Manipal School of Information Sciences (MSIS)

Manipal Academy of Higher Education, Manipal

Master of Engineering - ME (Big Data Analytics)

Course File

Course Name	:	Algorithms and Data Structures for Big Data
Course Code	:	BDA 5101
Academic Year	:	2024 – 2025
Semester	:	I
Name of the Course Coordinator	:	Mr. DEEPAK RAO B
Name of the Program Coordinator	:	Dr. PRATHVIRAJ N

Signature of Program Coordinator with Date	**Signature of Course Coordinator with Date**

Program Education Objectives (PEOs)

The overall objectives of the Learning Outcomes-based Curriculum Framework (LOCF) for **ME (Big Data Analytics)**, program are as follows.

PEO No.	Education Objective
PEO 1	Develop in depth understanding of the key technologies in data engineering, data science and business analytics.
PEO 2	Practice problem analysis and decision-making using machine learning techniques.
PEO 3	Gain practical, hands-on experience with statistics, programming languages and big data tools through coursework and applied research experiences.

Program Outcomes (POs)

By the end of the postgraduate program in Big Data Analytics, graduates will be able to:

PO1	Independently carry out research /investigation and development work to solve practical problems.
PO2	Write and present a substantial technical report/document.
PO3	Demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program.
PO4	Develop and implement big data analysis strategies based on theoretical principles, ethical considerations, and detailed knowledge of the underlying data.

PO5	Demonstrate knowledge of the underlying principles and evaluation methods for analyzing data for decision-making.

1. Course Plan

1.1 Primary Information

Course Name	**:**	Algorithms and Data Structures for Big Data
L-T-P-C	**:**	3-0-0-3
Contact Hours	**:**	36 Hours
Pre-requisite	**:**	Programming with Python or C
Core/ PE/OE	**:**	Core

1.2 Course Outcomes (COs), Program outcomes (POs) and Bloom's Taxonomy Mapping

CO	At the end of this course, the student should be able to:	No. of Contact Hours	Marks	Program Outcomes (PO's)	BL
CO1	**Analyze** recursive programs, solve a general class of recurrence relations	3	10	PO3	3
CO2	**Design** programs for implementation of linked lists, stack, queues and binary search tree	10	40	PO4	4

CO3	**Design** programs for dictionary, hash tables, graphs, shortest path techniques, sorting and searching.	14	40	PO4	4
CO4	**Design** string and text processing programs.	9	10	PO4	4
	Total		**100**		

1.3 Assessment Plan

Components	**Mid Semester**	**Flexible Assessments (2 – 3 in number)**	**End semester/ Makeup examination**
Duration	90 minutes	To be decided by the faculty.	180 minutes
Weightage	30%	20%	50%
Typology of questions	Applying; Analyzing and Evaluating	Applying; Analyzing. Evaluating.	Applying; Analyzing; Evaluating.

Pattern	Answer all 5 questions of 10 marks each. Each question may have 2 to 3 parts of 3/4/5/6/7 marks.	**Quiz / Test:** 10% **Assignment 1:** (Text processing algorithms, 5% weightage) **Assignment 2:** (Data Streaming algorithms, 5% weightage)	Answer all 10 full questions of 10 marks each. Each question may have 2 to 3 parts of 3/4/5/6/7 marks.
Schedule	As per academic calendar.	**Quiz / Test:** September 2024 **Assignment 1:** October 2024 **Assignment 2:** November 2024	As per academic calendar.
Topics covered	Algorithm Analysis Techniques Elementary Data structures Trees Sorting, Searching Dictionary and Hashing		Comprehensive examination covering the full syllabus. Students are expected to answer all questions.

1.4 Lesson Plan

L. No.	TOPICS	Course Outcome Addressed
L0	Course delivery plan, Course assessment plan, Course outcomes, Program outcomes, CO-PO mapping, reference books	---
L1	Analysis of recursive programs.	CO1
L2	Solving recurrence equations.	CO1
L3	General solution for a large class of recurrences.	CO1
L4	Implementation of lists.	CO2
L5	Implementation of lists.	CO2
L6	Implementation of lists.	CO2
L7	Implementation of stacks.	CO2
L8	Implementation of stacks.	CO2
L9	Implementation of queues.	CO2
L10	Implementation of queues.	CO2
L11	Implementation of Trees.	CO2
L12	Implementation of Trees.	CO2
L13	Implementation of Trees.	CO2
L14	Sorting – bubble, selection and insertion	CO3
L15	Sorting – Quick sort	CO3
L16	Sorting – Merge sort	CO3
L17	Sorting – Heap Sort	CO3

