



**Dr. G. Y. Pathrikar College of
Computer Science and Information Technology**
Chhatrapati Sambhajinagar

**M.Sc. Data Science
First & Second Year
(Syllabus)**

MGM University

Vision

- To ensure sustainable human development which encourages self-reliant and self-content society.
- To promote activities related to community services, social welfare and also Indian heritage and culture.
- To inculcate the culture of non-violence and truthfulness through vipassanna meditation and Gandhian Philosophy.
- To develop the culture of simple living and high thinking

Mission

- To impart state of art education and technical expertise to students and give necessary training to teachers to create self-reliant society for future.
- To encourage students to participate in Indian and International activities in sports, literature, etc. so that future generation becomes base for free and liberal society
- To educate students in areas like Management, Finance, Human relations to inculcate philosophy of simple living and high thinking value of simple economic society.
- To inculcate culture of non-violence and truthfulness through Vipassana.

To sustain activities of Indian culture (viz. classical dance, music and fine arts) through establishing institutes like Mahagami, Naturopathy, etc.

विद्यापीठगीत

अत्तदिपभवभवप्रदिपभव,
 स्वरूपरूपभवहो
 ज्ञानसब्बविज्ञानसब्बभव ,
 सब्बदिपभवहो
 अत्ताहिअत्तनोनाथो ,
 अत्ताहिअत्तनोगति
 अत्तमार्गपरअप्रमादसेहैतुझेचलना
 सब्बकाकल्याणहो ,
 वोकार्यकुशलकरना
 सब्बकाउत्तममंगल , पथप्रदर्शकहो
 अत्तदिपभवभवप्रदिपभव ,
 स्वरूपरूपभवहो
 ज्ञानसब्बविज्ञानसब्बभव ,
 सब्बदिपभवहो
 बुद्धमंशरनंगच्छामि :
 धम्मंशरनंगच्छामि :
 संघंशरनंगच्छामि :

Dr. G. Y. Pathrikar College of Computer Science & Information Technology

MGM college of Computer Science and Information Technology was established in 2001 offering undergraduate and postgraduate degree program in Computer Science and Information Technology. College was renamed as Dr.G.Y.Pathrikar College of Computer Science and Information Technology in 2003 in memory of great educationalist, one of the founder member and Ex-Secretary MGM, Dr.G.Y.Pathrikar Sir.

It is first self-financed ISO certified institution offering program dedicated to Computer science and Information technology in Maharashtra and has achieved status of 2f/12b. Ours was the only and first college to be re-accredited as A+ grade with NAAC in the year 2017. Experienced and qualified faculty with Ph.D is strength of our college. Starting with 77 student's College has crossed total students strength of 10,000 passing out. Student are doing well in various MNCs like Infosys, Tech-Mahindra, Wipro, Capgemini, Cognizant etc. Many have their own Startups. Some of the students have completed their Masters and Ph.D. program from foreign countries like US, UK, Australia. Now we are constituent college of MGM University, Chhatrapati Sambhajanagar.

Vision

To be an academic institution in dynamic equilibrium in social, ecological and economical environment striving continuously for excellence in total quality education, research and technological service to the nation.

Mission

- To create and sustain a community of learning in which students acquire knowledge and learn to apply it professionally with due consideration for ethical, and economical issues.
- To upgrade our students in all respect with the help of latest infrastructure in the area of Computer Science and Information Technology in order to build the National Capabilities.
- To understand the culture of Non-violance, truth, peace through Gandhian Philosophy.

Programs offered at Dr. G. Y. Pathrikar College of Computer Science & Information Technology

Undergraduate Programmes	Postgraduate Programmes	PhD Programmes
B.Sc. (Computer Science) Honours / Honours with Research	M.Sc(Computer Science)	Ph.D. in Computer Science and Information Technology
B.Sc(Information Technology) Honours/ Honours with Research	M.Sc(Information Technology)	
BCA(Science) Honours / Honours with Research	M.Sc(Data Science)	
B.Sc(Animation) Honours / Honours with Research	M.Sc(Animation)	
Integrated M.Sc. Data Science		
BCA(Digital Marketing) Honours		
B.Sc(Robotics) Honours		

Name of Program – M.Sc. Data Science

Duration – Two Years

Eligibility –Any Bachelor degree of minimum 3 years with at least two papers in Mathematics / Information Technology / Computer Science /Computer Application / Statistics / Electronics / Operational Research at 10+2 level or at Graduation level examination.

He / She must have obtained at least 50% marks in aggregate (45% in case of candidates of reserved categories) from any recognized University.

First Year- Semester I												
Course Category	Course Code	Course Title	Nature of Course	No. of Credits	Teaching (Contact hrs/ week)		Evaluation Scheme (Marks)			Minimum Passing (Marks)		
					L	P	Internal	External	Total	Internal	External	Total
Major	MDS41M ML501	Introduction to Data Science	Lecture	3	3		60	40	100	-	16	40
Major	MDS41M ML502	Data Structure and Algorithm	Lecture	3	3		60	40	100	-	16	40
Major	MDS41M ML503	Python Programming	Lecture	3	3		60	40	100	-	16	40
Major	MDS41R ML501	Research Methodology	Lecture	4	4		60	40	100	-	16	40
Elective		Elective from Basket-1	Lecture	3	3		60	40	100	-	16	40
Major	MDS41M MP501	Practical Based on Introduction to Data Science	Practical	1	-	2	30	20	50	-	8	20
Major	MDS41M MP502	Practical Based on Data Structure and Algorithm	Practical	1	-	2	30	20	50	-	8	20
Major	MDS41M MP503	Practical Based on Python Programming	Practical	1	-	2	30	20	50	-	8	20
Elective		Practical based on Elective from Basket-1	Practical	1	-	2	30	20	50	-	8	20
Total				20	16	8	420	280	700	-		

First Year- Semester II												
Course Category	Course Code	Course Title	Nature of Course	No. of Credits	Teaching (Contact hrs/ week)		Evaluation Scheme (Marks)			Minimum Passing (Marks)		
					L	P	Internal	External	Total	Internal	External	Total
Major	MDS41M ML504	Advances in Database Management System	Lecture	3	3		60	40	100	-	16	40
Major	MDS41M ML505	Data Mining and Visualization	Lecture	3	3		60	40	100	-	16	40
Major	MDS41M ML506	Machine Learning	Lecture	3	3		60	40	100	-	16	40
Elective		Elective from Basket-2	Lecture	3	3		60	40	100	-	16	40
Major	MDS41M MP504	Practical Based on Advances in Database Management System	Practical	1	-	2	30	20	50	-	8	20
Major	MDS41M MP505	Practical Based on Data Mining and Visualization	Practical	1	-	2	30	20	50	-	8	20
Major	MDS41M MP506	Practical Based on Machine Learning	Practical	1	-	2	30	20	50	-	8	20
Elective		Practical based on Elective from Basket-2	Practical	1	-	2	30	20	50	-	8	20
OJT	MDS41JT J501	On Job Training / Internship	Project	4		8	60	40	100	-	--	40
Total				20	12	16	420	280	700			

Note:

Nature of Course : L- Lecture, P-Practical, S-Seminar, J-Project, I-Internship, D-Dissertation,

Course Category: MM-Major Mandatory, ME-Major Elective, MI-Minor, OE-Generic / Open electives, VSC-Vocational skill course, SEC-Skill Enhancement course, AEC-Ability Enhancement course, IKS-Indian Knowledge system, VEC-Value Education course, OJT-On Job Training / Internship / Apprenticeship, FP-Field project, CEP-Community engagement and service, CC-Co – curricular course, RM-Research methodology, RP-Research project.

Semester III

Second Year- Semester III												
Course Category	Course Code	Course Title	Nature of Course	No. of Credits	Teaching (Contact hrs/ week)		Evaluation Scheme (Marks)			Minimum Passing (Marks)		
					L	P	Internal	External	Total	Internal	External	Total
Major	MDS41M ML601	Big Data	Lecture	3	3	-	60	40	100	-	16	40
Major	MDS41M ML602	Cloud Computing	Lecture	3	3	-	60	40	100	-	16	40
Major	MDS41M ML603	Natural Language Processing	Lecture	3	3	-	60	40	100	-	16	40
Elective		Elective from Basket-3	Lecture	3	3	-	60	40	100	-	16	40
Major	MDS41M MP601	Practical Based on Big Data	Practical	1	-	2	30	20	50	-	08	20
Major	MDS41M MP602	Practical Based on Cloud Computing	Practical	1	-	2	30	20	50	-	08	20
Major	MDS41M MP603	Practical Based on Natural Language Processing	Practical	1	-	2	30	20	50	-	08	20
Elective		Practical based on elective 3	Practical	1	-	2	30	20	50	-	08	20
RP	MDS41R PJ601	Research Project	Project	4	-	8	60	40	100	-	16	40
Total				20	12	16	420	280	700			

Nature of Course : L- Lecture, P-Practical, S-Seminar, J-Project, I-Internship, D-Dissertation,

Course Category: MM-Major Mandatory, ME-Major Elective, MI-Minor, OE-Generic / Open electives, VSC-Vocational skill course, SEC-Skill Enhancement course, AEC-Ability Enhancement course, IKS-Indian Knowledge system, VEC-Value Education course, OJT-On Job Training / Internship / Apprenticeship, FP-Field project, CEP-Community engagement and service, CC-Co – curricular course, RM-Research methodology, RP-Research project.

Semester IV

Second Year- Semester IV												
Course Category	Course Code	Course Title	Nature of Course	No. of Credits	Teaching (Contact hrs/ week)		Evaluation Scheme (Marks)			Minimum Passing (Marks)		
					L	P	Internal	External	Total	Internal	External	Total
MM	MDS41MM L604	Artificial Intelligence	Lecture	3	3	-	60	40	100	--	16	40
MM	MDS41MM L605	Development Operations	Lecture	3	3	-	60	40	100	--	16	40
MM	MDS41MM L606	Recommendation System	Lecture	2	2	-	30	20	50	--	08	20
ME		Elective from Basket-4	Lecture	3	3	-	60	40	100	--	16	40
ME	MDS41MM P604	Practical Based on Artificial Intelligence	Practical	1	-	2	30	20	50	--	08	40
MM	MDS41MM P605	Practical Based on Development Operations	Practical	1	-	2	30	20	50	--	08	20
		Elective from Basket-4	Practical	1	-	2	30	20	50	--	08	20
RP	MDS41RPJ 602	Research Project	Project	6		12	120	80	200	--	32	80
Total				20	11	18	420	280	700			

Nature of Course : L- Lecture, P-Practical, S-Seminar, J-Project, I-Internship, D-Dissertation,

Course Category: MM-Major Mandatory, ME-Major Elective, MI-Minor, OE-Generic / Open electives, VSC-Vocational skill course, SEC-Skill Enhancement course, AEC-Ability Enhancement course, IKS-Indian Knowledge system, VEC-Value Education course, OJT-On Job Training / Internship / Apprenticeship, FP-Field project, CEP-Community engagement and service, CC-Co – curricular course, RM-Research methodology, RP-Research project.

Basket I - (Elective)

Code	Title
MDS41MEL501	Mathematics and Statistics
MDS41MEL502	Data Warehousing
MDS41MEL503	Compiler Design
MDS41MEL504	Sentiment Analysis
MDS41MEP501	Practical Based on Mathematics and Statistics
MDS41MEP502	Practical Based on Data Warehousing
MDS41MEP503	Practical Based on Compiler Design
MDS41MEP504	Practical Based on Sentiment Analysis

Basket 2- (Elective)

Code	Title
MDS41MEL505	Google Analytics
MDS41MEL506	Neural Network
MDS41MEL507	Optimization Techniques
MDS41MEL508	Business Intelligence
MDS41MEP505	Practical Based on Google Analytics
MDS41MEP506	Practical Based on Neural Network
MDS41MEP507	Practical Based on Optimization Techniques
MDS41MEP508	Practical Based on Business Intelligence

Basket 3-(Elective)

Code	Title
MDS41MEL601	Cloud Services
MDS41MEL602	Data Analytics
MDS41MEL603	Internet Of Things
MDS41MEL604	Block chain
MDS41MEP601	Practical Based on Cloud Services
MDS41MEP602	Practical Based on Data Analytics
MDS41MEP603	Practical Based on Internet Of Things
MDS41MEP604	Practical Based on Block chain

- Basket 4- (Elective)**

Code	Title
MDS41MEL605	Exploratory Data Analysis
MDS41MEL606	Data Centre Technologies
MDS41MEL607	Social Media Analytics
MDS41MEL608	Deep Learning
MDS41MEP605	Practical Based on Exploratory Data Analysis
MDS41MEP606	Practical Based on Data Centre Technologies
MDS41MEP607	Practical Based on Social Media Analytics
MDS41MEP608	Practical Based on Deep Learning

Semester: FIRST

Semester-I

Course code: MDS41MML501	Course name: Introduction to Data Science
Course category: Major Mandatory	Credits: 3
Teaching scheme: L-3 P-0	Evaluation scheme: CA–60, ESE–40
Pre-requisites: Mathematical background and programming skills sufficient enough to learn new languages and software are required. Basic knowledge of statistics.	
Course Objectives:	
1. Understand and Learn the lifecycle and phases of data science and work comfortably with data science projects	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: To develop fundamental knowledge of concepts underlying data science	
CO2: To develop practical data analysis skills, which can be applied to practical problems	
CO3: To explain how math and information sciences can contribute to building better algorithms and software	
CO4: To develop applied experience with data science software, programming, applications and processes	

Contents –

Unit	Content	Teaching hours
1	Introduction to Data Science: Definition and applications, Skills needed basic steps in Data Science, Benefits and uses of data science and big data. Facets of Data: Structured data, unstructured data, Natural Language, Machine generated data; Graph based or network data, Audio, Image and Video, Streaming data, Data Science Process: Overview of data science process. Basics of retrieving data, data cleansing, transforming data, data modeling.	10
2	Steps for Data Science Process: Step 1: Setting the research goal, Step 2: Retrieving data, Step 3: Cleansing, Integrating and transforming data. Step 4: Exploratory data analysis, Step 5: Build the model, Step 6: Presentation and automation. Sentiment analysis and its applications: Definition, types of sentiments, applications, case study. Uses of Text mining.	10
3	Data with R Studio: Using an integrated development environment, installing R studio, Creating R Script, History, overview. Data types variable, operators, and string. Syntax of R objects. R functions: Statistical functions, Built in functions, User defined functions.	10
4	Objects in R Language: Creating and manipulation of R Objects Vector, List, Matrices, Arrays, Factors, Data frames, Data Interfaces: CSV file, Excel file, Text file, Charts and Graphs: Pie chart, Bar chart, Boxplot, Histogram, Advanced visualization techniques.	10
5	Machine learning: Concept, application, Supervised and Unsupervised Learning, Supervised learning via Support Vector Learning.	5

Text Books: 1. Jeffrey S.Saltz,Jeffrey M.Stanton Introduction to Data Science Ebook SAGE Publications
2. Galit Shmueli, Peter C. Bruce, Inbal Yahav, Nitin R. Patel, Kenneth C. Lichtendahl Jr.,Data Mining for Business Analytics: Concepts, Techniques and Applications in R, Wiley India, 2018.
Reference Books: 1. Rachel Schutt & Cathy O’Neil Doing Data Science, O’ Reilly, First Edition, 2013
2. B. Ram Computer Fundamental, BPB Publication
Online Resources: 1. NPTEL / SWAYAM lectures.

