RCOEM

Shri Ramdeobaba College of Engineering and Management, Nagpur

SHRI RAMDEOBABA COLLEGE OF ENGINEERING AND MANAGEMENT, NAGPUR – 440013

PROGRAMME SCHEME & SYLLABI

2023-24

B.Tech in Computer Science and Engineering (Data Science)

Published By

Dr. R. S. Pande

Principal
Shri Ramdeobaba College of Engineering & Management
Ph. : 0712-2580011 Fax : 0712 - 2583237
ISO 9001 : 2015 CERTIFIED ORGANISATION
2015 CERTIFIED ORGANISATION

ABOUT THE DEPARTMENT:

The Department of Computer Science & Engineering (Data Science) was established with the primary objective of imparting knowledge in the field of computer science and engineering with special emphasis on data science. The department is well-equipped with state-of-the-art infrastructure which includes latest configuration desktops organized in three different laboratories. Latest software, wireless access point, LCD projectors and separate routers are used in the laboratories for teaching purpose. The department has well qualified and highly experienced faculty with expertise in diversified areas of Data Science and AI ML with industry experience. Besides excellence in teaching, there is a committed pursuance of high-quality research and innovation. Faculty members of the department have been involved in research and patent filling. The curriculum is designed to impart students with in-depth knowledge and understanding of the key technologies in data science. Open Electives and MOOCs provide an opportunity for interdisciplinary learning. Option of full six months internship is provided in reputed industries and organizations for VIII semester students. Department takes pride in excellent internships and placements of the final year students and has the distinction of consistently getting good results in all semesters. The student club of the department, DASCA, organizes various technical and co-curricular activities which ensures all round development of the students.

DEPARTMENT VISION:

To achieve excellent standards of quality-education in computer science and engineering with a specialized focus on data science, foster innovation, nurture collaborative culture and contribute in the field of data science to serve the industry and society.

DEPARTMENT MISSION:

To empower the next generation of computer scientists and engineers with the knowledge, skills, and tools essential to excel in the rapidly evolving field of data science.

To impart the ability to devise globally recognized solutions for the problems of society and industry,

To provide an environment for interdisciplinary collaboration, research, and life long learning.

PROGRAM EDUCATION OBJECTIVES (PEOs):

1. Graduates will have strong foundation of knowledge and skills in mathematics, statistics, programming and computer science to solve problems in data science.
2. Graduates will have the ability and attitude to adapt to emerging technological changes with lifelong learning skills.
3. Graduates will demonstrate collaborative learning and spirit of team work through multidisciplinary Data Science projects ensuring ethical and moral values.
4. Graduates will demonstrate professionalism, ethical attitude and leadership skills with lifelong learning in their career.

PROGRAMME OUTCOMES (POs):

1. Engineering knowledge: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
2. Problem analysis: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.
3. Design/development of solutions: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
4. Conduct investigations of complex problems: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
5. Modern tool usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.
6. The engineer and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.
7. Environment and sustainability: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
8. Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
9. Individual and team work: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

10. Communication: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

11. Project management and finance: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

12. Life-long learning: Recognize the need for, and have the preparation and ability to engage in independent and lifelong learning in the broadest context of technological change.

PROGRAMME SPECIFIC OUTCOMES (PSOs):

1. Apply the concepts and practical knowledge of data science in analysis, design and development of computing systems and applications to multi-disciplinary problems.

2. Acquaint with the contemporary trends in industrial, research settings and thereby innovate novel solutions to existing problems.

