequisite

- Statistics
- Big data framework
- Java

Course Learning Outcomes

On successful completion of the course, the students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Demonstrate concepts of probability in R	**K2**
CLO2	Apply different probability distributions in SPSS	**K3**
CLO3	Apply specific Mapper and reducer functions for different situations	**K3**
CLO4	Design and build a hadoop cluster. Create NoSQL databases using MongoDB.	**K6**
CLO5	Develop DML,DDL commands in Hbase.	**K6**

Mapping with Program Outcomes

CLOs	PO1	PO2	PO3	PO4
CLO1	S	S	M	M
CLO2	S	S	M	M
CLO3	S	S	M	M
CLO4	S	S	S	M
CLO5	S	S	S	M

S- Strong; M- Medium; L- Low

Syllabus

- Setting up a hadoop environment
- Exercises to implement file management tasks using Hadoop
- Exercises to implement Map reduce program
- Exercises to implement Stock count Map reduce program.
- Exercises to implement CRUD, Aggregating and indexing operations in MongoDB.
- Exercises to apply the concept of MapReduce in MongoDB .
- Exercises to implement DDL, DML commands using HBase.
- Exercises to implement concepts of probability and distributions in R
- Exercises to implement concepts of probability and distributions in SPSS

Pedagogy: Demonstrations

Course Designers:

1. Dr. S.Meera
2. Dr. T.A. Albinaa

MDA23P4	MACHINE LEARNING LAB	Category	L	T	P	Credits
		Practical	-	-	75	3

Preamble

This course introduces Sci-kit learn, the popular machine learning libraries. It also provides exercises to implement Sci-kit learn syntax and tools to apply machine learning algorithms.

Prerequisite

- Knowledge in Open source - Python
- Machine Learning

Course Learning Outcomes

On successful completion of the course, students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Apply Machine Learning concepts to real time applications.	**K3**
CLO2	Build models using classification algorithm for real world problems	**K6**
CLO3	Build models using clustering algorithm for real world problems	**K6**
CLO4	Create classification and clustering models	**K6**
CLO5	Test and evaluate the models	**K3**

Mapping with Program Outcomes

CLOs	PO1	PO2	PO3	PO4
CLO1	S	S	S	M
CLO2	S	S	S	M
CLO3	S	S	S	M
CLO4	S	S	S	M
CLO5	S	S	S	M

S- Strong; M-Medium; L-Low

Syllabus

- Exercises to load dataset into sci-kit learn
- Exercise for Building models in sci-kit learn
- Exercise to extract features from datasets
- Exercise to implement Regression
- Exercise to implement Classification
- Exercise to implement Clustering
- Exercises for Model selection and evaluation
- Exercises to Build a data pipeline

Pedagogy: Demonstrations

Course Designer:

1. Mrs.G..Anitha
2. Dr.N.Radha

ELECTIVES

		Category	L	T	P	Credit
MDA22E1	**ARTIFICIAL INTELLIGENCE AND ANALYTICS IN BUSINESS**	**Theory**	**58**	**2**	**-**	**4**

Preamble:

The course covers an overview of implementation of Analytics and AI strategies in Business and provide details of disruption and transformation brought in various domains like Banking, Healthcare and Life sciences, Retail and Exponential technologies

Prerequisite

- Basic Knowledge of Analytics and Artificial Intelligence
- Knowledge in different domains

Course Learning Outcomes:

Upon the successful completion of the course, student will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1.	Understand the need of Analytics and AI strategy in Business Transformation	K2
CLO2.	Apply Analytics and AI strategy indifferent domains	K3
CLO3.	Analyse mitigation of Fraud and customer retention using AI in different Domains	K4
CLO4.	Assess the behavior of different cyber threats and attacks	K5
CLO5	Develop domain-based applications	K6

Mapping with programme Outcomes

CLOs	PO1	PO2	PO3	PO4
CLO1.	S	M	M	M
CLO2.	S	M	M	M
CLO3.	S	M	M	M
CLO4.	S	S	M	M
CLO5	S	S	M	M

S-Strong; M-Medium; L-Low

Syllabus

UNIT I **(11Hrs)**

Analytics and AI Strategy for Business Transfer: Re-engineering Business to think AI and Analytics –Robust Data Monetization Strategy–Accelerated Decision-making with Real-Time Analytics–Analytics as a Service Model–Analytics-Led Enterprise Transformation.

UNIT II **(12Hrs)**

Banking Industry Transformed by Analytics and AI: Redefining Banking Industry –**AI powered Financial services – Fraud Mitigation through AI** – Reorienting Customer Retention and Risk Management – Advantage of AI in Fintech Companies – AI-Driven Transformations in Insurance –**Adopting Digital Based Insurance Model**.

UNIT III **(12Hrs)**

Redefining Healthcare and Life Sciences: AI adoption in Healthcare – **Real-world Evidence Based Analytics improving Treatment outcomes**– Leveraging Patient and Drug similarity Analytics–AI :A Boon to the Life Science Industry – **Analytics and Genomics**.

UNIT IV **(12Hrs)**

Analytics and AI in Retail: AI-powered shopping experience – Emergence of Smart Consumers – Recommendation Engines for Personalizing Experiences–**Evolution of Smart Retailers–Omni channel Experiences–** Fluid Supply Chain Transformation.

UNIT V **(11Hrs)**

Exponential Technologies underpinned by Analytics and AI: Beating Cyber attacks with Analytics – Connected Car Technology reshaping Automotive Industry – **IoT Analytics – Crypto currency Analytics – Chat bots**– Redefining the Talent Landscape.

Text Book

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Sameer Dhanrajani	AI and Analytics- Accelerating Business Decisions	Wiley	2018,1st Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Stuart Russel and Peter Norvig	Artificial Intelligence – A Modern Approach	Pearson Education Press	2021,2nd Edition
2	Kevin Knight, Elaine Rich, B.Nair	Artificial Intelligence	Mc Graw Hill	2009,3rd Edition
3	George F.Luger	Artificial Intelligence	Pearson Education	2001,5th Edition
4	Nils J. Nilsson	Artificial Intelligence: A new Synthesis	Morgan Kaffman	2002,2nd Edition

Pedagogy: Lectures, Case Analysis, Group Discussions, Demonstrations

Course Designer

1. Dr.N.Radha
2. Dr.T.A.Albinaa

MDA22E6	MARKETING ANALYTICS	Category	L	T	P	Credit
		Theory	58	2	-	4

Preamble

This course contains the strategies of marketing analytics, forecasting techniques, Customer Needs and retailing. This will also offer the market research tools involved in retailing.