Syllabus **Semester-I**

Course code: MDS41MML502	Course name: Data Structure and Algorithm
Course category: Major Mandatory	Credits: 3
Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Programming language concepts, discrete mathematical structure.	
Course Objectives: This course provides an introduction to mathematical modeling of computational problems. It covers the common algorithms, algorithmic paradigms, and data structures used to solve these Problems. The course emphasizes the relationship between algorithms and programming, and Introduces basic performance measures and analysis techniques for these problems.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: To impart the basic concepts of data structures and algorithms	
CO2: To understand concepts about searching and sorting techniques	
CO3: To understand basic concepts about stacks, queues, lists, trees and graphs	
CO4: To understanding about writing algorithms and step by step approach in solving problems with the help of fundamental data structures	

Contents –

Unit	Content	Teaching hours
1	Overview: Introduction to Algorithm, Analysis of algorithm, Designing of algorithm, the Correctness of Algorithms and the Complexity of Algorithms	10
2	Linear Data Structures: Stack, Queue, Array, Linked list, Priority Queue, Deque, Doubly linked list, circular linked list, Searching and sorting Techniques	10
3	Non Linear Data Structure: Graphs: Introduction to Graph Theory, Graph isomorphism, Graph data structures: Adjacency lists, Adjacency matrices Elementary graph Algorithms: BFS, DFS, Topological sort, strongly connected, Components Trees: Introduction to Trees, Tree traversals (preorder, inorder and postorder), Binary trees, Balanced trees: Avl etc., B and B+ tree Application of trees, Minimum Spanning Trees, Single source shortest path, All pair shortest path.	10
4	Strings: The string abstract data type, Brute force string pattern matching, regular expression pattern matching, finite automata, Hashing : Hash function, collision resolution, Heap	10
5	Dynamic programming and greedy algorithms NP vs P: The spaces P and NP, polynomial reduction, NP complete problem	5

Text Books: 1. Thomas Cormen Introduction to Algorithms O'Really 2. Galit Shmueli, Peter C. Bruce, Inbal Yahav, Nitin R. Patel, Kenneth C. Lichtendahl Jr., Data Mining for Business Analytics: Concepts, Techniques and Applications in R, Wiley India, 2018.
Reference Books: 1. Ellis Horowitz, .Fundamentals of Data Structures in C++", O' Reilly
Online Resources: 1. NPTEL / SWAYAM lectures.

Syllabus Semester-I

Course code: MDS41MML503	Course name: Python Programming
Course category: Major Mandatory	Credits: 3
Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basic computer programming knowledge is essential	
Course Objectives: Understand and use variables. Work with common Python data types like integers, floats, strings, characters, lists, dictionaries, as well as pandas DataFrames. Use basic flow control, including for loops and conditionals. read data from text files. Obtain basic summary statistics from data files.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: To understand importance of Python for data science	
CO2: Develop a skill to implement Python Programming for data science.	
CO3: Hands-on Python experience for professional advancement	
CO4: To use and implement standard programming constructs like data structure, numerical computing.	

Contents –

Unit	Content	Teaching hours
1	<p>Introduction to Python:</p> <p>Getting Started: Introduction to Python- an interpreted high level language, interactive mode and script mode. Variables, Expressions and Statements. Variables and Types-mutable and Immutable variable and Keywords. Operators and Operands in Python. (Arithmetic, relational and logical operators), Operator precedence, Expressions and Statements (Assignment statement) Taking input (using raw input() and input()) and displaying output - print statement, Comments in Python.</p>	10
2	<p>Conditional and Looping Construct</p> <p>if - else statement and nested if – else while, for, use of range function in for, Nested loops break, continue, pass statement Use of compound expression in conditional constructs.</p> <p>Functions:</p> <p>Built-In Function, invoking built in functions Module(Importing entire module or selected objects using from statement) Functions from math, random, time & date module.</p> <p>Composition User Define Function:</p> <p>Defining , invoking functions, passing parameters (default parameter values, keyword arguments) Scope of variables, void functions and functions returning values</p>	10
3	<p>String :</p> <p>Creating, initializing and accessing the elements, String operators: +, *, in, not in, range, slice [n:m] String built in functions & methods: len, capitalize, find, isalnum, isalpha, isdigit, lower, islower, isupper, upper, lstrip, rstrip, isspace, istitle, partition, replace, join, split, count, decode, encode, swapcase Strings constants defined in string module.</p>	10

	Errors and Exceptions: Syntax Errors, Exceptions, Handling Exceptions, Raising Exceptions, User-defined Exceptions, Defining Clean-up Actions(try - finally), Predefined Clean-up Actions	
4	Concept of mutable list: creating, initializing and accessing the elements of list List operations (Concatenation, Repetation, Membership, list slices), List comprehensions List functions & methods: len, insert, append, extend, sort, remove, reverse, pop, Tuples, Immutable concept, creating, initializing and accessing the elements in a tuple, Tuple functions: cmp(), len(), max(), min(), tuple(), Sets , Concept of Sets: creating, initializing and accessing the elements of Sets operation(Membership, union, intersection, difference, and symmetric difference, Dictionaries: Concept of key-value pair, creating, initializing and accessing the elements in a dictionary, Traversing, appending, updating and deleting elements Dictionary functions & Methods: cmp, len, clear(), get(), has_key(), items(), keys(), update(), values()	10
5	Introduction to Object Oriented concepts in Python: Object Oriented concepts, Classes & Objects, Python Scopes and Namespaces Class Objects, Instance Objects, Method Objects, Class and Instances Variables, Concept of self Constructors, Inheritance	5

Text Books: 1. Mark Lutz's Learning Python O'Really

Reference Books: 1. Mark Lutz's Programming Python O'Really

2. Jake VanderPlas Python Data Science Handbook O' Reilly

Online Resources: 1. NPTEL / SWAYAM lectures.

<https://docs.python.org/3/tutorial/>

<https://wiki.python.org/moin/>

<https://numpy.org/devdocs/user/quickstart.html>

Syllabus Semester-I

Course code: MDS41RML501	Course name: Research Methodology
Course category: Major Mandatorys	Credits: 3
Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Novel Thinking and problem solving skills	
Course Objectives: To understand the state-of-the-art in research methodology. Survey the currently available systems.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: Demonstrate knowledge of research methodology	
CO2: Understand the Research Problem	
CO3: Understand the Research Design	
CO4: Understand Sampling Design, Measurement and Scaling Techniques	
CO4: Understand Methods of Data Collection, Processing and Analysis of Data	

Contents –

Unit	Content	Teaching hours
1	Introduction to Research Methodology: Meaning of Research, Objectives of Research, Motivation in Research, Types of Research, Research Approaches, Significance of Research, Research Methods Verses Methodology, Research and Scientific Method, Importance of Knowing How Research is Done, Research Process, Criteria of Good Research, Problems Encountered by Research in India.	10
2	Defining the Research Problem: What is Research Problem, Selecting the Problem, Necessity of Defining the Problem, Techniques Involved in Defining a Problem, An Illusion.	10
3	Research Design: Meaning of Research Design, Need of Research Design, Features of Good Design, Important Concepts Relating to Research Design, Different Research Design, Basic Principles of Experimental Designs.	10
4	Sampling Design Measurement and Scaling Techniques: Implications of a Sample Design, Steps in Sampling Design, Criteria of Selecting a Sampling Procedure, Characteristics of a Good Sample Design. Measurement in Research, Sources of Error in Measurement, Meaning and Scaling.	10
5	Methods of Data Collection, Processing and Analysis of Data: Collection of Primary Data, Observation Method, Interview Method, Collection of Data through Questionnaires, Collection of Data through Questionnaires Schedules, Processing Operations, Some Problems in Processing, Statistics in Research, Simple Regression Analysis.	5

Text Books: 1. C. R. Kothari Research Methodology Methods and Techniques, New Age International Publishers

Reference Books: 1. P. Sam Daniel, Aroma G. Sam Research Methodology Gyan Publishing House

Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-I

Course code: MDS41MMP501	Course name: Practical Based on Introduction to Data Science
Course category: Major Mandatory	Credits: 1
Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Course Objectives: Understand and learn the lifecycle and phases of data science and work comfortably with data science projects	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: To develop fundamental knowledge of concepts underlying data science	
CO2: To develop practical data analysis skills, which can be applied to practical problems	
CO3: To explain how math and information sciences can contribute to building better algorithms and software	
CO4: To develop applied experience with data science software, programming, applications and processes	

Contents –

Unit	Content	Teaching hours
1	To create records in excel using different data types.	1
2	To create student mark details by applying at least any ten statistical functions.	1
3	To Visualize data using graphs and charts	1
4	Write a Program in R to demonstrate Vector Object	1
5	Write a Program in R to demonstrate List	1
6	Write a Program in R to demonstrate Matrix	1
7	Write a Program in R to demonstrate Data Frame	1
8	Write a Program in R to demonstrate Arrays	1
9	Write a Program in R to show the results using statistical functions such as mean, mod, median, etc.,	1
10	Write a Program in R to Write data to CSV file.	1

Text Books: Jeffrey S.Saltz,Jeffrey M.Stanton Introduction to Data Science Ebook SAGE Publications

Reference Books:

Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-I

Course code: MDS41MMP502	Course name: Practical based on Data Structure and Algorithm
Course category: Major Mandatory	Credits: 1
Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Course Objectives: Learn and understand basic Data Structure operations	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: To impart the basic concepts of data structures and algorithms	
CO2: To understand concepts about searching and sorting techniques	
CO3: To understand basic concepts about stacks, queues, lists, trees and graphs	
CO4: To understanding about writing algorithms and step by step approach in solving problems with the help of fundamental data structures	

Contents –

Unit	Content	Teaching hours
1	Write a C program that uses functions to perform the following: a) Create a singly linked list of integers. b) Delete a given integer from the above linked list. c) Display the contents of the above list after deletion.	1
2	Write a C program that uses functions to perform the following: a) Create a doubly linked list of integers. b) Delete a given integer from the above doubly linked list. c) Display the contents of the above list after deletion	1
3	Write a C program that uses stack operations to convert a given infix expression into its postfix Equivalent, Implement the stack using an array.	1
4	Write C programs to implement a double ended queue ADT using i) array and ii) doubly linked list respectively.	1
5	Write a C program that uses functions to perform the following: a) Create a binary search tree of characters. b) Traverse the above Binary search tree recursively in Postorder	1
6	Write a C program that uses functions to perform the following: a) Create a binary search tree of integers. b) Traverse the above Binary search tree non recursively in inorder	1
7	Write C programs for implementing the following sorting methods to arrange a list of integers in ascending order: a) Insertion sort b) Merge sort	1
8	Write C programs for implementing the following sorting methods to arrange a list of integers in ascending order: a) Quick sort b) Selection sort	1
9	i) write a C program to perform the following operation:A) Insertion into a B-tree ii) Write a C program for implementing Heap sort algorithm for sorting a given List of integers in ascending order.	1
10	Write a C program to implement all the functions of a dictionary (ADT) using Hashing.	1

Text Books: Thomas Cormen. "Introduction to Algorithms"

Reference Books: Alfred V.Aho, Data structures and Algorithms O'Really

Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-I

Course code: MDS41MMP503	Course name: Practical Based on Python Programming
Course category: Major Mandatory	Credits: 1
Teaching scheme: L-0 P-2	Evaluation scheme: CA–30, ESE–20
Course Objectives: Understand and use variables. Work with common Python data types like integers, floats, strings, characters, lists, dictionaries, as well as pandas DataFrames.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: To understand importance of Python for data science	
CO2: Develop a skill to implement Python Programming for data science.	
CO3: Hands-on Python experience for professional advancement	
CO4: To use and implement standard programming constructs like data structure, numerical computing.	

Contents –

Unit	Content	Teaching hours
1	Python Program to Calculate the Area of a Triangle	1
2	Python Program to Swap Two Variables	1
3	Python Program to Find the Largest Among Three Numbers	1
4	Python Program to Print the Fibonacci sequence using function	1
5	Python program to interchange first and last elements in a list	1
6	Python Program to Display the multiplication Table using functions	1
7	Break a list into chunks of size N in Python	1
8	Python Counter Find all duplicate characters in string	1
9	Possible Words using given characters in Python	1
10	Python Program for Tower of Hanoi	1

Text Books: Mark Lutz's Learning Python O'Really

Reference Books: Mark Lutz's Programming Python O'Really

Online Resources: 1. NPTEL / SWAYAM lectures.

Elective from Basket-1 Syllabus
Semester-I

Course code: MDS41MEL501	Course name: Mathematics and Statistics
Course category: Elective	Credits: 3
Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basics of Mathematical and statistical concepts	
Course Objectives: The emphasis of course is on descriptive statistics. It gives an idea about the various statistical methods, measures of central tendency, measure of dispersion and correlation. Statistics is matter of science and logic. It mainly indulge on mathematics and logic.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: Understand the elementary statistical methods	
CO2: Apply the measures of central tendency, measure of dispersion and co-relation to solve our day-today life problem.	
CO3: Analyze the data to represent it graphically or tabulate and interpret it to generate information.	
CO4: Understanding of Mathematical optimization and Convex Optimization functions	

Contents –

Unit	Content	Teaching hours
1	Statistical Methods: Definition, scope and importance of Statistics, concepts of statistical population and sample. Data & Types of data: Primary and Secondary data, qualitative & quantitative data, Numerical (discrete, continuous), Categorical and Ordinal. Cross-section, time series, failure, industrial, directional data. attributes, variables, Processing of Data: Completeness, Consistency, Accuracy and Editing. Accuracy of Measurement. Classification, Tabulation and Graphical Representation: Preparation of Tables, Presentation of Data: Variable, Random Variable, Frequency, And Frequency Distribution. Diagrammatic representation of Data: Line and Bar Diagram, Histogram, Component Bar diagram, Pie Chart, Line Graph, Frequency polygon and Ogive. Measures of Skewness and Kurtosis	10
2	Measures of Central Tendency: Characteristics of Good measure of Central Tendency. Concept of central tendency- for Group and Ungroup data. Mean: Arithmetic mean (A.M.): simple and weighted Merits and demerits. Geometric mean (G.M.): computation for G M, Merits demerits and applications of G.M., Harmonic Mean (H.M.): computation for frequency, non-frequency data, merits and demerits of H.M., Median: Definition, Median for grouped and non-grouped data, Properties and Merits & demerits., Mode: Definition, Mode for grouped & Non-grouped data, Graphical Method for finding mode, Merits and demerits.	10
3	Measures of Dispersions: Purposes of Measure of Dispersion, Properties of Good measures of Dispersion. Range, Quartile Deviation & Mean Deviation: Variance, Standard Deviation, Coefficient of Variation: Bivariate data: Definition, scatter diagram, simple, partial and multiple correlation (3 variables only), rank correlation. Simple linear regression.	10
4	Introduction:	10

	Mathematical optimization, Least-squares and linear programming, Convex optimization, Nonlinear optimization. Convex sets: Affine and convex sets. Some important examples. Operations that preserve convexity. Generalized inequalities. Separating and supporting hyperplanes. Dual cones and generalized inequalities.	
5	Convex functions: Basic properties and examples. Operations that preserve convexity. The conjugate function. Quasiconvex functions. Log-concave and log-convex functions. Convexity with respect to generalized inequalities.	5

Text Books: 1. B.L.Agarwal Basic Statistics New Age International (P) Limited.
2. S. C. Gupta & V. K. Kapoor Fundamental of Mathematical Statistics Sultan Chand & Sons
Reference Books: S. C. Gupta Fundamental of Statistics
Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-I

Course code: MDS41MEL502	Course name: Data Warehousing
Course category: Elective	Credits: 3
Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Database management skills	
Course Objectives: The course exposes students to topics involving planning, designing, building, populating, and maintaining a successful data warehouse. Students learn the reasons why data warehousing is a compelling decision support solution, and study implementation approaches using proven methodologies and technologies.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: Analyze data warehouse characteristics and plan warehouse data (Dimensions, Facts, Hierarchies, Rollups)	
CO2: Illustrate trends towards data warehousing and data mining.	
CO3: Critically use all the data transformation processes.	
CO4: Estimate hardware infrastructure requirements.	

Contents –

Unit	Content	Teaching hours
1	Business Drivers: Business Drivers, A Concise History of Data Warehousing, What is OLAP & Decision Support Systems, Business Vs Data Requirements, Reasons to Build a Data Warehouse, The Technology Solution, Data Warehousing Concepts, Definition	10
2	Characteristics of a Data Warehouse: Data Marts, Identifying Business Data Flow & Processes, Data sources, The Data Extraction process, Transformation process, Assuring Data Quality, Transportation process, Maintaining Warehouse Data, Metadata	10
3	Data Warehouse Architecture and Models: Warehosuse architecture, Identifying Warehouse Data: Fact data, Dimension data, Hierarchies, Summaries (roll-ups)	10
4	Data Warehouse models: Star and Snowflake, Modelling the Data Warehouse Accessing a Data Warehouse, User query requirements and User query progression, OLAP Access, Relational OLAP (ROLAP) Access, Multidimensional OLAP (MOLAP) OLAP Query Techniques Definition of Data Mining	10
5	The Data Warehouse Challenge: The DW project: Scope Definition Prototyping ETT Tool selection, Implementation, Risks to: Organization, Business, Operations ,Technical considerations	5

Text Books: Connolly, Thomas, Carolyn Begg and Anne Strachan. Database Systems: A Practical Approach to Design, Implementation and Management. Addison Wesley, ISBN-13:9780321294012

Reference Books: William H. Inmon Building the Data Warehouse. ISBN 0-7645-9944-5 Wiley

Online Resources: 1. NPTEL / SWAYAM lectures.