B. Tech. Computer Science and Engineering (Data Science) [2023-24 NEP]
Teaching & Evaluation Scheme

Semester–I

Sr. No.	Category	Course Code	Course Name	Hours/week			Credits	Maximum marks			ESE Duration (Hrs)
				L	T	P		Continuous Evaluation	End Sem Exam	Total	
1.	BSC	CHT1001	Chemistry of Smart Materials	2	0	0	2	50	50	100	2
2.	BSC	CHP1001	Chemistry of Smart Materials Lab	0	0	2	1	50	-	50	-
3.	BSC	MAT1002	Calculus	3	0	0	3	50	50	100	3
4.	ESC	CDT1001	Digital Electronics	3	0	0	3	50	50	100	3
5.	ESC	CDP1001	Digital Electronics Lab	0	0	2	1	50	-	50	-
6.	ESC	CDT1002	Programming for problem solving	3	0	0	3	50	50	100	3
7.	ESC	CDP1002	Programming for problem solving Lab	0	0	2	1	50	-	50	-
8.	VSEC	CDT1003	Computer Workshop – I	1	0	0	1	50	-	50	-
9	VSEC	CDP1003	Computer Workshop – I Lab	0	0	2	1	50	-	50	-
10	HSSM-IKS	HUT1001	Foundational Literature of Indian Civilization	2	0	0	2	50	50	100	2
11	CCA	PET1001	Sports-Yoga-Recreation	1	0	0	1	50	-	50	-
12	CCA	PEP1001	Sports-Yoga-Recreation Lab	0	0	2	1	50	-	50	-
TOTAL				**15**	**0**	**10**	**20**			**850**	**-**

Semester - II

Sr. No.	Category	Course Code	Course Name	Hours/week			Credits	Maximum marks			ESE Duration (Hrs)
				L	T	P		Continuous Evaluation	End Sem Exam	Total	
1.	BSC	PHT2001	Introduction to Quantum Computing	2	1	0	3	50	50	100	3
2.	BSC	PHP2001	Introduction to Quantum Computing Lab	0	0	2	1	50	-	50	-
3.	BSC	MAT2002	Discrete Mathematics	3	0	0	3	50	50	100	3
4.	BSC	MAP2001	Computational Mathematics Lab	0	0	2	1	50	-	50	-
5.	BSC	CHT2007	Bioinformatics	2	0	0	2	50	50	100	2
6.	ESC	CDT2001	Object Oriented Programming	3	0	0	3	50	50	100	3
7.	ESC	CDP2001	Object Oriented Programming Lab	0	0	2	1	50	-	50	-
8.	PCC	CDT2002	Computer Architecture	2	0	0	2	50	50	100	2
9.	VSEC	CDT2003	Computer Workshop – II	1	0	0	1	50	-	50	-
10.	VSEC	CDP2003	Computer Workshop – II Lab	0	0	2	1	50	-	50	-
11.	AEC	HUT2002	English for Professional Communication	2	0	0	2	50	50	100	2
12.	AEC	HUP2002	English for Professional Communication Lab	0	0	2	1	50	-		
13.	CCA	HUP0001	Liberal/Performing Art	0	0	2	1	50	-	50	-
14.	VEC	HUT2004	Foundational Course in Universal Human Values	1	0	0	1	50	-	50	-
TOTAL				**16**	**1**	**12**	**23**			**1000**	**-**

Exit option : Award of UG Certificate in Major with 43 credits and an additional 8 credits.			
Exit Courses			
1	**Web Designer**	**Online/offline certification Course**	**8**
2	**IT Support Engineer**		**8**
3	**Certified Programmer (language learned in Sem-1 and/or Sem-2 [C,C++,Java, Python])**		**8**

Semester - III

Sr. No.	Category	Course code	Course Name	Hours/ week			Credits	Maximum marks			ESE Duration (Hrs)
				L	T	P		Continuous Evaluation	End Sem Exam	Total	
1	PCC	CDT3001	Data Structures and Algorithms	3	1	0	4	50	50	100	2
2	PCC	CDP3001	Data Structures and Algorithms Lab	0	0	2	1	50	-	50	-
3	PCC	CDT3002	Theory of Computations	3	0	0	3	50	50	100	2
5	PCC	CDP3003	Programming Language Lab	0	0	4	2	50	-	50	-
6	MDM	MAT3002	Probability and Statistics	3	0	0	3	50	50	100	2
8	VSEC	CDP3004	Software Laboratory - I	0	0	2	1	50	-	50	-
7	OE	CDT2980-1	Open Elective - I	4	0	0	4	50	50	100	2
8	AEC	HUT3001	Business Communication	2	0	0	2	50	50	100	2
9	HSSM	CDP3006	Idea Lab	0	0	4	2	50	-	50	-
10	VEC	CDT3007	Cyber Laws and Ethics in IT	2	0	0	2	50	50	100	2
TOTAL				**17**	**1**	**12**	**24**	**500**	**300**	**800**	