Prerequisite

- Foundations of Data science

Course Learning Outcomes

On successful completion of the course, the students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand the concepts of marketing data and effective use of Microsoft Excel	**K2**
CLO2	Apply Forecasting Techniques to improve response rates formarketing campaigns	**K3**
CLO3	Analyse Market segmentation based on Cluster Analysis	**K4**
CLO4	Estimate methods for processing marketing data in prediction and market segmentation.	**K5**
CLO5	Formulate solutions in excel for various use cases of marketinganalytics methods	**K6**

Mapping with Programme Outcomes

CLOs	PO1	PO2	PO3	PO4
CLO1	S	M	L	M
CLO2	S	L	L	M
CLO3	S	S	L	M
CLO4	M	M	S	M
CLO5	S	M	M	M

S- Strong; M-Medium; L-Low

Syllabus

UNIT I **(11 Hrs)**

Marketing Data Summarization - Slicing and Dicing Marketing Data with Pivot Tables - **Using Excel Charts to Summarize Marketing Data** - Using Excel Functions to Summarize Marketing Data.

UNIT II **(12 Hrs)**

Forecasting Techniques: Simple Linear Regression and Correlation - Using Multiple Regression to Forecast Sales - Forecasting in the Presence of Special Events - **Modeling Trend and Seasonality** -

Ratio to Moving Average Forecasting Method - Winter's Method - **Using Neural Networks to Forecast Sale**

UNIT III **(12 Hrs)**

Customer Needs: Conjoint Analysis - Logistic Regression - Discrete Choice Analysis – **CustomerValue - Introduction to Customer value, Benefits**

UNIT IV **(12 Hrs)**

Market segmentation: Cluster Analysis - User-Based Collaborative Filtering - Collaborative Filtering - **Using Classification Trees for Segmentation.**

UNIT V **(11 Hrs)**

Retailing and market research tools : Retailing - Introduction to retailing, Market Basket Analysis andLift - Marketing Research Tools - **Principal Components Analysis**

Text Book

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Wayne.L.Winston	Analytics: Data driven techniques with MS-Excel	Wiley Publications	2014,1st Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Stephan Sorger	Marketing Analytics: Strategic models and metrics	Create Space Independent Publishing Platform	2013,1st Edition
2	Chuck Hemann, Ken Burbary	Digital Marketing Analytics	Pearson Publication	2013,2nd Edition
3	Robert W. Palmatier, J. Andrew Petersen , Frank Germann	Marketing Analytics: Based onFirst Principles	Bloomsbury Academic	2022,2nd Edition

Pedagogy: Lectures, Demonstrations, Group Discussions, Case studies, Online References

Course Designers:

1. Dr. N. Radha
2. Mrs.G.Anitha

MDA23E3	TEXT ANALYTICS	Category	L	T	P	Credits
		Theory	58	2	-	4

Preamble

This course aims to introduce the principles of linguistics, natural language processing and text analytics. It focuses on models, techniques and tools for dealing with text data and to gain knowledge in recent advances in text analytics. It also provides text analytic framework and machine learning techniques for modelling the information in textual sources.

Prerequisite

- Statistics
- Data Structure
- Fundamentals of Information Retrieval

Course Learning Outcomes

On the successful completion of the course, students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand the basics, language syntax and structure of textanalytics and NLP	**K2**
CLO2	Apply the concept of text preprocessing, tokenization and normalization	**K3**
CLO3	Analyze and parsing techniques, text classification, summarization And semantic relations	**K4**
CLO4	Evaluate classification models and clustering techniques	**K5**
CLO5	Build the model using text analytics techniques	**K6**

Mapping with Program Outcomes

CLOs	PO1	PO2	PO3	PO4
CLO1	S	S	M	M
CLO2	S	S	S	M
CLO3	S	S	S	M
CLO4	S	S	S	M
CLO5	S	S	S	M

S- Strong; M-Medium; L-Low

Syllabus

UNIT I **(11 Hrs)**

Introduction to Text Analytics: Introduction – Origin - Analytics in business and Industry -skills - benefits - **Text Analytics Process Road Map** - The Fundamentals of Content Analysis - **Planning for Text Analytics**

UNIT II **(11 Hrs)**

Natural Language Basics- Language Syntax and Structure – Language Semantics - Language as Data: A Computational Model – Language, Contextual & Structural Features – Text Corpora – Working withText - String Operations and Methods - **Text Analytics Frameworks**

UNIT III **(12 Hrs)**

Processing and Understanding Text - Text Tokenization – Text Normalization – Understanding Text Syntax and Structure – **Parts of Speech Tagging** – Shallow parsing – **Dependency-based parsing – Constituency-based parsing**

UNIT IV **(12 Hrs)**

Text Classification - Feature Extraction – Classification Algorithms: **Multinomial Naïve Bayes** - Evaluating Classification Models – **Building a Multi-Class Classification System**-Application and uses – Text Summarization: Text Normalization – Feature Extraction – **Key phrase Extraction** - Text Similarity: Analyzing Term Similarity - Clustering Greatest Movies of All Time: **K-Means Clustering**

UNIT V **(12 Hrs)**

Semantic and Sentiment Analysis: Semantic Analysis- Exploring WordNet – **Analyzing lexical semantic relations – Word sense disambiguation** – Named Entity Recognition – **Analyzing Semantic Representations** – Case Study: Sentiment Analysis of Movie Reviews

Text Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Murugan Anandarajan, Chelsey Hill, Thomas Nolan	Practical Text Analytics: Maximizingthe Value of Text Data	Springer	2016,1st Edition
2	Dipanjan Sarkar	Text Analytics with Python: A Practical Real-World Approach to Gaining Actionable Insights from your Data	Apress Media LLC	2016,2nd Edition
3	David Robins	Text Mining with R: A Tidy Approach	O'Reilly Media Inc	2017,1st Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Steven Bird, Ewan Klein, Edward Loper	Natural Language Processing with Python	O'Reilly Media, Inc	2009,1st Edition
2	Benjamin Bengfort & Tony Ojeda & Rebecca Bilbro	Applied Text Analysis with Python:Enabling Language-Aware Data Products with Machine Learning	O'Reilly Media, Inc.	2018,1st Edition

Pedagogy: Lectures, Demonstrations, Guest Lecture, Video Lectures

Course Designers

1. Ms.M.Nandhini
2. D. K. Gandhimathi

SEMESTER III

Course Code	Course Name	Category	L	T	P	Credit
MDA2310	BIG DATA WITH SPARK AND HIVE	Theory	58	2	-	4

Preamble

This course introduces the fundamental concepts of Spark and Hive. It also provides knowledge on SQL in Spark and understanding Hive data model.The course also covers data definition and data manipulation operations using HiveQL and applying these to perform analytics.