<http://www.dw-institute.com/> and <http://www.dwinfocenter.org/>

Semester-I

Course code: MDS41MEL503	Course name: Compiler Design
Course category: Elective	Credits: 3
Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basics of Data and their structure	
Course Objectives: Understand the working of Computer System and Grammatical structures	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: Understand the different phases of compiler.	
CO2: Design a lexical analyzer for a sample language.	
CO3: Apply different parsing algorithms to develop the parsers for a given grammar.	
CO4: Understand syntax-directed translation and run-time environment.	

Contents –

Unit	Content	Teaching hours
1	Introduction To Compilers: Structure of a compiler – Lexical Analysis – Role of Lexical Analyzer – Input Buffering – Specification of Tokens – Recognition of Tokens – Lex – Finite Automata – Regular Expressions to Automata – Minimizing DFA.	10
2	Syntax Analysis: Role of Parser – Grammars – Error Handling – Context-free grammars – Writing a grammar – Top Down Parsing - General Strategies Recursive Descent Parser Predictive Parser-LL(1) Parser-Shift Reduce Parser-LR Parser-LR (0)Item Construction of SLR Parsing Table - Introduction to LALR Parser - Error Handling and Recovery in Syntax Analyzer-YACC.	10
3	Intermediate Code Generation: Syntax Directed Definitions, Evaluation Orders for Syntax Directed Definitions, Intermediate Languages: Syntax Tree, Three Address Code, Types and Declarations, Translation of Expressions, Type Checking.	10
4	Run-Time Environment And Code Generation: Storage Organization, Stack Allocation Space, Access to Non-local Data on the Stack, Heap Management - Issues in Code Generation - Design of a simple Code Generator.	10
5	Code Optimization: Principal Sources of Optimization – Peep-hole optimization - DAG-Optimization of Basic Blocks- Global Data Flow Analysis - Efficient Data Flow Algorithm.	5

Text Books: Alfred V. Aho, Monica S. Lam, Ravi Sethi, Jeffrey D. Ullman, Compilers: Principles, Techniques and Tools Pearson Education, 2009. Second Edition

Reference Books: Randy Allen, Ken Kennedy, Optimizing Compilers for Modern Architectures: A Dependence based Approach, Morgan Kaufmann Publishers, 2002.

Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-I

Course code: MDS41MEL504	Course name: Sentiment Analysis
Course category: Elective	Credits: 3
Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Data handling skills	
Course Objectives: This course is an introduction to Sentiment Analysis, with topics covered including relevant aspects of machine learning and Natural Language Processing.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: Understand the Data Analysis concepts	
CO2: Establish an understanding of concepts and theories of Sentiment Analysis.	
CO3: Understanding the types of Sentiment Analysis	
CO4: Understanding Sentiment Analysis how to implement Opinion Analysis	

Contents –

Unit	Content	Teaching hours
1	Sentiment Analysis: A Fascinating Problem, Sentiment Analysis Applications, Sentiment Analysis Research, Different Levels of Analysis, Sentiment Lexicon and Its Issues, Natural Language Processing Issues, Opinion Spam Detection, The Problem of Sentiment Analysis, Problem Definitions, Opinion Definition	10
2	Sentiment Analysis Tasks: Opinion Summarization, Different Types of Opinions, Regular and Comparative Opinions Explicit and Implicit Opinions, Subjectivity and Emotion. Document Sentiment Classification: Sentiment Classification Using Supervised Learning, Sentiment Classification Using Unsupervised Learning, Sentiment Rating Prediction.	10
3	Cross-Domain Sentiment Classification: Cross-Language Sentiment Classification, Sentence Subjectivity and Sentiment Classification, Sentiment Analysis and Opinion Mining, Subjectivity Classification, Sentence Sentiment Classification, Dealing with Conditional Sentences. Dealing with Sarcastic Sentences Cross-language Subjectivity and Sentiment Classification Using Discourse Information for Sentiment Classification.	10
4	Aspect-based Sentiment Analysis: Aspect Sentiment Classification, Basic Rules of Opinions and Compositional Semantics, Aspect Extraction, Finding Frequent Nouns and Noun Phrases, Using Opinion and Target Relations, Using Supervised Learning, Using Topic Models	10
5	Mapping Implicit Aspects: Identifying Resource Usage Aspect, Simultaneous Opinion Lexicon Expansion and Aspect, Extraction, Grouping Aspects into Categories, Entity, Opinion Holder and Time Extraction, Coreference Resolution and Word Sense Disambiguation	5

Text Books: Bing Liu Sentiment Analysis and Opinion Mining Morgan & Claypool Publishers, May 2012.

Reference Books: Bo Pang and Lillian Lee Opinion mining and sentiment analysis

Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-I

Course code: MDS41MEP501	Course name: Practical Based on Mathematics and Statistics
Course category: Elective	Credits: 1
Teaching scheme: L-0 P-2	Evaluation scheme: CA–30, ESE–20
Pre-requisites: Basics of Mathematical concepts	
Course Objectives: The emphasis of course is on descriptive statistics. It gives an idea about the various statistical methods, measures of central tendency, measure of dispersion and correlation. Statistics is matter of science and logic. It mainly indulge on mathematics and logic.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: Understand the elementary statistical methods.	
CO2: Apply the measures of central tendency, measure of dispersion and co-relation to solve our day-today life problem.	
CO3: Analyze the data to represent it graphically or tabulate and interpret it to generate information.	
CO4: Understanding of Mathematical optimization and Convex Optimization functions	

Contents –

Unit	Content	Teaching hours
1	Definition of Statistic, Data & Types of data	1
2	Presentation of Data: Frequency, And Frequency Distribution. Diagrammatic representation of Data: Line and Bar Diagram, Histogram, Component Bar diagram, Pie Chart, Line Graph, Frequency polygon and Ogive.	1
3	Arithmetic mean (A.M.): Formula, Problems based on AM	1
4	Tabulated & Class based AM Problems.	1
5	Geometric mean (G.M.): Formula, Problems based on GM	1
6	Harmonic Mean (H.M.): Formula, Problems based on HM	1
7	Median: Computation & Problem based on Median.	1
8	Mode: Computation & Problem based on Mode.	1
9	Range, Quartile Deviation Mean Deviation: Problems	1
10	Variance & Standard Deviation	1

Text Books: B.L. Agarwal Basic Statistics New Age International (P) Limited.
Reference Books: S. C. Gupta & V. K. Kapoor Fundamental of Mathematical Statistics Sultan Chand & Sons
Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-I

Course code: MDS41MEP502	Course name: Practical Based on Data Warehousing
Course category: Elective	Credits: 1
Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Pre-requisites: Basics of Data Structure	
Course Objectives: The course exposes students to topics involving planning, designing, building, populating, and maintaining a successful data warehouse. Students learn the reasons why data warehousing is a compelling decision support solution, and study implementation approaches using proven methodologies and technologies.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: Analyze data warehouse characteristics and plan warehouse data (Dimensions, Facts, Hierarchies, Rollups)	
CO2: Illustrate trends towards data warehousing and data mining.	
CO3: Critically use all the data transformation processes.	
CO4: Estimate hardware infrastructure requirements.	

Contents –

Unit	Content	Teaching hours
1	Getting Started with Learning About Data Warehousing	1
2	Performing Data Warehousing Software Evaluations	1
3	Design Data Warehouse Architecture	1
4	Design Data Warehouse Querying	1
5	Design how to Data Warehouse Loading	1
6	How to Save Money on Your Data Warehousing Efforts	1
7	Using Data Warehousing in Strategic Decision Making	1
8	Maintenance Issues for Data Warehousing Systems	1
9	What Decision Support Tools are Used For Dataset	1
10	Design warehouse system for Decision Support	1

Text Books: Connolly, Thomas, Carolyn Begg and Anne Strachan. Database Systems: A Practical Approach to Design, Implementation and Management. Addison Wesley, ISBN-13:9780321294012
Reference Books: William H. Inmon. Building the Data Warehouse.
Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-I

Course code: MDS41MEP503	Course name: Practical Based on Compiler Design
Course category: Elective	Credits: 1
Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Course Objectives: Understand the working of Computer System and Grammatical structures	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: Understand the different phases of compiler.	
CO2: Design a lexical analyzer for a sample language.	
CO3: Apply different parsing algorithms to develop the parsers for a given grammar.	
CO4: Understand syntax-directed translation and run-time environment.	

Contents –

Unit	Content	Teaching hours
1	Develop a lexical analyzer to recognize a patterns	1
2	Create a symbol table, while recognizing identifiers.	1
3	Create a symbol table, while recognizing constants	1
4	Create a symbol table, while recognizing comments	1
5	Create a symbol table, while recognizing operators	1
6	Implement a Lexical Analyzer using Lex Tool	1
7	Implement an Arithmetic Calculator using LEX and YACC	1
8	Generate three address code for a simple program using LEX.	1
9	Generate three address code for a simple program using YACC.	1
10	Implement simple code optimization techniques (Constant folding, Strength reduction and Algebraic transformation)	1

Text Books: Alfred V. Aho, Monica S. Lam, Ravi Sethi, Jeffrey D. Ullman, Compilers: Principles, Techniques and Tools,

Reference Books: Randy Allen, Ken Kennedy, Optimizing Compilers for Modern Architectures: A Dependence based Approach

Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-I

Course code: MDS41MEP504	Course name: Practical Based on Sentiment Analysis
Course category: Elective	Credits: 1
Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Course Objectives: This course is an introduction to Sentiment Analysis, with topics covered including relevant aspects of machine learning and Natural Language Processing.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: Understand the Data Analysis process	
CO2: Establish an understanding of concepts and theories of Sentiment Analysis.	
CO3: Understanding the types of Sentiment Analysis	
CO4: Understanding Sentiment Analysis how to implement Opinion Analysis	

Contents –

Unit	Content	Teaching hours
1	Installation of R Programing and Configuration	1
2	Installation of R Studio and configuration	1
3	Loading Sentiment Analysis Packages	1
4	Social media monitoring.	1
5	Customer support ticket analysis	1
6	Brand monitoring and reputation management.	1
7	Product analysis	1
8	Amazon Product Reviews	1
9	Analyze IMDb Reviews	1
10	Twitter Sentiment Analysis	1

Text Books: Bing Liu Sentiment Analysis and Opinion Mining Morgan & Claypool Publishers, May 2012.

Reference Books: Bo Pang and Lillian Lee Opinion mining and sentiment analysis

Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-II

Course code: MDS41MML504	Course name: Advances in Database Management System
Course category: Major Mandatory	Credits: 3
Teaching scheme: L- 3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basic knowledge about computers including some experience using UNIX or Windows. Understanding of creation and updating of tables.	
Course Objectives: Students can design new database and modify existing ones for new or existing applications	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: To know the different issues involved in the design and implementation of a database system.	
CO2: Can use data manipulation language to query, update, and manage a database.	
CO3: Essential DBMS concepts such as: database security, integrity, concurrency, etc.	
CO4: Understand Advanced Application Development	

Contents –

Unit	Content	Teaching hours
1	Introduction to DBMS: Conceptual Database Design: Database Administrator (DBA), Database User, Database System, Database Architecture, DBMS Components. Data Models: Schemas and Instances, Database Languages, Database Structure, E-R Model, Strong and Weak Entity Sets.	10
2	SQL and Normalization: Features of SQL, Data Definition Languages (DDL), Data Manipulation Languages (DML), Views, Functions, RollBack, Commit and Savepoint, Indexes. Normalization: 1NF, 2NF, 3NF, 4NF, 5NF, Flowchart of Normalization, Domain Key Normal Form (DKNF).	10
3	Concurrency Control Techniques and Recovery: Lock-based Protocols, Deadlock Handling, Recovery System, Failure Classification, Store Structure, Categorization of Recovery Algorithm, Log-based Recovery, Shadow Paging, Recovery with Concurrent Transaction, Check Points.	10
4	Parallel and Web Databases: Design of Parallel Databases, Parallel Query Evaluation, Advantages of Parallel Databases, Elements of Parallel Database Processing. Introduction to Web Databases, Accessing Database Through the Web, XML Database. Multimedia Databases, Mobile Databases, Digital Databases.	10
5	Advanced Application Development: Performance Tuning, Bottlenecks, Tuning the Database Design, Performance Simulation, Database Application Classes, Benchmarks Suites, Standardization, Database Connectivity Standards, Object Oriented Databases Standards, Marketplaces, Secure Payment Systems, Legacy Systems.	5

Text Books: 1. Rajiv Chopra Database Management System (DBMS) A Practical Approach S. Chand Publishing
Reference Books: Avi Silberschatz, Henry F. Korth and S. Sudarshan Database System Concepts McGraw-Hill Education
Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-II

Course code: MDS41MML505	Course name: Data Mining and Visualization
Course category: Major Mandatory	Credits: 3
Teaching scheme: L- 3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: students should know the basic concepts of data Structure and analysis	
Course Objectives: Students will be able to actively manage and participate in data mining projects. To develop research interest towards advances in data mining. Students will be able to understand the visualization techniques	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: Identify appropriate data mining algorithms to solve real world problems.	
CO2: Compare and evaluate different data mining techniques like classification, prediction, clustering and association rule mining.	
CO3: Describe complex data types with respect to spatial and Data Visualization.	
CO4: Benefit the user experiences towards research and innovation. Integration in Data Mining area.	

Contents –

Unit	Content	Teaching hours
1	Introduction to Data Mining: Why Mine Data, Commercial Viewpoint, Scientific Viewpoint Motivation, Definitions, Origins of Data Mining, Data Mining Tasks, Classification, Clustering, Association Rule Discovery, Sequential Pattern Discovery, Regression, Challenges of Data Mining, Data Mining Data: What is Data? Attribute Values, Measurement of Length, Types and Properties of Attributes, Discrete and Continuous Attributes, Types of data sets, Data Quality, Data Preprocessing, Aggregation, Sampling, Dimensionality Reduction, Feature subset selection, Feature creation, Discretization and Binarization, Attribute Transformation, Density.	10
2	Data Mining: Exploring Data, Data Exploration Techniques, Summary Statistics, Frequency and Mode, Percentiles, Measures of Location: Mean and Median, Measures of Spread: Range and Variance, Visualization, Representation, Arrangement, Selection, Visualization Techniques: Histograms, Box Plots, Scatter Plots, Contour Plots, Matrix Plots, Parallel Coordinates, Other Visualization Techniques, OLAP : OLAP Operations	10
3	Data Mining Classification: Basic Concepts, Decision Trees, and Model Evaluation: Classification: Definition, Classification Techniques, Tree Induction, Measures of Node Impurity, Practical Issues of Classification, ROC curve, Confidence Interval for Accuracy, Comparing Performance of Two Models, Comparing Performance of Two Algorithms.	10
4	Data Mining Classification:	10

	Alternative Techniques: Rule-Based Classifier, Rule Ordering Schemes, Building Classification Rules, Instance-Based Classifiers, Nearest Neighbor Classifiers, Bayes Classifier, Naive Bayes Classifier, Artificial Neural Networks (ANN), Support Vector Machines.	
5	<p>Introduction to Data Visualization:</p> <p>Classification of Visualization techniques – Structure and representation, Selection of a Visualization, Visualizations for high dimensional data, Graphics sand computing, Principles of Data Visualization, Multivariate data, Linked data, Visualizing trees and forests, Large Datasets – Plots and their variates, Visualizing cluster analysis ,contingency tables – finite mixture models, Methodologies: Visualization in Bayesian data analysis, Matrix visualization, Data visualization by kernel machines .Applications Visualization for genetic network reconstruction, medical images, financial dataset and Insurance risk processes.</p>	5

Text Books: Tan, Steinbach, Kumar. Introduction to Data Mining

Reference Books: Jiawei Han, MichelineKamber Data Mining: Concepts and Techniques Morgan Kaufmann Publishers

Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-II

Course code: MDS41MML506	Course name: Machine Learning
Course category: Major Mandatory	Credits: 3
Teaching scheme: L- 3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basic knowledge about Data Mining and data warehousing	
Course Objectives: To introduce students to the basic concepts and techniques of Machine Learning. To become familiar with regression methods, classification methods, clustering methods.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: To become familiar with Dimensionality reduction Techniques.	
CO2: Identify machine learning techniques suitable for a given problem	
CO3: At the end of the course the students should be able to design and implement machine learning solutions to classification, regression, and clustering problems; and be able to evaluate and interpret the results of the algorithms.	
CO4: Understand the Data Clustering	

Contents –

Unit	Content	Teaching hours
1	Introduction to Machine Learning- Introduction to Analytics and Machine Learning, What is Machine Learning, Types of Machine Learning, Applications of Machine Learning, Issues in Machine Learning, Why Machine Learning?, Framework for Developing Machine Learning Models, State-of-the-Art Languages / Tools In Machine Learning. Descriptive Analytics- Working with DataFrames in Python, Handling Missing Values, Exploration of Data using Visualization: Drawing Plots, Bar Chart, Histogram, Distribution or Density Plot, Box Plot, Comparing Distributions, Scatter Plot, Pair Plot, Correlation and Heatmap.	10
2	Linear Regression- Simple Linear Regression, Steps in Building a Regression Model, Building Simple Linear Regression Model, Model Diagnostics, Multiple Linear Regression.	10
3	Classification- Classification Overview, Binary Logistic Regression, Credit Classification, Gain Chart and Lift Chart, Classification Tree (Decision Tree Learning)	10
4	Clustering- Overview, Clustering Working, K-Means Clustering, Creating Product Segments Using Clustering, Hierarchical Clustering.	10
5	Advances in Machine Learning- Gradient Descent Algorithm, Scikit-Learn Library for Machine Learning, K-Nearest Neighbors (KNN) Algorithm, Random Forest.	5

Text Books:

1. Machine Learning using Python, Manaranjan Pradhan, U Dinesh Kumar, Wiley
2. Machine Learning, Author: Saikat Dutt, S. Chandramouli, Amit K.Das, Pearson Pub.

