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.

विद्यापीठ गीत

अत्त दिप भव भव प्रदिप भव,

स्वरूप रूप भव हो

ज्ञान सब्ब विज्ञान सब्ब भव ,

सब्ब दिप भव हो

अत्ताहि अत्त नो नाथो ,

अत्ताहि अत्त नो गति

अत्त मार्गपर अप्रमादसे है तुझे चलना

सब्ब का कल्याण हो ,

वो कार्यकुशल करना

सब्ब का उत्तम मंगल , पथप्रदर्शक हो

अत्त दिप भव भव प्रदिप भव ,

स्वरूप रूप भव हो

ज्ञान सब्ब विज्ञान सब्ब भव ,

सब्ब दिप भव हो

बुद्धमं शरनं गच्छामि :

धम्मं शरनं गच्छामि :

संघं शरनं गच्छामि :

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 Sambhajinagar.

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.

Name of Faculty: Basic and Applied Science

Name of the College: Dr. G. Y. Pathrikar College of Computer Science and Information Technology, MGM University Aurangabad

Name of the Programme: M.Sc. Data Science

Programme Type (UG/PG):PG

Duration: Two Years

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**
MM	MDS41MML501	Introduction to Data Science	Lecture	3	3		60	40	100		16	40
MM	MDS41MML502	Data Structure and Algorithm	Lecture	3	3		60	40	100		16	40
MM	MDS41MML503	Python Programming	Lecture	3	3		60	40	100		16	40
RM	MDS41RML501	Research Methodology	Lecture	4	4		60	40	100		16	40
ME		Elective from Basket-1	Lecture	3	3		60	40	100		16	40
MM	MDS41MMP501	Practical Based on Introduction to Data Science	Practical	1	-	2	30	20	50		8	20
MM	MDS41MMP502	Practical Based on Data Structure and Algorithm	Practical	1	-	2	30	20	50		8	20
MM	MDS41MMP503	Practical Based on Python Programming	Practical	1	-	2	30	20	50		8	20
ME		Practical based on Elective from Basket-1	Practical	1	-	2	30	20	50		8	20
Total				**20**	**16**	**8**	**420**	**280**	**100**			

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.

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**
MM	MDS41MML504	Advances in Database Management System	Lecture	3	3		60	40	100		16	40
MM	MDS41MML505	Data Mining and Visualization	Lecture	3	3		60	40	100		16	40
MM	MDS41MML506	Machine Learning	Lecture	3	3		60	40	100		16	40
ME		Elective from Basket-2	Lecture	3	3		60	40	100		16	40
MM	MDS41MMP504	Practical Based on Advances in Database Management System	Practical	1	-	2	30	20	50		8	20
MM	MDS41MMP505	Practical Based on Data Mining and Visualization	Practical	1	-	2	30	20	50		8	20
MM	MDS41MMP506	Practical Based on Machine Learning	Practical	1	-	2	30	20	50		8	20
ME		Practical based on Elective from Basket-2	Practical	1	-	2	30	20	50		8	20
OJT	MDS41JTJ501	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-I

Course code: MDS41MML501 **Course name:** Introduction to Data Science
Course category: Major Mandatory **Credits:** 3
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
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
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	**Introduction to Python:** 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.	10
2	**Conditional and Looping Construct** 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. **Functions:** Built-In Function, invoking built in functions Module(Importing entire module or selected objects using from statement) Functions from math, random, time & date module. **Composition User Define Function:** Defining , invoking functions, passing parameters (default parameter values, keyword arguments) Scope of variables, void functions and functions returning values	10
3	**String :** 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. **Errors and Exceptions**: Syntax Errors, Exceptions, Handling Exceptions, Raising Exceptions, User-defined Exceptions, Defining Clean-up Actions(try - finally), Predefined Clean-up Actions	10

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
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
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
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
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.

Semester-I

Course code: MDS41RMP501	**Course name:** Practical Based on Research Methodology
Course category: Major Mandatory	**Credits:** 1
Course Objectives: To understand the state-of-the-art in research methodology and 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	

Contents –

Unit	Content	Teaching hours
1	Identifying and defining a research problem and formulating the objective of your case study	1
2	Literature Study : Study of previous work done	1
3	Development of framework for Case Study	1
4	Preparation of Schedule	1
5	Collection of Data : Construction of Questionnaire/ Conduction of interview	1
6	Analysis of Data	1
7	Outcome of the study	1
8	Preparation of Report	1
9	Identifying and defining a research problem and formulating the objective of your experimental research	1
10	Literature Review of the selected research topic	1

Text Books: C. R. Kothari Research Methodology Methods and Techniques New Age International Publishers
Reference Books: P. Sam Daniel, Aroma G. Sam Research Methodology Gyan Publishing House
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
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:** 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.	10

5	**Convex functions:** Basic properties and examples. Operations that preserve convexity. The conjusgate 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
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
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
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
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
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
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
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
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, RoolBack, 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
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:** 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.	10
5	**Introduction to Data Visualization:**	5

	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.	

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
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:** Intelligent Machine, Machine Learning Problem, Applications, Data Representation, Domain Knowledge for Productive use of Machine Learning, Diversity of Data: Structured and Unstructured, Forms of Learning, Machine Learning and Data Mining.	10
2	**Supervised Learning:** Learning from Observations, Bias and Variance, Computational Learning Theory, Heuristic Search, Cross-Validation, Bootstrapping, Mean Square Error, Mean Absolute Error, Misclassification Error, Confusion Matrix, ROC Curves, Issues in Machine Learning.	10
3	**Learning with SVM and NN:** Learning with Support Vector Machine (SVM): Linear Discriminant Functions for Binary Classification, Perceptron Algorithm, Nonlinear Classifier, Linear Regression and Nonlinear Regression, SVM Techniques. Learning with Neural Network (NN): Neuron Model, Biological Neuron, Artificial Neuron, Network Architecture, Feedforward and Recurrent Network, Perceptron.	10
4	**Data Clustering:** Unsupervised Learning, Clustering, Data Analysis, Clustering Analysis, Data Transformation, Enhancing the Information Content of the Data, Partitional Clustering, Hierarchical Clustering, K-Means Clustering, Fuzzy K-Means Clustering, Expectation Maximization.	10
5	**Business Intelligence and Data Mining:** Basic Analytical Techniques, Data Warehousing, Intelligent Information Retrieval System: Text Retrieval, Image Retrieval, Audio Retrieval, Data Mining Applications, Data Mining Trends, Technologies for Big Data.	5

Text Books: M. Gopal Applied Machine Learning McGraw-Hill Education
Reference Books: Charu C. Aggarwal Machine Learning for Text Springer
Online Resources: 1. NPTEL / SWAYAM lectures.

Elective from Basket-2 Semester-II

Course code: MDS41MEL505 **Course name:** Google Analytics
Course category: Elective **Credits:** 3
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
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, Backpropsagation 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
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
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
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
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
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 drawgraphs.	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
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
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
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
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 Vercellis 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.