Prerequisite

- Big Data Framework
- SQL

Course Learning Outcomes

On the successful completion of the course, students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand the programming constructs of Spark and database management using HiveQL	K2
CLO2	Apply the concepts of Spark QL and Hive in simple tasks. Understand writing scripts using SQL in Spark and perform various HiveQL queries by applying RDBMS concepts	K3
CLO3	Analyze different types of data and its sources for use in Spark Applications.	K4
CLO4	Appraise techniques for data manipulation in Hive.	K5
CLO5	Design and develop scripts in Spark and Hive for various situations	K6

Mapping with Programme Learning Outcomes

CLOs/ PLOs	PLO1	PLO2	PLO3	PLO4
CLO1	S	S	M	M
CLO2	S	S	M	M
CLO3	S	S	S	M
CLO4	S	S	S	M
CLO5	S	S	M	M

S-Strong;M-Medium

Syllabus

UNIT I **(11 Hrs)**

Overview of Big data and spark - Spark Philosophy - **History of Spark** - Running Spark- **Spark Architecture**-**Data Frames**-Transformations-End to end example–**Spark Toolset**-Spark run on cluster-Developing spark Applications-Deploying Spark.

UNIT II **(11 Hrs)**

Structure API -Structured Spark Types -API Execution - Basic structured operations - Working with Different types of Data-Booleans, strings, dates, complex types-**Working with Json-Aggregations**– Grouping – Joins-**Data sources**.

UNIT III **(12 Hrs)**

Spark SQL-Running Spark SQL Queries-**Tables-Views-Databases**-Advanced sub queries-Dataset- Transformations-RDD:- Creating RDD-Transformations- Actions- saving files-Advanced RDD - Key-value RDD - **Distributed shared variables - Accumulators.** Streaming: Streaming Fundamentals - Processing Design Points - Structure Streaming Basics - core concepts - Inputandoutput.

UNIT IV **(12 Hrs)**

Developing Hive – Services in Hive – Understanding Hive Data model- HiveQL: Data definition –**Databases – Alter database – Create Tables – Alter Tables – Manage Tables – Drop Tables**. HiveQL: Data manipulation – Load Data into managed tables – **Insert into tables from Queries** -CreatingTables and Loading Them inOneQuery-**Exporting Data**.

UNIT V **(12 Hrs)**

HiveQL: Queries–SELECT…FROM Clauses–Where Clauses–Group by clauses-**Join Statements–Order by and Sort by-Distribute by with Sort by-Cluster by**–casting-UNION All -**HiveQL: Views**- HiveQL: Indexes- Case Studies.

Text Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Bill Chambers and Matei Zaharia	Spark: The Definitive Guide	O'Reilly Media	2018, 1st Edition
2	Jason Rutherglen, Dean Wampler, Edward Capriolo	Programming Hive	O'Reilly Media	2016, 2nd Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Hanish Bansal, Saurabh Chauhan	Apache Hive Cookbook	Packt publishing	2016, 1st Edition
2	Mike Frampton	Mastering Apache Spark	Pack Publishing	2015, 1st Edition
3	Jules.S, Brooke Wenig, Tathagata Das & Denny Lee	Learning Spark	O'Reilly Media	2020, 2nd Edition
4	Holden Karau & Rachel Warren	High Performance Spark	O'Reilly Media	2017, 1st Edition
5	Dayong Du	Apache Hive Essentials	Packt Pub Ltd	2015, 2nd Edition

Pedagogy: Lectures**,** Demonstrations, Case studies

Course Designers:

1. Mrs.G. Anitha
2. Dr.S. Meera

Course Code	Course Name	Category	L	T	P	Credits
MDA2311	APPLIED STATISTICS	Theory	43	2	-	3

Preamble

This course presents the different aspects of statistics like ANOVA, design of an experiments, control charts, multivariate analysis, Correlation and regression, which enables the student to analyze, organize, present and interpret data effectively.

Course Learning Outcomes

Upon the successful completion of the course, students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand the basic concepts of analysis of variance	K2
CLO2	Apply the concept of partial, multiple correlation and regression for solving problems	K3
CLO3	Apply statistical inference and control charts to applications	K4
CLO4	Analyse and interpret data using techniques like multivariate statistics	K5
CLO5	Design the experiments using methods like Randomized block and Latin squares	K6

Mapping with Programme Learning Outcomes

CLOs /PLOs	PLO1	PLO2	PLO3	PLO4
CLO1	S	M	S	S
CLO2	S	S	S	M
CLO3	S	S	M	M
CLO4	M	S	S	M
CLO5	S	S	S	M

S-Strong; M-Medium

Syllabus

UNIT I **(8 Hrs)**

Analysis of variance: **One way and Two way analysis: Introduction–Single factor experiment and linear statistical model for one way ANOVA**– Fixed effects mode and ANOVA – Computation for sum of squares – Single factor experimental and linear statistical model for two way ANOVA– Fixed effective model for two way ANOVA – Computation for sum of squares. Exploring ANOVA Test in PSPP.

UNIT II **(8 Hrs)**

Introduction – Randomized block design- Advantages of a completely randomized experimental design **– Latin squares – Significance of Latin squares** – Assumption in the Analysis of Latin squares–Randomized block Vs.Latin squares–**Latin Cubes–Factorial**

experiment. Case Studies to analyze Latin Square Design in PSPP.

UNIT III **(8 Hrs)**

Introduction – Statistical quality control: Relation Between Confidence Limit and control limit-Types of Control chart's – **Control charts for variables X-chart, R-chart, S-chart, X-chart** – Out of control situations in control chart and process monitoring–Process capability and process Capability index–Six sigma: Six sigma metrics. Case Studies to determine Control Charts in Excel.

UNIT IV **(9 Hrs)**

Multivariate Analysis – **Basic concepts – Measurement Scales – Measurement Error -- Statistical significance–Types of multivariate techniques**–graphical examination of the data– Missing data and its approaches–methods of detection of outliers– testing the assumptions of Multivariate analysis - Multivariate Analysis in Excel.

UNIT V **(12 Hrs)**

Correlation ratio: **Multiple and partial correlation - Coefficient of multiple correlation: Properties of multiple correlation coefficient -** Coefficient of partial correlation –Multiple Regression Analysis - Normal Equations for Least Square Regression Plane- Assumptions of Linear Multiple Regression Analysis – Generalizations of more than three variables. Correlation and Regression Analysis in Excel.