Reference Books:

1. Machine Learning with Python, By Abhishek Vijayvargia, BPB Publications
2. Machine Learning, By Mitchell Tom, McGraw Hill Pub.

Elective from Basket-2 Semester-II

Course code: MDS41MEL505	Course name: Google Analytics
Course category: Elective	Credits: 3
Teaching scheme: L- 3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Data Analysis techniques	
Course Objectives: Configuration of Google Analytics Understand different types of insights you can generate using reports available in Google Analytics to support improvement decisions	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: Determine basic navigation of Google Analytics Interface.	
CO2: Identify the benefits of a customer-centric approach to website design and optimization	
CO3: How to navigate the Google Analytics interface and reports, and set up dashboards and shortcuts.	
CO4: Determine site content in GA, filtering data techniques	

Contents –

Unit	Content	Teaching hours
1	Introducing Web Analytics : Defining Web Analytics, Quantitative and Qualitative Data, The Continuous Improvement Process, Measuring Outcomes, What Google Analytics Contributes, How Google Analytics Fits in the Analytics Ecosystem, Creating an Implementation Plan: Gather Business Requirements, Analyze and Document Website Architecture, Create an Account and Configure Your Profile, Configure the Tracking Code and Tag Pages, Tag Marketing Campaigns, Create Additional User Accounts and Configure Reporting Features, Perform Optional Configuration Steps.	10
2	How Google Analytics Works : Data Collection and Processing, Reports, About the Tracking Code, The Mobile Tracking Code, App Tracking, The (Very) Old Tracking Code: urchin.js, Understanding Page views. Tracking Visitor Clicks, Outbound Links, and Non-HTML Files: About the Tracking Cookies Designing Blogs for Google Analytics	10
3	Google Analytics Accounts and Profiles: Google Analytics Accounts, Creating a Google Analytics Account, Creating Additional Profiles, Access Levels, All About Profiles, Basic Profile Settings, Profile Name, Website URL, Time Zone, Default Page, Exclude URL Query Parameters, E-Commerce Settings, Tracking On-Site Search, Applying Cost Data. Filters: Filter Fields, Filter Patterns, Filter Type, Include/Exclude Filters, Search and Replace Filters, Lowercase/Uppercase Filters, Advanced Profile Filters, Predefined Filters.	10
4	Tracking Conversions with Goals and Funnels: Goals, Time on Site, Pages per Visit, URL Destinations, Additional Goal Settings, Tracking Defined Processes with Funnels.	10
5	Must-Have Profiles: Profile Roles, Raw Data Profile, Master Profile, Test Profile, Access-Based Profiles, Using Profiles to Segment Data, Exclude Internal Traffic, Include Valid Traffic, Force Request URI to Lowercase, Force Campaign Parameters to Lowercase, Keeping Track of Your Configuration Changes.	5

Text Books: Todd Kelsey Justin Cutroni Google Analytics O'Reilly
Reference Books: Introduction to Google Analytics: A Guide for Absolute Beginners APress
Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-II

Course code: MDS41MEL506	Course name: Neural Networks
Course category: Elective	Credits: 3
Teaching scheme: L- 3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Introduction to Machine Learning	
Course Objectives: To study learning and modeling of the algorithms of Neural Networks.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: To Learn to Feed Forward Neural Networks	
CO2: Learn working of Supervised Learning	
CO3: Understand the modeling and applications of Neural Networks	
CO4: Algorithms used in Neural Networks	

Contents –

Unit	Content	Teaching hours
1	Introduction to Feedforward Neural Networks: Artificial Neurons, Neural Networks and Architectures: Neuron Abstraction, Neuron Signal Functions, Mathematical Preliminaries, Neural Networks Defined, Architectures: Feed forward and Feedback, Salient Properties and Application Domains of Neural Network.	10
2	Geometry of Binary Threshold Neurons and Their Network: Patterns Recognition and Data Classification, Convex Sets, Convex Hulls and Linear Separability, Space of Boolean Functions, Binary Neurons are pattern Dichotomizes, Non-linearly separable Problems, Capacity of a simple Threshold Logic Neuron, Revisiting the XOR Problem, Multilayer Networks.	10
3	Supervised Learning: Supervised Learning I: Perceptrons and LMS: Learning and Memory, From Synapses to Behaviour: The Case of Aplysia, Learning Algorithms, Error Correction and Gradient Descent Rules, The Learning Objective for TLNs, Pattern space and Weight Space, Perceptron Learning Algorithm, Perceptron Convergence Theorem	10
4	Perceptron Learning: Perceptron learning and Non-separable Sets, Handling Linearly Non-Separable sets, α -Least Mean Square Learning, MSE Error Surface and its Geometry, Steepest Descent Search with Exact Gradient Information, μ -LMS: Approximate Gradient Descent, Application of LMS to Noise Cancellation.	10
5	Supervised Learning II: Backpropagation and Beyond: Multilayered Network Architectures, Backpropagation Learning Algorithm, Structure Growing Algorithms, Fast Relatives of Backpropagation, Universal Function Approximation and Neural Networks, Applications of Feedforward Neural Networks, Reinforcement Learning	5

Text Books: Satish Kumar Neural Network- A Classroom Approach. Tata McGraw Hill
Reference Books: Sivanandam, S Sumathi, S N Deepa Introduction to neural networks using MATLAB 6.0 TATA McGraw HILL
Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-II

Course code: MDS41MEL507	Course name: Optimization Techniques
Course category: Elective	Credits: 3
Teaching scheme: L- 3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basics of statistics	
Course Objectives: Introduction to optimization techniques using both linear and non-linear programming.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: - To Learn Linear algebra	
CO2: To Learn Linear Programming	
CO3: Cast problems into optimization framework	
CO4: Learn efficient computational procedures to solve optimization problems.	

Contents –

Unit	Content	Teaching hours
1	Mathematical preliminaries: Linear algebra and matrices, Vector space, eigen analysis, Elements of probability theory, Elementary multivariable calculus.	10
2	Linear Programming: Introduction to linear programming model, Simplex method, Duality, Karmarkar's method.	10
3	Unconstrained optimization: Conjugate direction and quasi-Newton methods, Gradient-based methods, One-dimensional search methods	10
4	Constrained Optimization: Constrained Optimization Lagrange theorem, FONC, SONC, and SOSC conditions, Projection methods	10
5	KKT: KKT conditions, Non-linear constrained optimization models, Non-linear problems	5

Text Books: Edwin P K Chong, Stainslaw Zak An introduction to Optimization
Reference Books: Dimitri Bertsekas Nonlinear Programming
Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-II

Course code: MDS41MEL508	Course name: Business Intelligence
Course category: Elective	Credits: 3
Teaching scheme: L- 3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Introduction to Analytics techniques	
Course Objectives: This course provides an introduction to the concepts of business intelligence (BI) as components and functionality of information systems.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: To learn organization's business operations through the use of relevant data.	
CO2: It explores how business problems can be solved effectively by using operational data to create data warehouses	
CO3: Applying data mining tools and analytics to gain new insights into organizational operations.	
CO4: Understand decision support systems and knowledge management systems	

Contents –

Unit	Content	Teaching hours
1	Business intelligence: Effective and timely decisions, Data, information and knowledge, The role of mathematical models, Business intelligence architectures, Ethics and business intelligence Decision support systems: Definition of system, Representation of the decision-making process, Evolution of information systems, Definition of decision support system, Development of a decision support system	10
2	Mathematical models for decision making: Structure of mathematical models, Development of a model, Classes of models Data mining: Definition of data mining, Representation of input data , Data mining process, Analysis methodologies Data preparation: Data validation, Data transformation, Data reduction	10
3	Classification: Classification problems, Evaluation of classification models, Bayesian methods, Logistic regression, Neural networks, Support vector machines Clustering: Clustering methods, Partition methods, Hierarchical methods, Evaluation of clustering models	10
4	Business intelligence applications: Marketing models: Relational marketing, Sales force management, Logistic and production models: Supply chain optimization, Optimization models for logistics planning, Revenue management systems. Data envelopment analysis: Efficiency measures, Efficient frontier, The CCR model, Identification of good operating practices	10
5	Knowledge Management: Introduction to Knowledge Management, Organizational Learning and Transformation, Knowledge Management Activities, Approaches to Knowledge Management, Information Technology (IT) In Knowledge Management, Knowledge Management Systems Implementation, Roles of People in Knowledge Management	5

Text Books: Carlo Vercellis Business Intelligence: Data Mining and Optimization for Decision Making Wiley First 2009

Reference Books: Efraim Turban, Ramesh Sharda, Dursun Delen Decision support and Business Intelligence Systems Pearson

Online Resources: 1. NPTEL / SWAYAM lectures.
--

Semester-II

Course code: MDS41MMP504	Course name: Practical Based on ADBMS
Course category: Major Mandatory	Credits: 1
Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Course Objectives: Students can design new database and modify existing ones for new or existing applications	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: To know the different issues involved in the design and implementation of a database system.	
CO2: Can use data manipulation language to query, update, and manage a database.	
CO3: Essential DBMS concepts such as: database security, integrity, concurrency, etc.	
CO4: Understand Advanced Application Development	

Contents –

Unit	Content	Teaching hours
1	Create Student Database	1
2	Create Student Detail and Marks Table	1
3	Insert Records to Student Detail and Marks Table	1
4	Display all records of Student Details and Marks Table at once	1
5	Index on table Student Details with the name and date of birth	1
6	Create Table and Perform the DDL Commands	1
7	Create Table and Perform the DML Commands	1
8	Create Database with Normalization	1
9	Demo for Recovery with Concurrent Transaction	1
10	Demo for Accessing Database Through the Web	1

Text Books: Rajiv Chopra Database Management System (DBMS)A Practical Approach S. Chand Publishing

Reference Books: Avi Silberschatz, Henry F. Korth and S. Sudarshan Database System Concepts McGraw-Hill Education

Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-II

Course code: MDS41MMP505	Course name: Practical Based on Data Mining and Visualization
Course category: Major Mandatory	Credits: 1
Teaching scheme: L-0 P-2	Evaluation scheme: CA–30, ESE–20
Pre-requisites: Basics concepts of Data Structure	
Course Objectives: Students will be able to actively manage and participate in data mining projects. To develop research interest towards advances in data mining. Students will be able to understand the visualization techniques	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: Identify appropriate data mining algorithms to solve real world problems.	
CO2: Compare and evaluate different data mining techniques like classification, prediction, clustering and association rule mining.	
CO3: Describe complex data types with respect to spatial and Data Visualization.	
CO4: Benefit the user experiences towards research and innovation. Integration in Data Mining area.	

Contents –

Unit	Content	Teaching hours
1	Demonstration of preprocessing on dataset student.arff	1
2	Demonstration of preprocessing on dataset labor.arff	1
3	Demonstration of Association rule process on dataset contactlenses.arff using apriori algorithm	1
4	Demonstration of Association rule process on dataset test.arff using apriori algorithm	1
5	Demonstration of classification rule process on dataset using Nearest neighbor algorithm	1
6	Demonstration of classification rule process on dataset using K-NN algorithm	1
7	Demonstration of classification rule process on dataset using Decision tree algorithm	1
8	Demonstration of classification rule process on dataset using Regression algorithm	1
9	Apply Visualization techniques for Various Dataset	1
10	Apply Visualization techniques for Various Dataset	1

Text Books: Tan, Steinbach, Kumar. Introduction to Data Mining
Reference Books: Jiawei Han, MichelineKamber Data Mining: Concepts and Techniques Morgan Kaufmann Publishers
Online Resources: 1. NPTEL / SWAYAM lectures.

Semester-II

Course code: MDS41MMP506	Course name: Practical Based on Machine Learning
Course category: Major Mandatory	Credits: 1
Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Course Objectives: Students can design new database and modify existing ones for new or existing applications	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: To become familiar with Dimensionality reduction Techniques.	
CO2: Identify machine learning techniques suitable for a given problem	
CO3: At the end of the course the students should be able to design and implement machine learning solutions to classification, regression, and clustering problems; and be able to evaluate and interpret the results of the algorithms.	
CO4: Understand the Data Clustering	

Contents –

Unit	Content	Teaching hours
1	Implement and demonstrate the FIND-S algorithm for finding the most specific hypothesis based on a given set of training data samples. Read the training data from a .CSV file.	1
2	For a given set of training data examples stored in a .CSV file, implement and demonstrate the Candidate-Elimination algorithm to output a description of the set of all hypotheses consistent with the training examples.	1
3	Write a program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate data set for building the decision tree and apply this knowledge to classify a new sample.	1
4	Build an Artificial Neural Network by implementing the Backpropagation algorithm and test the same using appropriate data sets.	1
5	Write a program to implement the naïve Bayesian classifier for a sample training data set stored as a .CSV file. Compute the accuracy of the classifier, considering few test datasets	1
6	Assuming a set of documents that need to be classified, use the naïve Bayesian Classifier model to perform this task. Built-in Java classes/API can be used to write the program. Calculate the accuracy, precision, and recall for your data set.	1
7	Write a program to construct a Bayesian network considering medical data. Use this model to demonstrate the diagnosis of heart patients using standard Heart Disease Data Set. You can use Java/Python ML library classes/API.	1
8	Apply EM algorithm to cluster a set of data stored in a .CSV file. Use the same data set for clustering using k-Means algorithm. Compare the results of these two algorithms and comment on the quality of clustering. You can add Java/Python ML library classes/API in the program.	1
9	Write a program to implement k-Nearest Neighbour algorithm to classify the iris data set. Print both correct and wrong predictions. Java/Python ML library classes can be used for this problem.	1
10	Implement the non-parametric Locally Weighted Regression algorithm in order to fit data points. Select appropriate data set for your experiment and draw graphs.	1

Text Books: M. Gopal Applied Machine Learning McGraw-Hill Education
--

Reference Books: Charu C. Aggarwal Machine Learning for Text Springer
--

Practical based on Elective-I Semester-II

Course code: MDS41MEP505	Course name: Practical Based on Google Analytics
Course category: Elective	Credits: 1
Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Course Objectives: Configuration of Google Analytics and Understand different types of insights you can generate using reports available in Google Analytics to support improvement decisions.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: Determine basic navigation of Google Analytics Interface	
CO2: Identify the benefits of a customer-centric approach to website design and optimization	
CO3: How to navigate the Google Analytics interface and reports, and set up dashboards and shortcuts.	
CO4: Determine site content in GA, filtering data techniques	

Contents –

Unit	Content	Teaching hours
1	Create Google Analytics Account	1
2	Configure Google Analytics Account	1
3	Create Web blogs	1
4	Create Google Analytics Tracking Code and insert in Blog	1
5	Creating an Implementation Plan of Business case	1
6	Configure Google Analytics Accounts and Profiles	1
7	Create Administrator and User Profiles	1
8	Create Filters in Google Analytics	1
9	Create Tracking Conversions with Goals	1
10	Create Tracking Conversions with Funnels	1

Text Books: Justin Cutroni Google Analytics O'Reilly

Reference Books: Todd Kelsey Introduction to Google Analytics: A Guide for Absolute Beginners
APress

Online Resources: 1. NPTEL / SWAYAM lectures.