Semester - IV

Sr. No.	Category	Course code	Course Name	Hours/ week			Credits	Maximum marks			ESE Duration (Hrs)
				L	T	P		Continuous Evaluation	End Sem Exam	Total	
1	PCC	CDT4001	Operating Systems	3	0	0	3	50	50	100	3
2	PCC	CDP4001	Operating Systems Lab	0	0	2	1	50	-	50	-
3	PCC	CDT4002	Artificial Intelligence	3	0	0	3	50	50	100	3
4	PCC	CDP4002	Artificial Intelligence Lab	0	0	2	1	50	-	50	-
5	PCC	CDT4003	Design and Analysis of Algorithms	3	0	0	3	50	50	100	3
6	MDM	MAT4001	Linear Algebra	3	0	0	3	50	50	100	3
7	OE	CDT2990-1	Open Elective - II	2	0	0	2	50	50	100	2
8	VSEC	CDP4005	Software Laboratory - II	0	0	2	1	50	-	50	-
9	HSSM	HUT4003	Managerial Economics	2	0	0	2	50	50	100	2
10	VEC	HUT4002	Environmental Education	2	0	0	2	50	50	100	2
11	CEP	CDP4006	Community Engagement Project	0	0	4	2	50	-	50	-
TOTAL				**18**	**0**	**10**	**23**	**550**	**350**	**900**	

Exit option : Award of UG Diploma in Major with 90 credits and an additional 8 credits			
Exit Courses			
1	Application Development (Android, Advanced Java)	Online/offline certification Course	8
2	Certified Software Engineer (DevOps)		8

Semester- V

Sr. No.	Category	Course Code	Course Name	Hours/ week			Credits	Maximum marks			ESE Duration (Hrs)
				L	T	P		Continuous Evaluation	End Sem Exam	Total	
1	PCC	CDT5001	Database Management Systems	3	0	0	3	50	50	100	3
2	PCC	CDP5001	Database Management Systems Lab	0	0	2	1	50	-	50	-
3	PCC	CDT5002	Machine Learning	3	0	0	3	50	50	100	3
4	PCC	CDP5002	Machine Learning Lab	0	0	2	1	50	-	50	-
5	PCC	CDT5003	Computer Networks	3	0	0	3	50	50	100	3
6	PEC	CDT5004	Program Elective - I	3	0	0	3	50	50	100	3
7	MDM	CDT5005	Financial Analytics	3	0	0	3	50	50	100	3
8	MDM	CDP5005	Financial Analytics Lab	0	0	2	1	50	-	50	-
9	OE	CDT3980-1	Open Elective - III	2	0	0	2	50	50	100	2
TOTAL				**17**	**0**	**6**	**20**	**450**	**300**	**750**	

Semester- VI

Sr. No.	Category	Course code	Course Name	Hours/ week			Credits	Maximum marks			ESE Duration (Hrs)
				L	T	P		Continuous Evaluation	End Sem Exam	Total	
1	PCC	CDT6001	Data Warehousing and Business Intelligence	3	0	0	3	50	50	100	3
2	PCC	CDP6001	Data Warehousing and Business Intelligence Lab	0	0	2	1	50	-	50	-
3	PCC	CDT6002	Deep Learning - I	3	0	0	3	50	50	100	3
4	PCC	CDP6002	Deep Learning - I Lab	0	0	2	1	50	-	50	-
5	PCC	CDT6003	Data Analysis and Visualization	1	0	0	1	50	-	50	-
6	PCC	CDP6003	Data Analysis and Visualization Lab	0	0	2	1	50	-	50	-
7	PEC	CDT6004	Program Elective - II	3	0	0	3	50	50	100	3