Text Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Dr.J.Ravichandran	Probability and Statistics for Engineers (Unit I: Chapter 13, Unit III : Chapter 15)	Wiley	2019, 2nd Edition
2	S.P.Gupta	Statistical Methods (Unit II: Chapter 6)	Sultan Chand & Sons	2021, Volume II, 46th Revised Edition
3	Joseph F.Hair, Rolph E Anderson, Ronald L Tatham, William C.Black	Multivariate Dataanalysis (Unit IV: Chapter1, Chapter 2:pg 1-82)	Pearson Education India	1998,7th Revised Edition
4	S.C.Gupta,V.K.Kapoor	Fundamentals of Mathematical Statistics (Unit V:Chapter12 : Sections12.4-12.10)	Sultan Chand & Sons.	2020, 4th Revised Edition
5	Price Michael	Excel 2019 In Easy Steps	Tata MC Graw - Hill	2018, 3rd Revised Edition

Pedagogy: Lectures**,** Demonstrations, Case studies

Course Designers:

1. Dr.T.A.Albinaa
2. Dr.T.Brindha

Course Code	Course Name	Category	L	T	P	Credit
MDA2312	DATA MINING	Theory	58	2	-	4

Preamble

This course presents the basic concepts of data mining and data warehousing, various data mining techniques like classification, clustering, association rule mining. The course also introduces various
Applications of data mining, data visualization using Tableau

Prerequisite

- Database Management Systems
- Probability and Statistics

Course Learning Outcomes

On successful completion of the course, the students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand basic concepts of data mining and data warehousing	K2
CLO2	Apply data mining techniques to carry out simple data mining tasks	K3
CLO3	Examine data mining techniques like classifications, clustering	K4
CLO4	Compare data mining algorithms using Tools	K5
CLO5	Design solutions with data visualization using Tableau	K6

Mapping with Programme Learning Outcomes

CLOs/ PLOs	PLO1	PLO2	PLO3	PLO4
CLO1	S	S	S	S
CLO2	S	S	S	S
CLO3	S	S	S	S
CLO4	S	S	S	S
CLO5	S	S	S	S

S-Strong;M-Medium

Syllabus

UNIT I **(11 Hrs)**

Introduction to data mining: Mining from database - Data mining functionalities - Mining patterns -**Classification of data mining systems - Major issues in Data mining**. Data warehouse - Multidimensional Data Model - Data Warehouse Architecture -Data Warehouse Implementation - **Datawarehousingto Data Mining.**

UNIT II **(12 Hrs)**

Data Preprocessing: Need for preprocessing–Data summarization–**Data cleaning**–Data integration-Data transformation–**Data reduction–Data discretization-**Ensemble Classifier: XG Boost.

UNIT III **(12 Hrs)**

Association Rule Mining: Basic concepts - Frequent Itemset Mining Methods –**Pattern Evaluation Methods**–Density Based Clustering- Grid-Based Methods–**Clustering High-Dimensional Data–Evaluation of clustering**

UNIT IV **(11 Hrs)**

Outlier Detection: Outliers and Outlier Analysis - Outlier Detection Methods - Proximity-Based Approaches - **Outlier Detection in High-Dimensional Data** – Data visualization : Foundations forbuilding visualizations-visualizing data – working with data in Tableau – **Moving from foundational to Advanced visualizations.**

UNIT V **(12 Hrs)**

Mining Complex Data-Time-Series, Symbolic Sequences, and Biological Sequences–Mining Graphs and Networks – Other Methodologies of Data Mining -**Applications of Data Mining: Financial Data Analysis–Retail, Marketing and Telecommunication Industries**-Intrusion Detection and Prevention-Recommender Systems– Data Mining Trends.

Text Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	JaiweiHan,Micheline Kamber	Data Mining-concepts and techniques	Morgan Kaufmann Publishers, San Francisco	2012, 3rd Edition
2	Joshua N.Milligan	Learning Tableau	Packt Publishing	2015, 1st Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Joshua N.Milligan	Learning Tableau	Morgan Kaufmann Publishers, San Francisco	2011, 3rd edition
2	Arun K.Pujari	Data Mining Techniques	Universities Press, Hyderabad	2001, 1st Edition
3	Soman K P	Data mining from theory to practice	PHI Learning Pvt. Ltd, New Delhi	2005, 2nd edition

Pedagogy: Lectures**,** Demonstrations, Case studies

Course Designers:

1. Dr.S. Meera
2. Mrs.G. Anitha

Course Code	Course Name	Category	L	T	P	Credit
MDA23S1	RESEARCH METHODOLOGY	Theory	30	-	-	2

Preamble

This course presents the concepts of research, types of research, research design, literature review, writing reports and adhering to research ethics.

Prerequisite

This course is most appropriate for postgraduate students who are interested in research but do not Have prior research experience.

Course Learning Outcomes

On successful completion of the course, the students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand the concepts of research, research design, research process for measurement, scaling and data collection	**K2**
CLO2	Apply statistical methods for processing the research data	**K3**
CLO3	Examine different methods of design and research approaches	**K4**
CLO4	Analyze the process of hypothesis testing and report writing	**K5**
CLO5	Design solutions for research problems in a responsible and ethical manner	**K6**

Mapping with Programme Learning Outcomes

CLOs/ PLOs	PLO1	PLO2	PLO3	PLO4
CLO1	S	M	M	M
CLO2	S	M	M	M
CLO3	S	S	S	M
CLO4	S	S	S	M
CLO5	S	S	S	S

S-Strong;M-Medium

Syllabus

UNIT I **(6 Hrs)**

An Introduction:–Meaning of research–Objective of research–**Types of research**–Research approaches – Significance of research – Research methods versus Methodology – Research and scientific method – Research process-Criteria of Good Research.

Defining the research problem: What is research problem? – **Selecting the problem** – Necessity of defining the problem. Research Design: Meaning of research design – Need for research design – Features of a good design– **Different research designs.**

UNIT II **(6 Hrs)**

Measurement and Scaling: Quantitative and Qualitative Data – **Classification of Measurement Scales**–Goodness of Measurement Scales-Sources of Error in Measurement–Technique of Developing Measurement Tools–Scaling – Scale Classification Bases. Data Collection: Introduction- Experiments and Surveys-**Collection of Primary Data– Collection of secondary data** – Selection of appropriate method for data collection.

UNIT III **(6 Hrs)**

Data Preparation: Data Preparation Process-Some problems in Preparation Process-Missing Values and Outliers-**Types of Analysis**– Statistics in Research.Testing of Hypotheses: Introduction to Hypothesis–**Basic concepts concerning testing of Hypotheses** - Procedure for Hypothesis testing – Measuring the power of a Hypothesis testing -**Tests of Hypothesis.**

UNIT IV **(6 Hrs)**

Interpretation and report writing: Meaning of interpretation–Technique of interpretation–Significance of report writing – Different steps in writing report – **Layout of the research report –Typesof reports.** Research Publications: Preparing Research papers for journals, Seminars and Conferences – Design of paper using template, Plagiarism, Calculations of Impact factor of a journal, citation Index, **ISBN&ISSN.**

UNIT V **(6 Hrs)**

Research Ethics –Brief history and analytical basis of research ethics, responsible conduct in research(Honesty in Science: Integrity, Authorship, Conflicts of Interest, Privacy and Confidentiality, Informed Consent, Risk/Benefit Assessment), The legal regulation of research ethicsin India (From UGC, MHRD and other governing agencies), Regulatory requirements relevant to international research.