Practical based on Elective-I Semester-II

Course code: MDS41MEP506	Course name: Practical Based on Neural Network
Course category: Elective	Credits: 1
Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Course Objectives: To study learning and modeling of the algorithms of Neural Networks.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: Introduction to Feed forward Neural Networks	
CO2: Learn working of Supervised Learning	
CO3: Understand the modeling and applications of Neural Networks	
CO4: Algorithms used in Neural Networks	

Contents –

Unit	Content	Teaching hours
1	To study about MATLAB.	1
2	Write a program to perform the basics matrix operations.	1
3	WAP to plot the Straight line.	1
4	WAP to plot the Sine curve	1
5	How the weight & bias value effects the output of neurons.	1
6	How the choice of activation function effect the output of neuron experiment with the following function purelin(n), bimary threshold(hardlim(n) haradlims(n)) ,Tansig(n) logsig(n)	1
7	How the weight and biased value are able to represent a decision boundary in the feature space.	1
8	How the Perceptron Learning rule works for Linearly Separable Problem.	1
9	How the Perceptron Learning rule works for Non-Linearly Separable Problem.	1
10	Write a program to draw a graph with multiple curve.	1

Text Books: Satish Kumar Neural Network- A Classroom Approach. Tata McGraw Hill
Reference Books: Sivanandam, S Sumathi, S N Deepa Introduction to neural networks using MATLAB 6.0 TATA McGraw HILL
Online Resources: 1. NPTEL / SWAYAM lectures.

Practical based on Elective-I Semester-II

Course code: MDS41MEP507	Course name: Practical Based on Optimization Techniques
Course category: Elective	Credits: 1
Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Course Outcomes: At the end of the course, the students will be able to -	
CO1: To Learn Linear algebra	
CO2: To Learn Linear Programming	
CO3: Cast problems into optimization framework	
CO4: Learn efficient computational procedures to solve optimization problems.	

Contents –

Unit	Content	Teaching hours
1	Installation and configuration of Matlab	1
2	Introduction to MATLAB	1
3	Understand basics and engineering applications of optimization	1
4	Get familiar with mathematical software to solve optimization problems in MATLAB	1
5	Understand basics of Classical Optimization Techniques	1
6	Understand Elimination Methods of Unconstrained Optimization	1
7	Understand Interpolation Methods of Unconstrained Optimization	1
8	Understand Direct Root Methods of Unconstrained Optimization	1
9	Understand and solve Equality Constraints problems of Constrained Optimization	1
10	Understand and solve Inequality Constraints problems of Constrained Optimization	1

Text Books: Edwin P K Chong, Stainslaw Zak An introduction to Optimization

Reference Books: Dimitri Bertsekas Nonlinear Programming

Online Resources: 1. NPTEL / SWAYAM lectures.

Practical based on Elective-I Semester-II

Course code: MDS41MEP508	Course name: Practical Based on Business Intelligence
Course category: Elective	Credits: 1
Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Course Objectives: This course provides an introduction to the concepts of business intelligence (BI) as components and functionality of information systems.	
Course Outcomes: At the end of the course, the students will be able to -	
CO1: To learn organization's business operations through the use of relevant data.	
CO2: It explores how business problems can be solved effectively by using operational data to create data warehouses	
CO3: Applying data mining tools and analytics to gain new insights into organizational operations.	
CO4: Understand decision support systems and knowledge management systems	

Contents –

Unit	Content	Teaching hours
1	Import the legacy data from different sources such as (Excel , SqlServer, Oracle etc.) and load in the target system. (You can download sample database such as Adventureworks, Northwind, foodmart etc.)	1
2	Perform the Extraction Transformation and Loading (ETL) process to construct the database in the Sqlserver.	1
3	Create the Data staging area for the selected database.	1
4	Create the cube with suitable dimension and fact tables based on ROLAP, MOLAP and HOLAP model.	1
5	Create the ETL map and setup the schedule for execution.	1
6	Execute the MDX queries to extract the data from the datawarehouse.	1
7	Import the datawarehouse data in Microsoft Excel and create the Pivot table and Pivot Chart.	1
8	Import the cube in Microsoft Excel and create the Pivot table and Pivot Chart to perform data analysis	1
9	Apply the what – if Analysis for data visualization. Design and generate necessary reports based on the data warehouse data	1
10	Perform the data classification using classification algorithm	1

Text Books: Carlo Verzellis Business Intelligence: Data Mining and Optimization for Decision Making Wiley

Reference Books Efraim Turban, Ramesh Sharda, Dursun Delen Decision support and Business Intelligence Systems Pearson

Online Resources: 1. NPTEL / SWAYAM lectures.

--

Semester: THIRD

Syllabus Semester-III

Course code: MDS41MML601	Course name: Big Data	Course category: Major Mandatory
Credits: 3	Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basic knowledge of Data science concepts and applications.		
Course Objectives:		
2. Understand the Big data platform and its use cases.		
Course Outcomes: At the end of the course, the students will be able to -		
CO1: Identify Big Data and its Business Implications.		
CO2: List the components of Hadoop and Hadoop Eco-System		
CO3: Access and Process Data on Distributed File System		
CO4: Solve ordinary differential equations analytically and numerically and apply these methods to solve engineering problems.		
CO5: Manage Job Execution in Hadoop Environment		

Contents-

Unit	Content	Teaching hours
1	Introduction to Big Data -Data, Data Storage and Analysis, Comparison with Other Systems, RDBMS, Grid Computing, Volunteer Computing, A Brief History of Hadoop, Apache Hadoop and the Hadoop Ecosystem	10
2	MapReduce -A Weather Dataset, Data Format, Analyzing the Data with Unix Tools, Analyzing the Data with Hadoop, Map and Reduce, Java MapReduce Scaling Out, Data Flow, Combiner Functions, Running a Distributed MapReduce Job, Hadoop Streaming, Ruby, Python, Hadoop Pipes, Compiling and Running	10
3	The Hadoop Distributed Filesystem :- The Design of HDFS,HDFS Concepts, Blocks, Namenodes and Datanodes, HDFS Federation, HDFS High-Availability, Failover and fencing, The Command-Line Interface, Basic Filesystem Operations, Hadoop Filesystems, Interfaces, HTTP, FUSE, The Java Interface, Reading Data from a Hadoop URL, Reading Data Using the FileSystem API, FSDataInputStream, Writing Data, Directories, Querying the Filesystem, Deleting Data, Data Flow, Coherency Model, Parallel Copying with distcp, Keeping an HDFS Cluster Balanced.	10
4	Hadoop I/O :- Data Integrity, Compression, Serialization, Writable Classes, Serialization Frameworks, File-Based Data Structures, MapFile	5
5	Developing a MapReduce Application :- The Configuration API, Configuring the Development Environment, Writing a Unit Test, Writing a Unit Test, Testing the Driver, Running on a Cluster, Running on a Cluster, The MapReduce Web UI, Retrieving the Results, Debugging a Job, Hadoop Logs, Remote Debugging, Tuning a Job, Profiling Tasks, MapReduce Workflows,	10

Text Books: 1.Hadoop: The Definitive Guide Tom WhiteO'reily Media, Third Edit on, 2012.
2. Big Data Analytics Seema Acharya, Subhasini Chellappan, Wiley 2015.

Reference Books: 1. "Big Data and Business Analytics" Jay Liebowitz Auerbach Publications, CRC press (2013).
3. "Big data principles and paradigms", edited by rajkumar buyya,Rodrigo N. calheiros, Amir Vahid Dastjerdi in 2016.

Syllabus Semester-III

Course code: MDS41MML602	Course name: Cloud Computing	Course category: Major Mandatory
Credits: 3	Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basic Knowledge of networking & Operating system		
Course Objectives: Understanding basics of cloud computing and Key concepts of Cloud Analytics and different cloud computing services.		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Understand Cloud Computing Basics		
CO 2: Understand the use of Cloud Computing in Data Analytics		
CO 3: Learn the Concept of Cloud Infrastructure		
CO 4: Understand Business imperative of Cloud Computing		
CO 5: Learn Cloud computing Security Components		

Contents –

Unit	Content	Teaching hours
1	Introduction to Cloud computing:- Evolution of computing paradigms, Concept of cloud, Introduction to virtualization and virtual machine, Virtualization in fabric/cluster/grid context, Virtual network, Information model & data model for virtual machine, Service Oriented Architecture, On Demand Computing, Web services: SOAP versus REST.	10
2	Cloud Computing Technologies:- Introduction to Cloud Computing, Cloud Architecture and Cloud Storage, Characteristics of cloud computing, Components and Organizational scenarios of clouds, Administering and Monitoring cloud services, Benefits and Limitations of cloud computing, Deploy application over cloud: Cloud computing technology, Accessing the cloud, Cloud Applications, Migrating to the Cloud, Software Licenses, Cloud Cost Model, Service Levels for Cloud Applications.	10
3	Web Services and Platforms:- Service Models, Software-as-a-Service, Platform-as-a-Service, Infrastructure -as-a-Service, Process-as-a-Service, Application-as-a-Service, Storage-as-a-Service, Information-as-a-Service, Integration-as-a-Service, Security-as-a-Service, Management/Governance-as-a-Service, Testing-as-a-Service, Comparison among IAAS, PAAS, SAAS Cloud computing platforms	10
4	Cloud Disaster Management:- Cloud Disaster Management, Disaster Recovery, Disaster Recovery Planning, Benefits of a cloud Disaster Recovery service, Disaster Recovery as a Cloud Service, Cloud data Centers, Comparing approaches	10
5	Cloud Computing Security:- Cloud security fundamentals, Vulnerability assessment tool for cloud, Privacy and Security in cloud, Cloud computing security architecture, Cloud computing security challenges	5

Text Books:

1. Cloud Computing Dr.Pandey U.S. & Dr. Chaudhary Kavita S. Chand Publishing.
2. Cloud Computing Miller Pearson Education India.

Reference Books:

1. Learning Amazon Web Services (AWS): A Hands-On Guide to the Fundamentals of AWS Cloud Mark Wilkins-First Edition.
2. "Handbook of cloud computing" by Borko Furht, Armando Escalante published by springer(2010)

Syllabus Semester-III

Course code: MDS41MML603	Course name: Natural Language Processing	Course category: Major Mandatory
Credits: 3	Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basic knowledge of Data structure & Algorithms, Theory of computer science, Probability Theory.		
Course Objectives: To Design, implement and Test algorithm for NLP based application problem.		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Have a broad understanding of the field of natural language processing.		
CO 2: Model linguistic phenomena with formal grammars		
CO 3: Design, implement and test algorithms for NLP problems		
CO 4: Design and implement NLP Based Application		
CO 5: Develop Language Modeling		

Contents –

Unit	Content	Teaching hours
1	Finding the Structure of Words:- Words and Their Components, Tokens, Lexemes, Morphemes, Typology, Issues and Challenges, Irregularity ,Ambiguity, Productivity, Morphological Models, Dictionary Lookup, Finite-State Morphology, Unification-Based Morphology, Functional Morphology, Morphology Induction	10
2	Finding the Structure of Documents:- Introduction, Sentence Boundary Detection, Topic Boundary Detection, Methods, Generative Sequence Classification Methods, Discriminative Local Classification Methods, Discriminative Sequence Classification Methods, Hybrid Approaches, Extensions for Global Modeling for Sentence Segmentation, Complexity of the Approaches, Performances of the Approaches, Features, Processing Stages	10
3	Syntax:- Parsing Natural Language, Treebanks: A Data-Driven Approach to Syntax, Representation of Syntactic Structure, Syntax Analysis Using Dependency Graphs, Syntax Analysis Using Phrase Structure Trees, Parsing Algorithms, Models for Ambiguity Resolution in Parsing, Multilingual Issues: What Is a Token	10
4	Semantic Parsing:- Introduction, Semantic Interpretation, Structural Ambiguity, Word Sense, Entity and Event Resolution ,System Paradigms, Predicate-Argument Structure, Meaning Representation, System Paradigms, Word Sense, Resources, Systems, Software, Predicate-Argument Structure, Resources, Systems, Software, Meaning Representation, Resources, Systems, Software	10
5	Language Modeling:- Introduction, Language Model Adaptation, Language-Specific Modeling Problems, Multilingual and Crosslingual Language Modeling	5

Text Books:

1. Multilingual Natural Language Processing Applications by Daniel M. Bikel Imed Zitouni Published by Pearson plc Publishing as IBM Press.

2. Natural Language Processing with Python Steven Bird, Ewan Klein O'Reilly.

Reference Books: Introduction to Natural Language Processing - A Practical Guide for Beginners by Ansari Sakil White Falcon Publishing.

Online Resources: 1.NPTEL / SWAYAM lectures.

1. <https://archive.nptel.ac.in/courses/106/101/106101007/Web Resources>

Syllabus Semester-III

Course code: MDS41MEL601	Course name: Cloud Services	Course category: Elective
Credits: 3	Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basic knowledge of cloud concepts.		
Course Objectives: Gain foundational knowledge of core data concepts such as relational and nonrelational data, big data, warehousing, real-time analytics, and data visualization.		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Students learn Cloud Services		
CO 2: Working and configurations of Cloud services		
CO 3: Explore relational data concepts and relational database services in Azure		
CO 4: learn nonrelational database fundamentals with Azure Cosmos DB.		
CO 5: Design Application relevant Virtual Networks		

Contents –

Unit	Content	Teaching hours
1	Introduction to Cloud Services: - Getting started with Azure, What is Azure, Overview of cloud computing, Comparison of on-premises versus Azure, Cloud offering, SaaS: Software as a Service, PaaS: Platform as a Service, IaaS: Infrastructure as a Service, IaaS: Infrastructure as a Service, Azure services, Introduction to portals.	10
2	Azure Website: - Azure Websites and Azure Cloud Services, Creating and configuring websites, What is Azure Websites, Creating a new website, Using the portal, Websites gallery, Configure and scale a website, Configuration, Scaling, Deploying and monitoring websites, Options for creating websites, Notepad or an HTML editor, Web Matrix, Visual Studio, Publishing a website from Visual Studio, Monitoring a website, PaaS Cloud Services, Miscellaneous points.	10
3	Virtual Machine:- Azure Virtual Machines, What is Azure Virtual Machines, VM status, IP address, Billing, Create and configure virtual machines, Create a virtual machine with the Azure Preview Portal, Connect to a virtual machine, Configure disks, Caching, Attach a disk, Formatting disks, Disk performance, Endpoints, Virtual machine management, Availability set, Service level agreement, Scalability, Image capture.	10
4	Cloud Storage:- Azure Storage, Blob storage, Azure Files, Table storage, Queues, Redundancy, Creating and managing storage, Create a storage account, Create a container and add blobs, Create a table and add records.	10
5	Virtual Networks: - Azure Virtual Networks, What is a virtual network, Creating a virtual network, Using a network configuration file, Cross-premises connection options, Point-to-site network.	5

Text Books:

1. Fundamentals of Azure by Michael Collier Robin Shahan PUBLISHED BY Microsoft Press ISBN: 978-0-7356-9722-5.
4. Cloud Computing Miller Pearson Education India.

Reference Books:

1. Learning Microsoft Azure by Jonah Carrio Andersson Released November 2023 Publisher(s): O'Reilly Media, Inc. ISBN: 9781098113322

Syllabus Semester-III

Course code: MDS41MEL602	Course name: Data Analytics	Course category: Elective
Credits: 3	Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basics of Programming and Data Science Applications.		
Course Objectives: The objective of this course is to provide comprehensive knowledge of python programming paradigms and Data Analytics.		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Demonstrate the use of built-in objects of Python		
CO 2: Demonstrate significant experience with python program development environment		
CO 3: Implement numerical programming, data handling and visualization through NumPy modules		
CO 4: Implement numerical programming, data handling and visualization through Pandas and Matplotlib modules		
CO 5: Develop Dashboard using Matplotlib		

Contents –

Unit	Content	Teaching hours
1	Introduction To Python:- Structure of Python Program-Underlying mechanism of Module Execution-Branching and Looping-Problem Solving Using Branches and Loops-Functions - Lists and Mutability Problem Solving Using Lists and Functions.	10
2	Data Types:- Sequence Datatypes And Object-Oriented Programming Sequences, Mapping and Sets- Dictionaries- -Classes: Classes and Instances-Inheritance Exceptional Handling-Introduction to Regular Expressions using “re” module.	10
3	Data Analytics Library:- Using NumpyBasics of NumPy-Computation on NumPy-Aggregations-Computation on Arrays Comparisons, Masks and Boolean Arrays-Fancy Indexing-Sorting Arrays-Structured Data: NumPy’s Structured Array.	10
4	Data Manipulation With Pandas:- Introduction to Pandas Objects - Data indexing and Selection - Operating on Data in Pandas -Handling Missing Data - Hierarchical Indexing - Combining Data Sets - Aggregation and Grouping - Pivot Tables.	10
5	Visualization And Matplotlib:- Basic functions of matplotlib - Simple Line Plot, Scatter Plot - Density and Contour Plots -Histograms, Binnings and Density - Customizing Plot Legends, Colour Bars - Three-dimensional Plotting in Matplotlib.	5

Text Books: 1.Jake VanderPlas, Python Data Science Handbook - Essential Tools for Working with Data, O’Reilly Media Inc., 2016.