8	PEC	CDP6004	Program Elective - II Lab	0	0	2	1	50	-	50	-
9	PEC	CDT6005	Program Elective - III	3	0	0	3	50	50	100	3
10	MDM	CDT6006	Applied Econometrics	2	0	0	2	50	50	100	2
11	VSEC	CDP6007	Mini Project	0	0	4	2	25	25	50	-
TOTAL				**15**	**0**	**12**	**22**	**525**	**275**	**800**	

Exit option : Award of UG Degree in Major with 131 credits and an additional 8 credits			
Exit Courses			
1	**Certified Database Engineer (Oracle, DB2)**	**Online/offline certification Course**	**8**
2	**Certified Cloud Engineer (AWS, AZURE)**		**8**
3	**Certified Data Science Engineer**		**8**

Semester-VII

Sr. No.	**Category**	**Course code**	**Course Name**	**Hours/ week**			**Credits**	**Maximum marks**			**ESE Duration (Hrs)**
				L	**T**	**P**		**Continuous Evaluation**	**End Sem Exam**	**Total**	
1	PCC	CDT7001	Deep Learning - II	3	0	0	3	50	50	100	3
2	PCC	CDP7001	Deep Learning - II Lab	0	0	2	1	50	-	50	-
3	PCC	CDT7002	Cloud Computing	3	0	0	3	50	50	100	3
4	PCC	CDP7002	Cloud Computing Lab	0	0	2	1	50	-	50	-
5	PEC	CDT7003	Program Elective - IV	3	0	0	3	50	50	100	3
6	PEC	CDP7003	Program Elective - IV Lab	0	0	2	1	50	-	50	-
7	MDM	CDT7004	Customer Management	2	0	0	2	50	50	100	2
8	PRJ	CDP7005	Major Project - I	0	0	8	4	50	50	100	-
			TOTAL	**10**	**0**	**16**	**18**	**400**	**250**	**650**	

Semester- VIII

Sr. No.	Category	Course code	Course Name	Hours/ week			Credits	Maximum marks			ESE Duration (Hrs)
				L	T	P		Continuous Evaluation	End Sem Exam	Total	
1.	PEC	CDT8001	Program Elective-V	3	0	0	3	50	50	100	3
2.	PEC	CDT8002	Program Elective-VI	3	0	0	3	50	50	100	3
3.	Project	CDP8003	Major Project- II	0	0	12	6	50	50	100	-
			TOTAL	**6**	**0**	**12**	**12**	**150**	**150**	**300**	
OR											
1	PEC	CDT8001	Program Elective-V	3	0	0	3	50	50	100	3
2	PEC	CDT8004	Research Methodology	3	0	0	3	50	50	100	3
3	Project	CDP8005	Research Project	0	0	12	6	50	50	100	-
			TOTAL	**6**	**0**	**12**	**12**	**150**	**150**	**300**	
OR											
								Continuous Evaluation	**Industry Evaluation**	**Total**	
1.	INTR-801	CDP8006	Industry / TBI Internship	0	0	24	12	100	100	200	-
			TOTAL	**0**	**0**	**24**	**12**	100	100	**200**	

List of Program Electives

Elective-I	Elective-II	Elective-III	Elective-IV	Elective-V	Elective-VI
CDT5004-1 Compiler Design	CDT6004-1 Natural Language Processing	CDT6005-1 Data Science for Health Care	CDT7003-1 Image Analysis and Computer Vision	CDT8001-3 Mining Massive Datasets	CDT8002-1 Human Computer Interaction
CDT5004-2 Design Patterns	CDT6004-2 Blockchain Technology	CDT6005-2 Data Science for Genomics	CDT7003-2 Spatial Data Management	CDT8001-2 Generative AI	CDT8002-2 Optimization for Data Science
CDT5004-3 Software Engineering	CDT6004-3 Distributed and Parallel Computing	CDT6005-3 Data Science for Marketing	CDT7003-3 Information Retrieval	CDT8001-3Information Security and Data Privacy	CDT8002-3 Reinforcement Learning

B. Tech. Computer Science and Engineering (Data Science) [2023-24 NEP]
Honors and Minors Scheme

Following courses are proposed under Honors in CSE(Data Science)