Text Book

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Kothari,C.R Gaurav Garg	Research Methodology–Methods and Techniques	New Age International(P) Ltd.	2020, 4th edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	R.Panneerselvam	Research Methodology	Prentice Hall India Learning Private Limited	2014, 4th edition
2	Ranjit Kumar	Research Methodology – A step- by-step guide for beginners	Pearson Education	2019, 5th edition
3	Deepak Chawla and Neena Sondh	Research Methodology, Concepts and Cases	Vikas Publishing House Pvt. Ltd.	2011
4	Committee on Science, Engineering, and Public Policy; National Academy of Sciences; National Academy of Engineering; Institute of Medicine	On BeingaScientist,A Guide to Responsible Conduct in Research	National Academy of Sciences	2009, 3rd edition
5	Mark Suckow, Bill Yates	Research Regulatory Compliance		E Book ISBN: 9780124200654,1st edition

Pedagogy: Lectures, Demonstrations, Case Studies, Group Discussions

Course Designers:

1. Dr. S. Meera
2. Mrs.G.Anitha

Course Code	Course Name	Category	L	T	P	Credit
MDA2313	BUSINESS ANALYTICS	Theory	43	2	-	3

Preamble

This course provides the fundamentals of domain knowledge in the functional areas of Data Analytics. Various important functional areas such as Health care analytics, Banking and Finance,Telecommunication and Retail Analytics are elaborated in this course. Few Use cases are also covered.

Prerequisite

- Foundations of Data Science
- Big data framework

Course Learning Outcomes

Upon successful completion of this course, students should be able to:

CLO Number	CLO Statement	Knowledge Level
CLO1	Compare various domain areas and their challenges	**K2**
CLO2	Apply the concepts of analytics to make better decisions	**K3**
CLO3	Examine use cases for different domains	**K4**
CLO4	Evaluate the challenges faced in various domains and choose appropriate analytics solutions in all domains	**K5**
CLO5	Propose suitable analytics solutions as required by the use cases	**K6**

Mapping with Programme Learning Outcomes

CLOs/ PLOs	PLO1	PLO2	PLO3	PLO4
CLO1	S	M	M	M
CLO2	S	M	M	M
CLO3	S	S	M	M
CLO4	S	M	S	M
CLO5	S	S	S	M

S-Strong;M-Medium

Syllabus

UNIT I **(8 Hrs)**

Healthcare analytics–Introduction-Potential contributions-Challenges of healthcare industry-current and future state of healthcare analytics–top healthcare analytics adaptations

Use Cases: Predictive Analytics for Patient Health Management, Clinical Decision Support Systems.

UNIT II **(8 Hrs)**

Banking and Finance: Systems of Banking–Commercial Banking–New Financial Services: Overview of Analytics in Insurance: Key Insurance Analytics–Challenges–**Health Insurance Analytics, Life Insurance Analytics**- Types of Insurance –Housing Finance.

Use Cases: Credit Risk Assessment, Fraud Detection and Prevention, Churn Prediction and Customer Retention.

UNIT III **(8 Hrs)**

Telecommunication: Introduction -End-User Needs and Demands-**Telecom Business Use Cases: Network Performance Optimization, Customer Experience Management, Customer Segmentation and Target Marketing.**

UNIT IV **(8 Hrs)**

Retail analytics–Understanding the new consumer–Marketing in a consumer-driven era-Managing the brand to drive loyalty

Use Cases: Customer Segmentation and Personalization, Market Basket Analysis and Cross-Selling, E- commerce Analytics

UNIT V **(11 Hrs)**

Case studies: Walmart (Inventory Management, Supply chain Optimization), Netflix(Content Recommendation,Content Production), Facebook (User Engagement**, Ad Targeting),**Uber (Dynamic Pricing, Route Optimization, Driver Performance Optimization), Amazon (Product Recommendation, Customer Service Analytics), Kaggle(Data Science Competitions)

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Dwight McNeill	A Framework for Applying Analytics in Healthcare:What Can Be Learned from Best Practices in Banking, Retail, Politics and Sports	Pearson Education	2013, 1st Edition
2	Gomez Clifford	Banking and Finance Theory Law and practice	PHI Learning	2011, 1st Edition
3	Patricia L.Saporito	Applied Insurance Analytics: A Framework for Driving More ValuefromData Assets, Technologies and Tools	Pearson Education LTD	2014, 1st Edition
4	Anders Olsson	Understanding Changing Telecommunications	Wiley Publications	2005, 1st Edition
5	Jennifer Le Claire, Danielle Dahlstrom, Vivian Braun	Business Analytics in Retail for dummies	Wiley Publications	2013, 2nd IBM Limited edition
6	Alistair Croll	Lean Analytics: Use Data to Build a Better Startup faster	OReilly Publishers	2013, 2nd Edition
7	Bernard Marr	Big Data in Practice–How 45 successful companies used big data analytics to deliver extraordinary results	Wiley Publications	2016, 1st Edition
8	Purba Halady Rao	Business Analytics. An application Focus	PHI Learning private ltd	2013, 1st Edition

Pedagogy: Lectures, Group Discussions, Demonstrations, Case studies

Course Designers

1. Mrs. G. Anitha
2. Dr. N. Radha

EVALUATION PATTERN		
CA	-	10
Case study Discussion	-	15
Mini-Project	-	25
External Viva Voce	-	50
Total	-	**100**

Question Paper Pattern for CA

One question with a weightage of 2 Marks: 2 x 3 = 6
One question with a weightage of 5 Marks (Internal Choice at the same CLO level) : 5 x 3 =15
One question with a weightage of 8 Marks (Internal Choice at the same CLO level) : 8 x 3 =24
Total : 45 Marks

Course Code	Course Name	Category	L	T	P	Credit
MDA23P5	DATA ANALYTICS LAB III	Practical	-	-	75	3

Preamble

This course provides implementation of the Hadoop components like Hive and Spark. This course also provides various exercises to implement the components in the distributed environment through map reduce programming.