2. Zhang.Y, An Introduction to Python and Computer Programming, Springer Publications, 2016

Reference Books: 1.Joel Grus, Data Science from Scratch First Principles with Python, O’Reilly Media, 2016.

Syllabus Semester-III

Course code: MDS41MEL603	Course name: Internet of Things	Course category: Elective
Credits: 3	Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basic understanding of Microcontroller and Microprocessors.		
Course Objectives: To understand the state of art Internet of Things.		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Describe the fundamentals of Internet of Things		
CO 2: Understand the state of art design principles and internet principles of IoT		
CO 3: Extend the understanding of sensors, sensor technology, actuators and RFID in IoT		
CO 4: Develop a better understanding of data acquisition, storage, and cloud computing platforms, privacy, and security and vulnerability solutions in IoT.		
CO 5: Develop IOT based Case study		

Contents –

Unit	Content	Teaching hours
1	Overview of Internet of Things:- IoT definition and vision, Characteristics of IoT ,IoT conceptual framework, IoT architectural view, Technology behind IoT, Sources of IoT,M2M communication, Wearable smart watch, an example of IoT, Smart home, an example of IoT, Smart cities, an application of IoT	10
2	Design Principles and Internet Connectivity Principles:- Fundamentals of design principles, Communication technologies, IoT/M2M systems, layers and designs standardization, Data Enrichment, Data Consolidation and Device Management at Gateway, Ease of designing and affordability, Web communication protocol for connected devices, Message communication protocol for connected devices,Web Connectivity for Connected-Devices Network using Gateway, SOAP, REST, HTTP, RESTful and WebSockets, Internet –based communication, IP addressing in IoT	10
3	Sensors, Participatory Sensing, RFIDs, and Wireless Sensor Networks: Sensor technology, Analog sensors, Digital sensors, Participatory Sensing, Industrial IoT, Automotive IoT, Actuators, Sensor data communication protocols, Radio Frequency identification Technology (RFID),Wireless Sensor Network (WSN) concept, WSN architecture	10
4	Data Acquisition, Organization, Storage and Computing as Cloud Platforms: Data acquisition , storage and organization, Cloud computing paradigm for data collection and storage and computing, IoT Cloud-based services and models IoT Privacy, Security and Vulnerabilities Solutions: Defining basic terms associated to privacy & security in IoT, Vulnerabilities, security requirements and threat analysis, Use cases and misuse cases, IoT security tomography and layered attacker model, Identity management, and establishment, Access control and secure message communication, Security models, Profiles and protocols for IoT	10
5	Applications of Internet of Things:- Connected RFIDs Supply Chain Monitoring Project, Customer monitoring in IoT Applications/Services Project, Connected car and its applications and services: Tesla an example, IoT applications for smart homes, cities, environment monitoring and agriculture, Case study: smart city streetlight control and monitoring	5

Text Books:

1. Internet of Things, Architecture and Design Principles Raj Kamal McGraw Hill Education (India) Private Limited.
2. IoT Fundamentals: Networking Technologies, Protocols, and Use Cases for the Internet of Things David Hanes, CISCO Press.

Reference Books:

1. Internet of Things, A Hands on Approach, Arshdeep Bagha, Vijay Madiseti Universities Press Private Limited.

Syllabus Semester-III

Course code: MDS41MEL604	Course name: Block chain	Course category: Elective
Credits: 3	Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Good programming Skill and domain knowledge of Networking & routing.		
Course Objectives: Understand and describe how block chain works and explain the underlying technology of transactions, blocks, proof-of-work, and consensus building.		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Describe block chain technology and its key concepts		
CO 2: Bit coin transaction processes		
CO 3: Ethereum deployment tools		
CO 4: business network using Hyperledger Composer		
CO 5: Develop Enterprise Blockchain Application		

Contents –

Unit	Content	Teaching hours
1	Introduction to Blockchain:- Backstory of Blockchain, What is Blockchain, Centralized vs. Decentralized Systems, Layers of Blockchain, Why is Blockchain Important, Blockchain Uses and Use Cases, How Blockchain Works, Laying the Blockchain, Foundation, Cryptography, Game Theory, Computer Science Engineering, Properties of Blockchain Solutions, Blockchain Transactions, Distributed, on senses, Mechanisms, Blockchain Applications, Scaling Blockchain.	10
2	Bitcoin Mechanics and Optimizations:- How Bitcoin Works: The History of Money, Dawn of Bitcoin, TheBitcoin Blockchain, The Bitcoin Network, Bitcoin Scripts, Full Nodes vs. SPVs, Bitcoin Wallets.	10
3	Smart Contracts:- How Ethereum Works, Bitcoin to Ethereum, Enter the Ethereum Blockchain, Ethereum, Smart Contracts, Ethereum Virtual Machine and Code Execution, Ethereum Ecosystem.	10
4	Scaling Blockchain:- Decentralized Applications, Blockchain, Application Development, Interacting with the Bitcoin Blockchain, Interacting, Programmatically with Ethereum Sending, Transactions, Interacting Programmatically, with Ethereum Creating a Smart Contract, Interacting Programmatically with Ethereum Executing Smart, Contract, Functions, Blockchain Concepts Revisited, Public vs. Private Blockchains, Decentralized Application Architecture.	10
5	Enterprise Blockchain:- Building an Ethereum DApp: The DApp, Setting Up a Private Ethereum Network, Creating the Smart Contract, Deploying the Smart Contract, Client Application.	5

Text Books:

1. Beginning Blockchain ABeginner’s Guide to Building Blockchain Solutions, Bikramaditya Singhal,Gautam Dhameja, PriyansuSekhar PandaApress.
2. “Mastering blockchain” by Imran bashir published expert insight.

Reference Books:

-
1. Bitcoin and Cryptocurrency Technologies Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller, and Steven Goldfeder.

Syllabus Semester-III

Course code: MDS41MMP601 Course name: Practical Based on Big Data Course category: Major Mandatory		
Credits: 1	Teaching scheme: L-0 P-2	Evaluation scheme: CA–30, ESE–20
Pre-requisites: Data science concepts and applications.		
Course Objectives: Understand the Big Data Platform and its Use cases.		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Identify Big Data and its Business Implications.		
CO 2: List the components of Hadoop and Hadoop Eco-System		
CO 3: Access and Process Data on Distributed File System		
CO 4: Manage Job Execution in Hadoop Environment		
CO 5: Developing MapReduce Application		

List of Practical's:

Sr.No.	Title of the Experiment	Practical hours
1	Install and Configure Apache Hadoop	2
2	Write Script for MapReduce program to calculate the frequency of a given word in a given file.	2
3	Write Script for MapReduce program to find the maximum temperature in each year.	2
4	Write Script for MapReduce program to find the grades of student's.	2
5	Write Script for MapReduce program to implement Matrix Multiplication.	2
6	Write Script for MapReduce to find the maximum electrical consumption in each year given electrical consumption for each month in each year.	2
7	Write Script for MapReduce to analyze weather data set and print whether the day is shiny or cool day.	2
8	Write Script for MapReduce program to find the number of products sold in each country by considering sales data containing fields like Transaction _Date Product Price Payment _Type Name City State Country Account _Created Last_Login Latitude Longitude	2
9	Develop a MapReduce program to find the tags associated with each movie by analyzing movie lens data.	2
10	<p>XYZ.com is an online music website where users listen to various tracks, the data gets collected which is given below. The data is coming in log files and looks like as shown below. UserId TrackId Shared Radio Skip 111115 222 0 1 0 111113 225 1 0 0 111117 223 0 1 1 111115 225 1 0 0</p> <p>Write a MapReduce program to get the following</p> <ul style="list-style-type: none"> • Number of unique listeners and Number of times the track was shared with others • Number of times the track was listened to on the radio • Number of times the track was listened to in total 	2

	• Number of times the track was skipped on the radio	
11	Project	10

Text Books: 1.Hadoop: The Definitive Guide Tom WhiteO'reily Media, Third Edit on, 2012.

2. Big Data Analytics Seema Acharya, Subhasini Chellappan,Wiley 2015

Reference Books: 1.“Big Data and Business Analytics” Jay Liebowitz Auerbach Publications, CRC press (2013)

Syllabus Semester-III

Course code: MDS41MMP602 Course name: Practical Based on Cloud Computing Course category: Major Mandatory		
Credits: 1	Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Pre-requisites: Basic Knowledge of networking & Operating system.		
Course Objectives: Understanding basics of cloud computing and Key concepts of Cloud Analytics and different cloud computing services.		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Understand Cloud Computing Basics		
CO 2: Understand the use of Cloud Computing in Data Analytics		
CO 3: Learn the Concept of Cloud Infrastructure		
CO 4: Understand Business imperative of Cloud Computing		
CO 5: Learn Cloud computing Security Components		

List of Practical's:

Sr.No.	Title of the Experiment	Practical hours
1	Install Virtual box/VMware Workstation with different flavors of Linux or windows Operating System	2
2	Install a C compiler in the virtual machine created using virtual box and execute Simple Programs	2
3	Install Google App Engine. Create hello world app and other simple web applications using python/java.	2
4	Use Google App Engine launcher to launch the web applications.	2
5	Simulate a cloud scenario using CloudSim and run a scheduling algorithm that is not present in CloudSim	2
6	Find a procedure to transfer the Text files from one virtual machine to another virtual machine	2
7	Find a procedure to transfer the Image files from one virtual machine to another virtual machine	2
8	Find a procedure to transfer the Video files from one virtual machine to another virtual machine	2
9	Find a procedure to launch virtual machine using trystack (Online Openstack Demo Version)	2
10	Install Hadoop single node cluster and run simple applications like wordcount.	2
11	Project	10

Reference Book / Hand Books/ Lab Manual

1. Cloud Computing Dr. Pandey U.S. & Dr. Chaudhary Kavita S. Chand Publishing.
2. Cloud Computing Miller Pearson Education India.
3. Learning Amazon Web Services (AWS): A Hands-On Guide to the Fundamentals of AWS Cloud Mark Wilkins-First Edition.

Syllabus Semester-III

Course code: MDS41MMP603	Course name: Practical Based on Natural Language Processing	
Course category: Major Mandatory		
Credits: 1	Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Pre-requisites: Data structure & Algorithms, Theory of computer science, Probability Theory.		
Course Objectives: To Design, implement and Test algorithm for NLP based application problem.		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Have a broad understanding of the field of natural language processing.		
CO 2: Model linguistic phenomena with formal grammars		
CO 3: Design, implement and test algorithms for NLP problems		
CO 4: Design and implement NLP Based Application		
CO 5: Develop Language Modeling		

List of Practical's:

SR.No.	Title of the Experiment	Practical hours
1	Install and Setup Python Programming	2
2	Read Text Dataset and Implement Pre-Processing text(Word and Sentence Tokenization) using TextBlob	2
3	Implementation of Morphological analysis (Stop word removal) using Natural Language Toolkit	2
4	Implementation of Stemming on Unstructured Data using CoreNLP	2
5	Implementation of Lemmatization on Unstructured Data using Gensim	2
6	Implementation of Parts of Speech Tagging in a sentence using spaCy	2
7	Implementation of Parts of Speech Chunking in a sentence using spaCy	2
8	Implementation of NER NameEntity Recognition using scikit-learn	2
9	Implement N-Gram LanguageModel using scikit-learn	2
10	Implement Wordnet-Lesk Algorithm for detecting shared vocabulary between the definitions of words using scikit-learn	2
11	Project	10

Reference Book / Hand Books/ Lab Manual

- Multilingual Natural Language Processing Applications by Daniel M. Bikel Imed Zitouni Published by Pearson plc Publishing as IBM Press.
- Natural Language Processing with Python Steven Bird, Ewan Klein O'Reilly.
- <https://archive.nptel.ac.in/courses/106/101/106101007/Web Resources>.

Syllabus Semester-III

Course code: MDS41MEP601	Course name: Practical Based on Cloud Services	
Course category: Elective		
Credits: 1	Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Pre-requisites: Cloud Concepts.		
Course Objectives: Gain foundational knowledge of core data concepts such as relational and nonrelational data, big data, warehousing, real-time analytics, and data visualization.		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Students learn Cloud Services		
CO 2: Working and configurations of Cloud services		
CO 3: Explore relational data concepts and relational database services in Azure		
CO 4: learn nonrelational database fundamentals with Azure Cosmos DB.		
CO 5: Design Application relevant Virtual Networks		

List of Practical's:

Sr.No.	Title of the Experiment	Practical hours
1	Installation and configuration of own Cloud	2
2	Implementation of Virtualization in Cloud Computing to Learn Virtualization Basics, Benefits of Virtualization in Cloud using Open Source Operating System	2
3	Implementation of infrastructure as Service using Open Stack.	2
4	Deployment of OpenStack using DevStack	2
5	Write a program for Web feed using PHP and HTML	2
6	Write a Program to Create, Manage and groups User accounts in own Cloud by Installing Administrative Features.	2
7	Assignment to install and configure Google App Engine.	2
8	Design an Assignment to retrieve, verify, and store user credentials using Firebase Authentication, the Google App Engine standard environment, and Google Cloud Data store.	2
9	Design an Assignment based on Working with Mangrasoft AnekaSoftware	2
10	Creating a virtual network, using a network configuration file	2
11	Project	10

Reference Book / Hand Books/ Lab Manual

1.Fundamentals of Azure by Michael Collier Robin Shahan PUBLISHED BY Microsoft Press ISBN: 978-0-7356-9722-5

2.Cloud Computing: Concepts, Technology & Architecture" by Thomas Erl, Ricardo Puttini, and Zaigham Mahmood published by Prentice Hall ISBN-13: 978-0133387520

Syllabus Semester-III

Course code: MDS41MEP602 Course name: Practical Based on Data Analytics Course category: Elective		
Credits: 1	Teaching scheme: L-0 P-2	Evaluation scheme: CA–30, ESE–20
Pre-requisites: Basics of Programming and Data Science Applications.		
Course Objectives: The objective of this course is to provide comprehensive knowledge of python programming paradigms and Data Analytics.		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Demonstrate the use of built-in objects of Python		
CO 2: Demonstrate significant experience with python program development environment		
CO 3: Implement numerical programming, data handling and visualization through NumPy modules		
CO 4: Implement numerical programming, data handling and visualization throughPandas and Matplotlib modules		
CO 5: Develop Dashboard using Matplotlib		

List of Practical's:

Sr.No.	Title of the Experiment	Practical hours
1	Installing Python and Shell or Notebook, Launching the IPythonShell, Launching the Jupiter Notebook	2
2	Passing Values to and from the Shell, Shell-Related Magic Command, Errors and Debugging, Controlling Exceptions: %mode	2
3	Aggregations: Min, Max, and Everything in Between, Summing the Values in an Array, Minimum and Maximum, Example: What Is the Average Height of US Presidents	2
4	Sorting Arrays, Fast Sorting in NumPy: np.sort and np.argsort, Partial Sorts:Partitioning, Example: k-Nearest Neighbors	2
5	Data Indexing and Selection, Data Selection in Series, Data Selection in DataFrame, Operating on Data in Pandas, Ufuncs: Index Preservation, UFuncs: Index Alignment	2
6	Handling Missing Data, Trade-Offs in Missing Data Conventions, Missing Data in Pandas, Operating on Null Values	2
7	Vectorized String Operations , Introducing Pandas String Operations,Tables of Pandas String Methods, Example: Recipe Database	2
8	General Matplotlib Tips, Importing matplotlib, Setting Styles, show() or Noshow() How to Display Your Plots, Saving Figures to File	2
9	Text and Annotation, Example: Effect of Holidays on US Births,Transforms and Text Position, Arrows and Annotation, Customizing Ticks,Major and Minor Ticks, Hiding Ticks or Labels, Reducing or Increasing theNumber of Ticks	2
10	Three-Dimensional Plotting in Matplotlib, Three-Dimensional Points andLines, Three-Dimensional Contour Plots	2
11	Project	10

Reference Book / Hand Books/ Lab Manual

1. Jake VanderPlas, Python Data Science Handbook - Essential Tools for Working with Data, O'Reilly Media Inc., 2016.
2. Zhang.Y, An Introduction to Python and Computer Programming, Springer Publications, 2016
3. Joel Grus , Data Science from Scratch First Principles with Python, O'Reilly Media, 2016.