L18	Searching – Linear and binary	CO3
L19	Dictionary and Hash Tables	CO3
L20	Dictionary and Hash Tables	CO3
L21	Dictionary and Hash Tables	CO3
	Mid Semester Evaluation	CO1, CO2, CO3
L22	Graph Terminology	CO3
L23	Representation of graphs	CO3
L24	Traversing Graphs	CO3
L25	Shortest Path algorithm	CO3
L26	Shortest Path algorithm	CO3
L27	String and Text processing	CO4
L28	String and Text processing	CO4
L29	String and Text processing	CO4
L30	String and Text processing	CO4
L31	Data streaming algorithms	CO4
L32	Data streaming algorithms	CO4
L33	Data streaming algorithms	CO4
L34	Data streaming algorithms	CO4
L35	Data streaming algorithms	CO4
L36	Data streaming algorithms	CO4

1.5 References

1. Introduction to Algorithms - Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest. MIT Press.
2. Data Structures and Algorithms - Aho, Hopcroft and Ulmann. Pearson Publishers.
3. Data Structures and Algorithms in Python - Michael T. Goodrich, Roberto Tamassia, and Michael H. Goldwasser. John Wiley & Sons.
4. Data Streams: Algorithms and Applications - S. Muthukrishnan. Foundations and Trends in Theoretical Computer Science archive, Volume 1 Issue 2, August 2005, Pages 117 – 236
5. https://in.coursera.org/specializations/boulder-data-structures-algorithms

1.6 Other Resources (Online, Text, Multimedia, etc.)

6. Web Resources: Blog, Online tools and cloud resources.
7. Journal Articles.

1.7 Course Timetable

1st Semester Big Data Analytics				Room: LG1 LH 8		Lab: Data Science Lab		
	9-10	10-11	11-12	12-1	1-2	2-3	3-4	4-5
MON		ADS						
TUE								
WED		ADS						
THU								
FRI		ADS						
SAT								

1.8 Assessment Plan

Cos		**Marks & Weightage**			
CO No.	CO Name	IT-1 (Max. 50)	Assignment (Max. 10)	End Semester (Max. 100)	CO wise Weightage
CO1	Analyze recursive programs, solve a general class of recurrence relations	5	-	10	**0.09**
CO2	Design programs for implementation of linked lists, stack, queues, binary search tree	30	10	50	**0.53**

CO3	Design programs for sorting and searching, dictionary, hash tables, graphs and shortest path techniques.	15		40	**0.32**
CO4	Design string and text processing programs.	-	10		**0.06**
	Marks (weightage)	**0.3**	**0.2**	**0.5**	**1.0**

Note:

- In-semester Assessment is considered as the Internal Assessment (IA) in this course for 50 marks, which includes the performances in class participation, assignment work, class tests, mid-term tests, quizzes etc.
- End-semester examination (ESE) for this course is conducted for a maximum of 100 and the same will be scaled down to 50.

End-semester marks for a maximum of 50 and IA marks for a maximum of 50 are added for a maximum of 100 marks to decide upon the grade in this course.

1.9 Assessment Details

The assessment tools to be used for the Current Academic Year (CAY) are as follows:

SI. No.	Tools (TLP)	Weightage	Frequency	Details of Measurement (Weightage/Rubrics/Duration, etc.)
1	Mid Semester	0.3	1	• Performance is measured using sessional attainment level. • Reference: question paper and answer scheme.

				• Mid semester exam is assessed for a maximum of 50 marks and scaled down to 30 marks
2	Assignments	0.2	1	• Performance is measured using assignments/quiz attainment level. • Assignments/quiz are evaluated for a maximum of 20 marks.
3	ESE	0.5	1	• Performance is measured using ESE attainment level. • Reference: question paper and answer scheme. • ESE is assessed for a maximum of 100 marks and scaled down to 50 marks.

1.10 Course Articulation Matrix

CO	PO1	PO2	PO3	PO4	PO5
CO1			**Y**		
CO2				**Y**	
CO3				**Y**	
CO4				**Y**	
Average Articulation Level			*	*	