Prerequisite

- Big data framework
- No SQL concepts

Course Learning Outcomes

On successful completion of the course, the students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand Hadoop components for big data processing	**K2**
CLO2	Apply simple processing operations in Spark	**K3**
CLO3	Analyze the data concepts to transfer various file formats into Hive for Processing	**K4**
CLO4	Develop Spark SQL for data processing	**K5**
CLO5	Create Hive commands for big data applications using Hadoop	**K6**

Mapping with Programme Learning Outcomes

CLOs/ PLOs	PLO1	PLO2	PLO3	PLO4
CLO1	S	S	S	M
CLO2	S	S	S	M
CLO3	S	S	S	M
CLO4	S	S	S	M
CLO5	S	S	S	M

S-Strong; M-Medium

Syllabus

- Exercises to implement HiveQL to sort, order, group, distribute and cluster.
- Exercises to implement partitioning and bucketing in Hive.
- Exercises to create joins, views and indexes in Hive.
- Exercises to transfer the contents of XML, JSON and ORC files into Hive for processing.
- Exercises to implement simple processing tasks in Spark.
- Exercises to implement basic operations in Spark SQL.
- Exercise to create a RDD and perform simple basic operations.
- Exercise to implement spark streaming using Scala.
- Exercise to Broadcast the variable using Scala.

Pedagogy: Demonstrations

Course Designer:

1. Mrs.G.Anitha
2. Dr.N.Radha

Course Code	Course Name	Category	L	T	P	Credit
MDA23P6	DATA MINING AND VISUALIZATION LAB	Practical	-	-	75	3

Preamble

This course provides exercises to implement data mining techniques such as classification, clustering, association rule mining, text mining using R and also introduces data visualization using Tableau and Power BI

Prerequisite

- Data Mining
- Data Visualization concepts

Course Learning Outcomes

On successful completion of the course, students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand the concepts of association rule mining, classification, clustering, Prediction algorithm using R	**K2**
CLO2	Apply data mining techniques to real world problem	**K3**
CLO3	Analyze data using data visualization tool and provide interpretation	**K4**
CLO4	Evaluate the features of various data mining and visualization tools using Tableau and Power BI	**K5**
CLO5	Develop algorithm-based solutions and visualizations for the given problems	**K6**

Mapping with Programme Learning Outcomes

CLOs/ PLOs	PLO1	PLO2	PLO3	PLO4
CLO1	S	S	M	M
CLO2	S	S	M	M
CLO3	S	S	M	M
CLO4	S	S	S	M
CLO5	S	S	S	M

S-Strong;M-Medium

Syllabus

- Exercise to implement the concept of classification in R
- Exercise to implement the concept of clustering in R
- Exercise to find associated items in dataset using R.
- Exercise to perform text classification using the movie review dataset in R.
- Exercise to create a corpus of documents and preprocess them in R using stemming, stop word removal, whitespace removal, convert them to lowercase and remove punctuations.

- Exercise to create a term document matrix for a corpus in R.
- Exercise to find the frequent terms in a document and remove sparse terms in R.
- Exercises to visualize data using Bar chart, Line chart, Pie chart, Scatterplot and Histogram in tableau & Power BI.
- Exercises to create Dashboard, analytics report for a dataset in Power BI.
- Exercises to create Story by combining worksheets/dashboards in Tableau.

Pedagogy: Lectures**,**Demonstrations

Course Designer:

1. Dr. S. Meera
2. Mrs.G.Anitha

ELECTIVES

Course Code	Course Name	Category	L	T	P	Credit
MDA23E4	**IOT ANALYTICS**	**Theory**	**58**	**2**	**-**	**4**

Preamble

The course covers basic concepts for IOT Analytics, collection of data for IOT, Integration of IOT with Cloud, Big Data Environments. Students can learn about applying geospatial analytics and applying machine learning to the IOT data. The course also covers the organization of the IOT data, cost benefits of using IOT and review of IOT invarious sectors.

Prerequisite

- Foundations of Data Science
- Basics of IOT

Course Learning Outcomes

On the successful completion of the course, students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand IOT Data Analytics, life cycle, IOT based geo spatial analytics and machine learning application in IOT	**K2**
CLO2	Apply IOT concepts in Geo spatial analytics and Machine learning	**K3**
CLO3	Examine concepts of cloud based IOT, Big data and IOT in various domains	**K4**
CLO4	Appraise techniques and strategies for data collection with reference to big data	**K5**
CLO5	Propose strategies for organization of IOT data and optimize cost benefits in using IOT data.	**K6**

Mapping with Programme Learning Outcomes

CLOs/ PLOs	PLO1	PLO2	PLO3	PLO4
CLO1	S	M	M	M
CLO2	S	S	M	M
CLO3	S	S	S	M
CLO4	S	S	S	M
CLO5	S	S	S	M

S-Strong;M-Medium

Syllabus

UNIT I **(12 Hrs)**

Introducing IOT Analytics : Introduction – IOT Data and Big Data – Challenges of IOT analytics Applications – **IOT analytics Lifecycle and Techniques**.IOT Cloud and Big Data Integration –Cloud based IOT platform–Data Analytics for IOT–Data Collection–**WAZIUP software Platform – Ikaas Software Platform**. Searching the Internet of things: Introduction– **Search architectures for social and physical sensors** – local Event Retrieval – Use of Sensor meta datastreams.

UNIT II: **(12 Hrs)**

IOT Devices and Networking Protocols: IOT devices in different domains.IOT Analytics for the Cloud – Building elastic analytics – elastic analytics concepts – designing for scale – **Cloud security and analytics** – AWS overview - **AWS key services for IOT analytics**. Thingworx overview .**Creating an AWS Cloud Analytics environment.**

UNIT III **(12 Hrs)**

Strategies and Techniques in Data collection :Designing data processing for analytics – Applyingbig data to storage – **Apache Spark for IOT data processing**. Exploring IOT Data: **Explore andvisualizedata** – Tableau– Attributeidentification–Using R for statistical analysis – Introduction to R packages – Data Import and Data Manipulation-Statistical Inference – Analysis of Variance (ANOVA) - Solving industry specificproblems

UNIT IV **(11 Hrs)**

Geospatial Analytics to IOT Data: Basics–Vector and Raster based methods–Storing geospatial data - **Processing geospatial data.** Data Science for IOT analytics – Machine learning basic – **Forecasting IOT data using ARIMA –** Analyzing time series IOT data using PSPP **- Deep learning with IOT data.**

UNIT V **(11 Hrs)**

Organize IOT data – Linked analytics datasets – Managing data lakes – **data retention strategy for IOT data**. Economics of IOT data – Cloud computing and open source – cost considerations –Revenue – **Predictive maintenance**.IOT review : IOT data flow – **IOT exploratory analytics** –IOT data science– Building revenue – Sample project.

Use Cases: Real time data analysis for manufacturing sector, IOT analytics for healthcare gamechanger.