Sr. No.	Category	Course code	Course Name	Hours/ week			Credits	Maximum marks			ESE Duration (Hrs)
				L	T	P		Continuous Evaluation	End Sem Exam	Total	
1	Honor	CDTH	Data Science Essentials	3	0	0	3	50	50	100	-
2	Honor	CDTH	System Architecture Analysis	3	0	0	3	50	50	100	-
3	Honor	CDTH	Data Engineering	4	0	0	4	50	50	100	-
4	Honor	CDTH	Business and Web Analytics	4	0	0	4	50	50	100	-
5	Honor	CDTH	Project	4	0	0	4	50	50	100	-
			TOTAL	**18**	**0**	**0**	**18**	**250**	**250**	**500**	

Note : Marks division for all the courses mentioned above (CSE(DS) Honors) will be of 50 marks for Continuous Evaluation and 50 marks for End Sem Exam. Evaluation pattern for Continuous Evaluation and ESE can be lab conduction, viva voce, execution based test and MCQ.

Scheme for Minors

Following courses are proposed under Minors in CSE(Data Science)

Sr. No.	Category	Course code	Course Name	Hours/ week			Credits	Maximum marks			ESE Duration (Hrs)
				L	T	P		Continuous Evaluation	End Sem Exam	Total	
1	Minor	CDTM	Foundation of Data Science	3	0	0	3	50	50	100	3
2	Minor	CDTM	Statistics for Data Science	3	0	0	3	50	50	100	3
3	Minor	CDTM	Exploratory Data Analysis	4	0	0	4	50	50	100	3
4	Minor	CDTM	AI tools for Data Science	4	0	0	4	50	50	100	3
8	Minor	CDTM	Project	4	0	0	4	50	50	100	-
			TOTAL	**18**	**0**	**0**	**18**	**250**	**250**	**500**	

B. Tech. Computer Science and Engineering (Data Science) [2023-24 NEP]
Teaching & Evaluation Scheme

OPEN ELECTIVE SCHEME

Sr. No.	Category	Course code	Course Name	Hours/ week			Credits	Maximum marks			ESE Duration (Hrs)
				L	T	P		Continuous Evaluation	End Sem Exam	Total	
1	OE	CDT3005	Introduction to Data Science	4	0	0	4	50	50	100	2
2	OE	CDT4004	Tools for Data Science	2	0	0	2	50	50	100	2
3	OE	CDT5006	Data Analytics Visualization	2	0	0	2	50	50	100	2
			TOTAL	**08**	**0**	**0**	**08**	**150**	**150**	**300**	

Syllabus for Semester III, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDT3001 **Category:** Program Core Course (PCC)

Course: Data Structures and Algorithms **L:** 3Hrs**, T:** 1Hr**, P:** 0Hr**, Per Week, Credits:** 4

Course Objectives

The objective of this course is to familiarize the prospective engineers with:

1. To impart to students the basic concepts of data structures and algorithms.
2. To familiarize students on different searching and sorting techniques.
3. To prepare students to use linear (stacks, queues, linked lists) and non-linear (trees, graphs) data structures.
4. To enable students to devise algorithms for solving real-world problems.

SYLLABUS

UNIT I Data Structures and Algorithms Basics

Introduction: basic terminologies, elementary data organizations, data structure operations; abstractdata types (ADT) and their characteristics.Algorithms: definition, characteristics, analysis of an algorithm, asymptotic notations, time and spacetradeoffs.Array ADT: definition, operations and representations – row-major and column- major.

UNIT II Stacks and Queues

Stack ADT: allowable operations, algorithms and their complexity analysis, applications of stacks –expression conversion and evaluation (algorithmic analysis), multiple stacks. Queue ADT: allowable operations, algorithms and their complexity analysis for simple queue andcircular queue, introduction to double-ended queues and priority queues.

UNIT III Linked Lists

Singly Linked Lists: representation in memory, algorithms of several operations: traversing, searching,insertion, deletion, reversal, ordering, etc.Doubly and Circular Linked Lists: operations and algorithmic analysis. Linked representation of stacksand queues, header node linked lists.

UNIT IV Sorting and Searching

Sorting: different approaches to sorting, properties of different sorting algorithms (insertion, Shell,quick, merge, heap, counting), performance analysis and comparison.