Syllabus Semester-III

Course code: MDS41MEP603	Course name: Practical Based on Internet of Things	Course category: Elective
Credits: 1	Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Pre-requisites: Basic understanding of Microcontroller and Microprocessors.		
Course Objectives: To understand the state of art Internet of Things.		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Describe the fundamentals of Internet of Things		
CO 2: Understand the state of art design principles and internet principles of IoT		
CO 3: Extend the understanding of sensors, sensor technology, actuators and RFID in IoT		
CO 4: Develop a better understanding of data acquisition, storage, Cloud computing platforms, privacy, and security and vulnerability solutions in IoT.		
CO 5: Develop IOT based Case study		

List of Practical's:

Sr.No.	Title of the Experiment	Practical hours
1	To study of Arduino Uno Microcontrollers and Arduino IDE	2
2	To demonstrate blinking of LED using Arduino Uno	2
3	To demonstrate sensing the available networks using Arduino	2
4	To demonstrate the use of ultrasonic sensor integrating with Arduino board	2
5	To demonstrate the use of ultrasonic sensor integrating with Arduino board	2
6	To demonstrate the use of vibration sensor integrating with Arduino board	2
7	To establish the connection with the available Wi-Fi using Arduino and wifi module	2
8	Demonstrate the touch sensor using Arduino board Demonstrate the use of temperature and humidity sensor integrating with Arduino board	2
9	Demonstrate the working of LDR sensor integrating with Arduino board	2
10	Study on Amazon web Services for IoT Study on Cloud Storage models & Communication APIs	2
11	Project	10

Reference Book / Hand Books/ Lab Manual

- Internet of Things, Architecture and Design Principles Raj Kamal McGraw Hill Education (India) Private Limited.
- IoT Fundamentals: Networking Technologies, Protocols, and Use Cases for the Internet of Things David Hanes, CISCO Press.
- Internet of Things, A Hands on Approach Arshdeep Bagha, Vijay Madiseti Universities Press Private Limited.

Syllabus Semester-III

Course code: MDS41MEP604 Course name: Practical Based on Block chain Course category: Elective		
Credits: 1	Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Pre-requisites: Good programming Skill and domain knowledge of Networking & routing.		
Course Objectives: Understand and describe how blockchain works and explain the underlying technology of transactions, blocks, proof-of-work, and andconsensus building.		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Describe blockchain technology and its key concepts		
CO 2: Bitcoin transaction processes		
CO 3: Ethereum deployment tools		
CO 4: business network using Hyperledger Composer		
CO 5: Develop Enterprise Blockchain Application		

List of Practical's:

Sr.No.	Title of the Experiment	Practical hours
1	Install and Configure Ethereum for Block chain applications Create a list of those applications and the industries/businesses they are impacting	2
2	Explore the bitcoin blockchain on blockchain.info	2
3	Use an online service to generate hashes for content	2
4	Build a transaction and then hash it. Generate public and private keys. Digitally sign a transaction	2
5	Explore the bitcoin blockchain on blockchain.info for block generation. Explore how long it takes a block to be confirmed.	2
6	Use an online service to illustrate how consensus is built in a distributed system with no central authority.	2
7	What is the computing power needed to mine and generate bitcoin? Explore if miner pools are dominating bitcoin mining. Compare incentives from mining activity vs transaction fees.	2
8	Install a bitcoin wallet. Generate and secure your private key. Send a small transaction amount (to be monetized by instructor) to another user. Track the transaction through blockchain. Verify the confirmation and commitment of the transaction to the bitcoin blockchain.	2
9	Pick three industries. Research the application of blockchain in those industries. Describe how blockchain could be successful in those industries.	2
10	Deploy a business network using Hyperledger Composer	2
11	Project	10

Reference Book / Hand Books/ Lab Manual

1.Beginning Blockchain ABeginner’s Guide to Building Blockchain Solutions

Bikramaditya Singhal,GautamDhameja,PriyansuSekhar PandaApress,ISBN-13(pbk): 978-1-4842-3443-3

2.Bitcoin and Cryptocurrency Technologies Arvind Narayanan,Joseph Bonneau, Edward Felten, Andrew Miller,and Steven Goldfeder

Semester: FOURTH

Syllabus Semester-IV

Course code: MDS41MML604	Course name: Artificial Intelligence	Course category: Major Mandatory
Credits: 3	Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basic knowledge of algorithms & data structures.		
Course Objectives: Understand basic principles, techniques, and applications of Artificial Intelligence and Problem solving, inference, perception, knowledge representation, and learning.		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Understanding of the history of artificial intelligence (AI) and its foundations.		
CO 2: Apply basic principles of AI in solutions that require problem solving, inference		
CO 3: Perception, knowledge representation, and learning.		
CO 4: Demonstrate awareness and a fundamental understanding of various applications of AI		
CO 5: Develop Knowledge based System		

Contents –

Unit	Content	Teaching hours
1	Introduction:- Introduction to Artificial Intelligence, Historical Background, What is Intelligence? Depth First Search (DFS), Breadth First Search (BFS), Comparison of Depth First Search and Breadth First Search, Quality of Solution	10
2	Heuristic Search:- Heuristic Functions, Best First Search, Hill Climbing, Local Maxima, Solution Space Search, Variable Neighbourhood Decent, Beam Search, Tabu Search, Peak to Peak Method.	10
3	Randomized Search and Emergent Systems:- Iterated Hill Climbing, Simulated Annealing, Genetic Algorithm, The Travelling Salesman Problem, Neural Network, Emergent System, Ant Colony Optimization	10
4	Finding Optimal Paths:- Brute Force, Branch and Bound, Refinement Search, Dijkstra's Algorithm, Algorithm of A*, Admissibility of A*, Recursive Best First Search (RBFS), Pruning the CLOSED List, Pruning the OPEN List, Divide and Conquer Beam Stack Search	10
5	Knowledge Representation:- Knowledge Representation, Representation and Mapping, Approaches to Knowledge Representation, Issues in Knowledge Representation, Predicate Logic, Representing Knowledge Using Rules	5

Text Books:

1. A First Course in Artificial Intelligence Deepak Khemani McGraw-Hill Education.
2. Artificial Intelligence A Modern Approach by Stuart Russell, Peter Norvig, third edition.

- Reference Books:** 1. Artificial Intelligence Elaine Rich Tata McGraw-Hill Education
2. Artificial Intelligence: A Modern Approach" by Stuart Russell and Peter Norvig Publisher: Pearson ISBN-13: 978-0134610993

Syllabus Semester-IV

Course code: MDS41MML605	Course name: Development Operations	Course category: Major Mandatory
Credits: 3	Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basic knowledge of Cloud Computing.		
Course Objectives: learn how these tools are used in software development operations		
Course Outcomes: At the end of the course, the students will be able to -		
CO 1: Understand the Fundamentals of DevOps		
CO 2: Lean DevOps Capabilities		
CO 3: Adopt DevOps resources and infrastructures		
CO 4: How DevOps Works with real time Application		
CO 5: Develop Multiple-Tier Applications		

Contents –

Unit	Content	Teaching hours
1	What Is DevOps:- Understanding the Business Need for DevOps, Recognizing the Business Value of DevOps, Enhanced customer experience, Increased capacity to innovate, Faster time to value, How DevOps Works, Develop and test against production-like systems, Deploy with repeatable, reliable processes, Monitor and validate operational quality, Amplify feedback loops.	10
2	Looking at DevOps Capabilities:- Paths to DevOps Adoption, Steer, Develop/Test, Collaborative development, Continuous testing, Deploy, Operate, Continuous monitoring, Continuous customer feedback and optimization.	10
3	Adopting DevOps:- Where to Begin, Identifying business objectives, Identifying bottlenecks in the delivery pipeline, People in DevOps, DevOps culture, DevOps team, Process in DevOps, DevOps as a business process, Change management process, DevOps techniques, Technology in DevOps, Infrastructure as code, Delivery pipeline, Deployment automation and release management.	10
4	How Cloud Accelerates DevOps:- Using Cloud as an Enabler for DevOps, Full-Stack Deployments, Choosing a Cloud Service Model for DevOps, Understanding What a Hybrid Cloud.	10
5	DevOps to Solve New Challenges:- Mobile Applications, ALM Processes, Scaling Agile, Multiple-Tier Applications, DevOps in the Enterprise, Supply Chains, The Internet of Things	5

Text Books:

1. DevOps By Sanjeev Sharma and Bernie Coyne 2nd IBM Limited Edition.
2. engineering Devops by marc hoenbeek.

Reference Books: 1. Effective DevOps By Jennifer Davis and Ryn Daniels O'Reilly

Syllabus Semester-IV

Course code: MDS41MML606 Course name: Recommendation System Course category: Major Mandatory		
Credits: 2	Teaching scheme: L-3 P-0	Evaluation scheme: CA–60, ESE–40
Pre-requisites: Basic knowledge of Machine Learning.		
Course Objectives: Study various recommender system and to generate relevant recommendations.		
Course Outcomes: At the end of the course, the students will be able to -,		
CO1: Familiarize with recommender systems and their applications		
CO2: Analyze the different approaches towards recommendation		
CO3: Evaluate the effectiveness of recommender system		
CO4: Design recommender system		

Contents –

Unit	Content	Teaching hours
1	An Introduction to Recommender Systems:- Introduction, Goals of Recommender Systems, The Spectrum of Recommendation Applications, Basic Models of Recommender Systems, Collaborative Filtering Models, Content-Based Recommender Systems.	6
2	Knowledge-Based Recommender Systems:- Demographic Recommender Systems, Hybrid and Ensemble-Based Recommender Systems, Domain-Specific Challenges in Recommender Systems, Advanced Topics and Applications	8
3	Neighborhood-Based Collaborative Filtering:- Introduction, Key Properties of Ratings Matrices, Predicting Ratings with Neighborhood-Based Methods, User-Based Neighborhood Models, Item-Based Neighborhood Models, Clustering and Neighborhood-Based Methods, Dimensionality Reduction and Neighborhood Methods, A Regression Modeling View of Neighborhood Methods, Graph Models for Neighborhood-Based Methods	8
4	Model-Based Collaborative Filtering:- Introduction, Decision and Regression Trees, Rule-Based Collaborative Filtering, Naive Bayes Collaborative Filtering, Using an Arbitrary Classification Model as a Black-Box, Latent Factor Models, Integrating Factorization and Neighborhood Models.	8

Text Books:

- 1.Recommender Systems By Charu C. Aggarwal ISBN 978-3-319-29657-9
2. recommender Systems By Francesco Ricci · Lior Rokach · Bracha Shapira · Paul B. Kantor.

Reference Books: 1.Recommender Systems Francesco Ricci · Lior Rokach · Bracha Shapira · Paul B. Kantor ISBN 978-0-387-85819-7 Springer New York Dordrecht Heidelberg London

Syllabus Semester-IV

Course code: MDS41MEL605	Course name: Exploratory Data Analysis	Course category: Elective
Credits: 3	Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basic knowledge of Data Mining.		
Course Objectives: Exploratory data analysis, students will learn how to perform an initial analysis on data.		
Course Outcomes: At the end of the course, the students will be able to -		
CO1: Learn EDA tool installations and configurations		
CO2: Understand concepts of Data Explorations		
CO3: Explore Exploratory Data Analysis through Unstructured		
CO4: Design Dashboard for Sales prediction dataset		
CO5: Use developed Dashboard for decision process		

Contents –

Unit	Content	Teaching hours
1	Introduction to Exploratory Data Analysis:- What is exploratory data analysis, Why is exploratory data analysis important in data science, Exploratory data analysis tools, Types of exploratory data analysis, Exploratory Data Analysis Tools. R Programming Installation:- Introduction, Installation of R and R Studio, Managing Data Frames with the dplyr package, Common dplyr Function Properties, Installing the dplyr package, Exploratory Data Analysis Checklist, Principles of Analytic Graphics.	10
2	Data Explorations:- Exploratory Graphs, Characteristics of exploratory graphs, One Dimensions: Five Number Summary, Boxplots, Barplot, Histograms, Density plot, Two Dimensions: Multiple or overlaid 1-D plots, Scatterplots, Smooth scatterplots, Overlaid or multiple 2-D plots; conditioning plots (coplots), Use color, size, shape to add dimensions, Spinning/interactive plots, Actual 3-D plots.	10
3	Plotting Systems:- The Base Plotting System, The Lattice System, and The ggplot2 System. Graphics Devices: The Process of Making a Plot, How Does a Plot Get Created, Graphics File Devices, Multiple Open Graphics Devices, Copying Plots.	10
4	The Base Plotting System: Base Graphics, Simple Base Graphics, Histogram, Boxplot, Scatterplot, Important Base Graphics Parameters, Base Plotting Functions, Base Plot with Regression Line, Multiple Base Plots.	10
5	Plotting and Color in R: Colors 1, 2, and 3, Connecting colors with data, Color Utilities in R, RColorBrewer Package, Using the RColorBrewer palettes, The smoothScatter() function, Adding transparency.	5

Text Books:

1. Exploratory Data Analysis with R By Roger D. Peng.
2. Exploratory Data Analysis by Matthieu Kumorowski, Justin D Saliccioli.

Reference Books: 1.Hands-On Exploratory Data Analysis with Python by Suresh Kumar Mukhiya, Usman Ahmed Publisher(s): Packt Publishing ISBN: 9781789537253

Syllabus Semester-IV

Course code: MDS41MEL606	Course name: Data Centre Technologies	Course category: Elective
Credits: 3	Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basics knowledge of Computer Networks and Data Networking.		
Course Objectives: A data center is a facility that centralizes an organization's IT operations and equipment for the purposes of storing, processing and disseminating data and applications.		
Course Outcomes: At the end of the course, the students will be able to -		
CO1: Learn Data Center concepts		
CO2: Install and Configuration steps of Data Centers		
CO3: Identify the Data center platforms		
CO4: Check Availability of suitable Hardware for Data center		
CO5: Deploy Developed Data center		

Contents –

Unit	Content	Teaching hours
1	Introduction to Data Center:- Smarter data center, Technology trends, Data center evolution, Data center infrastructure, Data center network reference architecture, Next-generation data center, IBM addressed market segments.	10
2	Product introduction:- Product overview, Product description of the IBM b-type m-series Ethernet/IP Routers, Product description of the IBM b-type r-series Ethernet Switches, Product description of the IBM b-type c-series Ethernet Switches, Product description of the IBM b-type s-series Ethernet Switches, Product description of the IBM b-type y-series Ethernet Switches, Stacking y-series Ethernet Switches, IBM b-type documentation	10
3	Networking introduction: Brief network history, Network communications, Introduction to data communications, ISO OSI communications model, Layer 1 - Physical Layer, Layer 2 - Data Link Layer, TRILL, Layer 3 - Network Layer, VPLS, Layer 4 to Layer 7.	10
4	Making the network highly available:- Hardware resiliency, Design for redundancy, Device-level redundancy, Hitless upgrades and hitless switchover, Infrastructure resiliency, Layer 1 availability features, Layer 1 availability features, Layer 2 availability features, Layer 3 availability features, Power over Ethernet redundancy.	10
5	Quality of Service:- QoS introduction, Why QoS is used, QoS architecture, QoS components, QoS on b-type networking products	5

Text Books:

1. IBM b-type Data Center Networking: Design and Best Practices Introduction By Jon Tate Norman Bogard, Michal Holenia, Sebastian Oglaza, Steven Tong.
2. data center fundamentals by Mauricio aggoes.

Reference Books: 1. Cisco Data Center Fundamentals By Somit Maloo, Iskren Nikolov

Published Sep 28, 2022 by Cisco Press. Part of the Networking Technology series.