Text Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Andrew Minteer	Analytics for the Internet of things	Packt Publishing	2017, 1st Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Prasant Kumar Pattnaik, Raghvendra Kumar, Souvik Pal, S. N. Panda	IOT and Analytics in Agriculture	Springer	2020, 1st Edition
2	John Soldatos	Building blocks for IOT Analytics: Internet-of-Things Analytics	River Publisher Series in Signal Image and Speech Processing	2017, 1st Edition

Pedagogy: Lectures, Demonstrations, Group Discussions, Case studies

Course Designers:

1. Mrs. K. Gandhimathi
2. Dr. N. Radha

Course Code	Course Name	Category	L	T	P	Credit
MDA23E5	SOCIAL MEDIA ANALYTICS	Theory	58	2	-	4

Preamble

The course covers concepts and techniques for retrieving, exploring, visualizing, and analyzing social network and social media data. Students learn the key metrics to assess social media goals, perform social network analysis to apply social media analytics process and formulate effective strategies based on the analytics.

Prerequisite

- Foundations of Data Science

Course Learning Outcomes

On the successful completion of the course, students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand sources and limitations of social media data	**K2**
CLO2	Apply social media analytics process and evaluate metrics	**K3**
CLO3	Examine different social media platforms and their associated tools	**K4**
CLO4	Apply social media information to create dashboards and reports for visualization	**K5**
CLO5	Design effective strategy based on the social media analytics data	**K6**

Mapping with Programme Learning Outcomes

CLOs/ PLOs	PLO1	PLO2	PLO3	PLO4
CLO1	S	M	M	M
CLO2	S	S	M	M
CLO3	S	M	S	M
CLO4	S	S	S	M
CLO5	S	S	M	M

S-Strong;M-Medium

Syllabus

UNIT I **(12 Hrs)**

Foundation for Analytics:–Digital Gap–Social Media Data Sources–Defining Social Media Data – **Data Sources**–Estimated vs Factual Data Sources–**Data Gathering in Social Media Analytics.** From Data to Insights : Actionable Analytics – Focus on objective – Plan to shape data to insights– Choosing a good analytics tool–Data Aggregation calculations and display–Data display – Social media and Big data – Potential Challenges. **Data Identification: Professional networking sites-social sites – information sharing sites –micro blogging sites –blogs/wikis.**

UNIT II **(12 Hrs)**

Analytics in social Media: Types of analytics. Dedicated Vs. Hybrid Tools – Dedicated tools –Hybrid tools – Data Integration Tools – Best Setup. **Social Network Landscape : Concept and UX on social**

networks – Interactivity of social network – Content flow on social network – **Interaction Pattern between users–Social Media as a twoway channel.**

UNIT III **(11 Hrs)**

Analytics Process: Analysis – Insight – Investigation beyond social analytics – Shaping a method – analysis cycle – Community Activity – Resources – Attention span – **Dynamic cycles – Short Periods – Long Periods** – Analyst Mindset – Instinctive Analyst. Metrics: Introduction – **Defaultand custom metrics** – Metrics Categories – **Graph Types – Metric Capabilities** – Metrics andStrategy–Estimated Metrics– Metrics and Tactics.

UNIT IV **(12 Hrs)**

Dashboards: Purpose and Objectives – **Default Vs. Custom Dashboards** – Linearity and order of metrics–Metrics Positioning and Correlation–Metric and dashboard layout–Graphic design–Data Integration dashboards. Reports: Elements of reporting**–Reporting approaches and formats –Animation and effects in reporting –Stakeholders and feedback–Reporting with teams.**

UNIT V **(11 Hrs)**

Strategy: Strategy in social media analytics – Strategic planning – **Data availability and data sources**– Knowledge beyond social media–Tools and technology preparation–**Team Preparation** – **Goals and objectives** – Contingency plans – **application of social media analytics strategy**–Strategy and tactics – Evaluation of a strategic analytics cycle.

Case Studies: Targeting the audience using Facebook Analytics, Tracking profile analytics in LinkedIn, Analysis of Political Tweets, ROI Analytics using Facebook, Marketing Strategy in Pinterest.

Text Book

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Alex Goncalves	Social Media Analytics Strategy-Using Data to Optimize Business Performance.	Apress	2017, 1st Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Ganis,Kohirkar	Social media Analytics	IBM Press PTG.	2016, 1stEdition
2	Nancy Flynn	The Social Media Handbook Policies and Best Practices	Wiley	2012, 1st Edition

Pedagogy: Lectures, Demonstrations, Group Discussions, Case studies

Course Designers:

1. Dr.N.Radha
2. Mrs.G.Anitha

Course Code	Course Name	Category	L	T	P	Credit
MDA23E6	WEB DATA ANALYTICS	Theory	58	2	-	4

Preamble

This course gives insights about leveraging web data to achieve strategic business objectives. It deals with the various techniques for analyzing web data like click stream analysis. The course also provides ways to execute competitive intelligence analysis and to analyze emerging social, mobile and video data.

Prerequisite

- Foundations of Data Science
- Information Retrieval

Course Learning Outcomes

On the successful completion of the course, students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand the techniques of web data analytics	K2
CLO2	Apply web data analytics on social, mobile and video data	K3
CLO3	Analyze techniques for measuring the success of a website	K4
CLO4	Assess the various cases to apply web data analytics	K5
CLO5	Propose new metrics based solutions for user websites	K6

Mapping with Programme Learning Outcomes

CLOs/ PLOs	PLO1	PLO2	PLO3	PLO4
CLO1	S	M	M	M
CLO2	S	S	M	M
CLO3	S	M	S	M
CLO4	S	S	S	M
CLO5	S	S	M	M

S-Strong;M-Medium

Syllabus

UNIT I **(11 Hrs)**

Introduction :**Web Analytics 2.0** - Clickstream- multiple outcome analysis-experimentation and testing-voice of customer–competitive intelligence-the tactical shift-**Optimal strategy for choosing web analytics**

UNIT II **(11 Hrs)**

Clickstream analysis: Metrics-**Eight critical web metrics**-web metrics demystified –strategically aligned tactics for impactful web–Web analytics report-**Foundational analytical strategies**-clickstream analysis made actionable-challenges

UNIT III **(12 Hrs)**

Measuring Success-Actionable Outcome KPIs-Moving beyond conversion rates-Micro and macro

conversion- **Measuring success for a non–ecommerce website**-Leveraging qualitative data:Surveys-Web enabled emerging user research options

UNIT IV **(12 Hrs)**

A/B Testing - Multivariate testing-Actionable testing ideas-Controlled experiments-Competitive intelligence analysis-CI data source, types, secrets- website traffic analysis-**Search and keyword analysis**-audience identification and segmentation analysis

UNIT V **(12 Hrs)**

Emerging analytics: Social. mobile, video: Measuring social web - the data challenge- analyzing mobile customer experiences-**measuring the success of blogs**- quantifying the impact of Twitter – **Analyzing the performance of videos.**

Exploring Web Analytical Tools: Google Analytics, Bitly, Open Web Analytics, Similar Web, SEMrush, Talkwalker Platform.

Text Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Avinash Kaushik	Web Analytics 2.0:The Art of Online Accountability and Science of Customer Centricity	Wiley Publishing	2010, 1st Edition
2	Dietmar Jannach, Markus Zanker	Recommender system-An introduction	Cambridge University Press	2011, 1st Edition
3	Bing Liu	Sentiment Analysis and opinion mining	Morgan and claypool Publishing	2012, 1st Edition

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Eric Enge, Stephan Spencer, Jessie Stricchiola	The Art of SEO:Mastering Search Engine Optimization	Oreilly Publishing	2015, 3rdEdition
2	Kristina Halvors	Content Strategy for the Web	Melissa Rach New Riders	2012, 1st Edition
3	Clifton B	Advanced Web Metrics with Google Analytics	Wiley Publishing,Inc	2012, 2nd Edition
4	Sterne J	Web Metrics: Proven methods for measuring website success	John WileyandSons	2002

Pedagogy: Lectures, Demonstrations, Group Discussions, Online References

Course Designers:

1. Mrs.G.Anitha
2. Ms. M. Nandhini

Job Oriented Course – M.Sc Data Analytics

Title: Microsoft Power BI Data Analyst (PL-300)

Duration :60 Hrs

Topic	Details
Prepare the Data (25-30%)	
Get data from data sources	- Identify and connect to a data source - Change data source settings, including credentials, privacy levels, and data source locations - Select a shared dataset, or create a local dataset - Choose between Direct Query, Import, and Dual mode - Change the value in a parameter
Clean the data	- Evaluate data, including data statistics and column properties - Resolve inconsistencies, unexpected or null values, and data quality issues - Resolve data import errors
Transform and load the data	- Select appropriate column data types - Create and transform columns - Transform a query - Design a star schema that contains facts and dimensions - Identify when to use reference or duplicate queries and the resulting impact - Merge and append queries - Identify and create appropriate keys for relationships - Configure data loading for queries
Model the Data (25-30%)	
Design and implement a data model	- Configure table and column properties - Implement role-playing dimensions - Define a relationship's cardinality and cross-filter direction - Create a common date table - Implement row-level security roles
Create model calculations by using DAX	- Create single aggregation measures - Use CALCULATE to manipulate filters - Implement time intelligence measures - Identify implicit measures and replace with explicit measures - Use basic statistical functions - Create semi-additive measures - Create a measure by using quick measures - Create calculated tables
Optimize model performance	- Improve performance by identifying and removing unnecessary rows and columns - Identify poorly performing measures, relationships, and visuals by using Performance Analyzer - Improve performance by choosing optimal data types - Improve performance by summarizing data

Topic	Details
Visualize and Analyze the Data (25-30%)	
Create reports	- Identify and implement appropriate visualizations - Format and configure visualizations - Use a custom visual - Apply and customize a theme - Configure conditional formatting - Apply slicing and filtering - Configure the report page - Use the Analyze in Excel feature - Choose when to use a paginated report
Enhance reports for usability and story telling	- Configure bookmarks - Create custom tooltips - Edit and configure interactions between visuals - Configure navigation for a report - Apply sorting - Configure Sync Slicers - Group and layer visuals by using the selection pane - Drill down into data using interactive visuals - Configure export of report content, and perform an export - Design reports for mobile devices - Incorporate the Q&A feature in a report
Identify patterns and trends	- Use the Analyze feature in Power BI - Use grouping, binning, and clustering - Use AI visuals - Use reference lines, error bars, and forecasting - Detect outliers and anomalies - Create and share scorecards and metrics
Deploy and Maintain Assets (15-20%)	
Create and manage workspaces and assets	- Create and configure a workspace - Assign workspace roles - Configure and update a workspace app - Publish, import, or update assets in a workspace - Create dashboards - Choose a distribution method - Apply sensitivity labels to workspace content - Configure subscriptions and data alerts - Promote or certify Power BI content - Manage global options for files
Manage datasets	- Identify when a gateway is required - Configure a dataset scheduled refresh - Configure row-level security group membership - Provide access to datasets

Course Code	Course Name	Category	L	T	P	Credit
MNM22CS2	CYBER SECURITY - II	Theory	-	-	-	Grade

Preamble

This course provides the classification of cyber security and cyber crime and its laws and data privacy and security in social media

Prerequisite

Basics of Internet

Course Learning Outcomes

On successful completion of the course, the students will be able to

CLO Number	CLO Statement	Knowledge Level
CLO1	Understand the basic concepts of cybersecurity and Cybersecurity threat landscape	K2
CLO2	Apply the methods to identify the cyber-attacks and crimes.	K3
CLO3	Analyze the legal framework that exist in India for cybercrime and legal frame work followed by other countries	K4
CLO4	Estimate the data privacy and security issues related to personal data privacy and security	K5
CLO5	Create a privacy settings on social media platform and register complaints on a social media platform.	K6

UNIT I

Overview of cyber security: Cyber security terminologies- Cyberspace- Cyber-attack- Cyber threats - Cyber terrorism – Cyber warfare.

UNIT II

Cyber crimes : Cyber Crimes targeting computer system and mobiles- Online scam frauds: emails Scams- Phishing- Vishing- Smishing- Online job fraud- online sextortion- Debit and credit card fraud- Online payment fraud- cyberbullying. Social Media Scam & Fraud: Impersonation- Identify theft -Job scams- Misinformation - Fake new cyber crime against persons -Cyber grooming -Child pornography - cyber stalking-Cyber police station -Crime reporting produce.

UNIT III

Cyber law: Cyber laws and legal and ethical aspects related to new technologies: AI/ML-IoT-Block chain- Darknet and social media- Cyber law of other countries.

UNIT IV

Data privacy and Data security: Defining data- Metro-Big data- Non personal data- Data protection- media platforms, Registering complaints on social media platforms, prepare password policy for computer and mobile device, security controls for computer and mobile phones , digital Forenics, Cyber Bulling, Phishing, Facebook Attack, Cyber Security audit and Compliance and National Security Policies. General Data Protection Regulations (GDPR)- 2016 Personal Information Protection and the Electronic document Act(PIPEDA)- Social media Data privacy and Security issues.

UNIT V

Case Studies: Case Study on Platform for reporting Cyber Crimes, Checklist for reporting cybercrimes online, Setting privacy settings on social media platforms, Do's and Don'ts for posting content on social

Reference Books

S.No	Name of the Authors	Title of the Book	Publishers	Year & Edition
1	Sumit Belapure and Nina God Bole	Cyber security understanding cyber crimes computer forensics and Legal perspectives	Wiley India Pvt Ltd	2011, 1st Edition
2	Dorothy F. Denning	Information warfare and security	Addison Wesley	1998
3	Henry A. Oliver	Security in the digital age: social media security threads and vulnerabilities	Create Space Independence publishing platform	2015, 1st Edition