Syllabus Semester-IV

Course code: MDS41MEL607	Course name: Social Media Analytics	Course category: Elective
Credits: 3	Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basic knowledge of Data Analytics.		
Course Objectives: Understand the role of social media data and analytics in helping organizations achieve their goals and understand their publics.		
Course Outcomes: At the end of the course, the students will be able to -		
CO1: To know the Users on Social Media		
CO2: Understand the Social Media Networks		
CO3: Social Media Configurations		
CO4: Design Content on Social Media		
CO5: Apply Social media tools on Processing Large Datasets		

Contents –

Unit	Content	Teaching hours
1	Users: The Who of Social Media:- Measuring Variations in User Behavior in Wikipedia, The Diversity of User Activities, The Long Tail in Human Activities, Long Tails Everywhere: The 80/20 Rule (p/q Rule), Online Behavior on Twitter, Retrieving Tweets for Users, Logarithmic Binning, and User Activities on Twitter.	10
2	Networks: The How of Social Media:- Introduction, Types and Properties of Social Networks, When Users Create the Connections: Explicit Networks, Directed Versus Undirected Graphs, Node and Edge Properties, Weighted Graphs, Creating Graphs from Activities: Implicit Networks, Visualizing Networks, Degrees: The Winner Takes All, Counting the Number of Connections, The Long Tail in User Connections, Beyond the Idealized Network Model, Capturing Correlations: Triangles, Clustering, and Assortativity.	10
3	Temporal Processes: The When of Social Media:- Traditional Models, Inter-Event Times, Comparing to a Memoryless Process, Autocorrelations, Deviations from Memorylessness, Periodicities in Time in User Activities, Bursty Activities of Individuals, Correlations and Bursts, Forecasting Metrics in Time, Finding Trends.	10
4	Content: The What of Social Media:- Defining Content: Focus on Text and Unstructured Data, The Basics of Natural Language Processing, The Basic Statistics of Term Occurrences in Text, Using Content Features to Identify Topics, The Popularity of Topics, How Diverse Are Individual Users' Interests, Extracting Low-Dimensional Information from High-Dimensional Text, Topic Modeling	10
5	Processing Large Datasets:- MapReduce: Structuring Parallel and Sequential Operations, Counting Words, Multi-Stage MapReduce Flows, Merging Data Streams, Joining Against Small Datasets, Patterns in MapReduce Programming, Incremental MapReduce Jobs, Temporal MapReduce Jobs, Challenges with Processing Long-Tailed Social Media Data. Executing on a Hadoop Cluster (Amazon EC2)	5

Text Books:

- 1.Social Media Data Mining and Analytics by Gabor Szabo,Gungor Polatkan Oscar Boykin,Antonios Chalkiopoulos Published byJohn Wiley & Sons, IncISBN: 978-1-118-82485-6.
- 2.social media analytics by Matthews ganis, avinash kohirkar.

Reference Books: 1.Social Media Analytics By Matthew Ganis Avinash Kohirkar IBM Press
Pearson plc ISBN-13: 978-0-13-389256-7

Syllabus Semester-IV

Course code: MDS41MEL608	Course name: Deep Learning	Course category: Elective
Credits: 3	Teaching scheme: L-3 P-0	Evaluation scheme: CA-60, ESE-40
Pre-requisites: Basic knowledge of Machine Learning.		
Course Objectives: Deep learning techniques have been used successfully for variety of applications.		
Course Outcomes: At the end of the course, the students will be able to -		
CO1: Basics of Machine Learning		
CO2: Learn Neural Network and its Applications		
CO3: Understand the need of Deep Learning		
CO4: Apply Deep Networks		
CO5: Develop Tuning Deep Network		

Contents –

Unit	Content	Teaching hours
1	Introduction: -The Learning Machine, How Can Machine Learn, Biological Inspiration, What is Deep Learning. Math, Behind Machine Learning: Scalars, Vectors, Matrices, Tensors, Hyperplanes, Probability, Conditional Probability, Posterior Probability, Distribution, Samples Verses, Population.	10
2	Foundation of Neural Network and Deep Learning: Neural Network, The Biological Neuron, The Perceptron, Multilayer Feed-Forward Networks, Training Neural, Network, Activation Function: Linear, Sigmoid, Tanh, Hard Tanh, Softmax, Loss Functions, Hyperparameters.	10
3	Fundamentals of Deep Networks: -Unsupervised Pretrained Networks, Deep Belief Network, Convolution Neural Networks (CNNs): Biological Inspiration, Intuition, Input Layer. Recurrent Neural Network, Modeling the Time Dimension. Recursive Neural Networks.	10
4	Building Deep Network: -Matching Deep Networks to the Right Problem, Columnar Data and Multilayer Perceptron's, The DL4J Suite of Tools, Basic Concept of DL4J API, Modelling CSV Data with Multilayer Perceptron Networks, Modelling Handwritten Images Using CNNs.	10
5	Tuning Deep Network: -Basic Concept in Tuning Deep Network, Matching Input Data and Network Architectures, Relating Model Goal and Output Layers, Working with Layer Count, Parameter Count and Memory, Feed-Forward Multilayer Neural Networks, Controlling Layer and Parameter Counts.	5

Text Books:

1. Deep Learning Ian Good fellow, Yoshua Bengio, Aaron Courville MIT Press.
2. Deep learning by Ian Goodfellow, Yoshua Bengio, Aaron Courville.

Reference Books: 1. Deep Learning: A Practitioner's Approach Josh Patterson, Adam Gibson O'Reilly Media, Inc.

Syllabus Semester-IV

Course code: MDS41MMP604	Course name: Practical Based on Artificial Intelligence	Course category: Major Mandatory
Credits: 1	Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Pre-requisites: Basic knowledge of algorithms & data structures.		
Course Objectives: Understand basic principles, techniques, and applications of Artificial Intelligence and Problem solving, inference, perception, knowledge representation, and learning.		
Course Outcomes: At the end of the course, the students will be able to -		
CO1: Understanding of the history of artificial intelligence (AI) and its foundations.		
CO2: Apply basic principles of AI in solutions that require problem solving, inference		
CO3: Perception, knowledge representation, and learning.		
CO4: Demonstrate awareness and a fundamental understanding of various applications of AI		
CO5: Develop Knowledge based System		

List of Practical:

Sr.No.	Title of the Experiment	Practical hours
1	Install and configure the Python Program to implement the Breadth First Search algorithm in Python	2
2	Program to implement the Depth First Search algorithm in Python	2
3	Design a program to solve water jug problem and implement it in python	2
4	Design a program to solve the Tower of Hanoi problem using Python	2
5	Design a program to solve Best First Search Algorithm using Heuristic Search techniques	2
6	Design a program to implement Randomized Search	2
7	Program to implement solution of The Travelling Salesman Problem	2
8	Program to implement Ant Colony Optimization Technique	2
9	Design solution for shortest path for finding two geographical locations	2
10	Design solution for Knowledge Representation using Predicate Logic	2
11	Design Project on AI Applications	10

Reference Book / Hand Books/ Lab Manual

1.A First Course in Artificial Intelligence Deepak Khemani McGraw-Hill
Education

2.Artificial Intelligence Elaine Rich Tata McGraw-Hill Education
Artificial Intelligence: A Modern Approach" by Stuart Russell and Peter NorvigPublisher: Pearson
ISBN-13: 978-0134610993

Syllabus Semester-IV

Course code: MDS41MMP605	Course name: Practical Based on Development Operations	
Course category: Major Mandatory		
Credits: 1	Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Pre-requisites: Basic knowledge of Cloud Computing.		
Course Objectives: learn how these tools are used in software development operations.		
Course Outcomes: At the end of the course, the students will be able to -		
CO1: Understand the Fundamentals of DevOps		
CO2: Lean DevOps Capabilities		
CO3: Adopt DevOps resources and infrastructures		
CO4: How DevOps Works with real time Application		
CO5: Develop Multiple-Tier Applications		

List of Practical:

Sr. No.	Title of the Experiment	Practical hours
1	Install and Configure DevOps platforms(GitLab, GitHub, Bitbucket)	2
2	Exploring Git Commands through Collaborative Coding.	2
3	Implement GitHub Operations	2
4	Implement BitBucket Operations	2
5	Demonstrate Maven Build Life Cycle	2
6	Applying CI/CD Pipelines to Web Development Using Jenkins, Git, and Local HTTP Server.	2
7	Create the GitHub Account to demonstrate CI/CD pipeline using Cloud Platform.	2
8	Applying CI/CD Pipelines to Web Development Using Jenkins, Git, using Docker Containers	2
9	Exploring Containerization and Application Deployment with Docker	2
10	Demonstrate Container Orchestration using Kubernetes.	2
11	Project	10

Reference Book / Hand Books/ Lab Manual

- 1.DevOps By by Sanjeev Sharma andBernie Coyne2nd IBM Limited Edition
- 2 .Effective DevOps By Jennifer Davis and Ryn Daniels O'Reilly

Syllabus Semester-IV

Course code: MDS41MEP605 Course name: Practical Based on Exploratory Data Analysis Course category: Elective		
Credits: 1	Teaching scheme: L-0 P-2	Evaluation scheme: CA-30, ESE-20
Pre-requisites: Basic knowledge of Data Mining.		
Course Objectives: exploratory data analysis, students will learn how to perform an initial analysis on data.		
Course Outcomes: At the end of the course, the students will be able to -		
CO1: Learn EDA tool installations and configurations		
CO2: Learn concepts of Data Explorations		
CO3: Learn various plotting systems		
CO4: learn Data analysis and Visualization		
CO5: Use developed Dashboard for decision process		

List of Practical:

Sr. No.	Title of the Experiment	Practical hours
1	Installing the dplyr package and execute Basics functions	2
2	Design Dataset and implement One Dimensions Exploratory Data Graphs	2
3	Design Dataset and implement Two Dimensions Exploratory Data Graphs	2
4	Design Graphics Devices: The Process of Making a Plot	2
5	Design Multiple Open Graphics Devices	2
6	Design code for Simple Base Graphics, Histogram, Boxplot, Scatterplot	2
7	Design code for Base Plot with Regression Line	2
8	Write code using RColorBrewer Package,	2
9	Write code Using the RColorBrewer palettes	2
10	Write code Using smoothScatter() function	2
11	Project	10

Reference Book / Hand Books/ Lab Manual

- 1.Exploratory Data Analysis with R By Roger D. Peng
 - 2.Hands-On Exploratory Data Analysis with Python by Suresh Kumar Mukhiya, Usman Ahmed
- Publisher(s): Packt Publishing ISBN: 9781789537253

Syllabus Semester-IV

Course code: MDS41MEP606 Course name: Practical Based on Data Centre Technologies Course category: Elective
Credits: 1 Teaching scheme: L-0 P-2 Evaluation scheme: CA-30, ESE-20
Pre-requisites: Basic knowledge of Computer Networks and Data Networking.
Course Objectives: A data center is a facility that centralizes an organization's IT operations and equipment for the purposes of storing, processing and disseminating data and applications.
Course Outcomes: At the end of the course, the students will be able to -
CO1: Understanding Data Center concepts
CO2: Learn Configuration steps
CO3: Learn Data Center specific Networking
CO4: Check Availability of Hardware's
CO5: Deploy Developed Data center

List of Practical:

Sr. No.	Title of the Experiment	Practical hours
1	Study Data Center Technology trends	2
2	Study Data center network reference architecture	2
3	Configure Product overview, Product description of the IBM b-type m-series	2
4	Configure Ethernet/IP Routers, Product description of the IBM b-type r-series	2
5	Model Configuration ISO OSI communications model Layer 1- Physical Layer,	2
6	Model Configuration ISO OSI communications model Layer 2 - Data Link Layer	2
7	Configure Design for redundancy,	2
8	Configure Device-level redundancy	2
9	Study Quality of Service architecture,	2
10	Study Quality of Service components	2
11	Project	10

Reference Book / Hand Books/ Lab Manual

1. IBM b-type Data Center Networking: Design and Best Practices Introduction By Jon Tate Norman Bogard, Michal Holenia, Sebastian Oglaza, Steven Tong
2. Cisco Data Center Fundamentals By Somit Maloo, Iskren Nikolov Published Sep 28, 2022 by Cisco Press. Part of the Networking Technology series.

Syllabus Semester-IV

Course code: MDS41MEP607	Course name: Practical Based on Social Media Analytics	
Course category: Elective		
Credits: 1	Teaching scheme: L-0 P-2	Evaluation scheme: CA–30, ESE–20
Pre-requisites: Basic knowledge of Data Analytics.		
Course Objectives: Understand the role of social media data and analytics in helping organizations achieve their goals and understand their publics.		
Course Outcomes: At the end of the course, the students will be able to -		
CO1: To know the Users on Social Media		
CO2: Understand the Social Media Networks		
CO3: Social Media Configurations		
CO4: Content on Social Media		
CO5: Apply Social media tools on Processing Large Datasets		

List of Practical:

Sr. No.	Title of the Experiment	Practical hours
1	Data Collection and API Access Set up API access (e.g., Twitter API) for social media data. Retrieve tweets or posts related to a specific hashtag or topic.	2
2	Data Cleaning and Preprocessing Clean and preprocess the collected social media data. Handle missing values, remove duplicates, and format text data.	2
3	Sentiment Analysis Perform sentiment analysis on social media text. Classify tweets or posts as positive, negative, or neutral.	2
4	Trend Analysis Identify trending topics or hashtags. Analyze the frequency and changes in usage over time.	2
5	Network Analysis Build a network graph based on social media connections (e.g., retweets, mentions). Analyze network properties and identify influential users.	2
6	Hashtag Analysis Analyze the usage and popularity of specific hashtags. Visualize the relationships between hashtags.	2
7	User Engagement Metrics Calculate engagement metrics (likes, shares, comments) for social media posts. Identify highly engaging content.	2
8	Geographic Analysis Extract location information from social media data. Visualize the geographic distribution of posts.	2
9	Social Media Analytics Dashboard Create a simple dashboard using tools like Tableau or Power BI. Display key metrics and visualizations for social media data.	2
10	Predictive Analytics Build a predictive model to forecast social media engagement. Use historical data to predict future trends.	2

11	Project	10
----	---------	----

Reference Book / Hand Books/ Lab Manual

1.Social Media Data Miningand Analytics by Gabor Szabo,Gungor PolatkanOscar Boykin,Antonios ChalkiopoulosPublished byJohn Wiley & Sons, IncISBN: 978-1-118-82485-6

2.Social Media Analytics By Matthew Ganis Avinash Kohirkar IBM Press
Pearson plc ISBN-13: 978-0-13-389256-7

Syllabus Semester-IV

Course code: MDS41MEP608 Course name: Practical Based on Deep Learning Course category: Elective
Credits: 1 Teaching scheme: L-0 P-2 Evaluation scheme: CA–30, ESE–20
Pre-requisites: Basic knowledge of Machine Learning.
Course Objectives: Deep learning techniques have been used successfully for variety of applications.
Course Outcomes: At the end of the course, the students will be able to -
CO1: Understand the basic knowledge about Machine Learning
CO2: Foundation of Neural Network and Deep Learning
CO3: Fundamentals of Deep Networks
CO4: Can Building Deep Network
CO5: Develop Tuning Deep Network

List of Practical:

Sr. No.	Title of the Experiment	Practical hours
1	Image Classification: Use deep learning models like Convolutional Neural Networks (CNNs) to classify images into different categories. This is widely used in areas such as medical imaging, object recognition, and autonomous vehicles.	2
2	Natural Language Processing (NLP): Employ deep learning models like Recurrent Neural Networks (RNNs) or Transformer models for tasks like sentiment analysis, language translation, and chatbot development.	2
3	Speech Recognition: Develop deep learning models for accurately transcribing spoken words into text. This is applied in voice assistants, transcription services, and other voice-controlled applications.	2
4	Recommendation Systems: Utilize deep learning to create personalized recommendation systems, improving user experience on platforms like e-commerce websites, streaming services, or social media.	2
5	Generative Models: Explore generative models like Generative Adversarial Networks (GANs) for tasks such as image generation, style transfer, and even creating synthetic data for training other models.	2
6	Healthcare Diagnostics: Apply deep learning in medical image analysis for tasks like tumor detection, organ segmentation, and disease classification, aiding in early diagnosis and treatment.	2
7	Autonomous Vehicles: Develop deep learning models for object detection, scene understanding, and decision-making to enable self-driving cars and autonomous vehicles.	2
8	Fraud Detection: Implement deep learning algorithms to detect fraudulent activities in financial transactions, identifying patterns that may indicate potential fraud.	2
9	Human Pose Estimation: Use deep learning models to estimate and track the poses of human bodies in images or videos. This is applied in areas like fitness tracking, gesture recognition, and animation.	2
10	Industrial Automation: Apply deep learning in manufacturing for quality control, predictive maintenance, and process optimization, enhancing efficiency and reducing errors in production.	2

11	Project	10
----	---------	----

Reference Book / Hand Books/ Lab Manual

1. Deep Learning Ian Goodfellow, Yoshua Bengio, Aaron Courville MIT Press

2. Deep Learning: A Practitioner's Approach Josh Patterson, Adam Gibson O'Reilly Media, Inc