Searching: necessity of a robust search mechanism, searching linear lists (linear search, binary search)and complexity analysis of search methods.

UNIT V Trees

Trees: basic tree terminologies, binary tree and operations, binary search tree [BST] and operations with time analysis of algorithms, threaded binary trees. Self-balancing Search Trees: tree rotations, AVL tree and operations, B+-tree: definitions, characteristics, and operations (introductory).

UNIT VI Graphs and Hashing

Graphs: basic terminologies, representation of graphs, traversals (DFS, BFS) with complexity analysis, path finding (Dijkstra's SSSP, Floyd's APSP), and spanning tree (Prim's method) algorithms.

Hashing: hash functions and hash tables, closed and open hashing, randomization methods (division method, mid-square method, folding), collision resolution techniques.

Course Outcomes

After successful completion of this course, the student will be able to,

1. Recognize different ADTs and their operations and specify their complexities.
2. Design and realize linear data structures (stacks, queues, linked lists) and analyze their computation complexity.
3. Devise different sorting (comparison based, divide-and-conquer, distributive, and tree-based) and searching (linear, binary) methods and analyze their time and space requirements.
4. Design traversal and path finding algorithms for Trees and Graphs.

Text Books

1. Ellis Horowitz, Sartaj Sahni& Susan Anderson-Freed, Fundamentals of Data Structures in C,Second Edition, Universities Press, 2008.
2. **Robert L. Kruse and Bruce P. Leung, Data Structures and Program Design in C, Pearson Education;**
3. Mark Allen Weiss; Data Structures and Algorithm Analysis in C; Second Edition; Pearson Education; 2002.
4. G.A.V. Pai; Data Structures and Algorithms: Concepts, Techniques and Application; FirstEdition; McGraw Hill; 2008.

Reference Books

1. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein; Introduction to Algorithms; Third Edition; PHI Learning; 2009.

2. Ellis Horowitz, Sartaj Sahni and Sanguthevar Rajasekaran; Fundamentals of Computer Algorithms; Second Edition; Universities Press; 2008.
3. A. K. Sharma; Data Structures using C, Second Edition, Pearson Education, 2013.

Syllabus for Semester III, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDP3001 **Category:** Program Core Course (PCC)

Course: Data Structures and Algorithms Lab **L:** 0Hr**, T:** 0Hr**, P:** 2Hrs**, Per Week, Credits:** 1

Course Outcome

On Successful completion of course, students will be able to:

1. To enable students to employ different searching and sorting methods.
2. To prepare students to identify and apply linear (stacks, queues, linked lists) and non-Linear(trees, graphs) data structures in solving problems.
3. To encourage students to design and execute tree-based algorithms for solving real-world problems.

SYLLABUS

Experiments based on CDT3001 Syllabus in C / C++ / Java / Python

Course Outcomes

On completion of the course the student will be able to

1. Design and realize different linear data structures.
2. Identify and apply specific methods of searching and sorting to solve a problem.
3. Implement and analyze operations on binary search trees and AVL trees.
4. Implement graph traversal algorithms, find shortest paths and analyze them.

Reference Books

1. K R. Venugopal and Sudeep. R Prasad; Mastering C; Second Edition; McGraw Hill; 2015.

2. Ellis Horowitz, SartajSahni& Susan Anderson-Freed, Fundamentals of Data Structures in C,Second Edition, Universities Press, 2008.

3. Mark Allen Weiss; Data Structures and Algorithm Analysis in C; Second Edition; PearsonEducation; 2002.

Syllabus for Semester III, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDT3002 **Category:** Program Core Course (PCC)

Course: Theory of Computations **L:** 3Hrs**, T:** 0Hr**, P:** 0Hr**, Per Week, Credits:** 3

Course Objectives

The objective of this course is to familiarize the prospective engineers with:
1. To provide students an understanding of basic concepts in the theory of computation.
2. To teach formal languages and various models of computation.
3. To exhibit fundamental concepts related with computability theory.

SYLLABUS

UNIT-I Basics of Theory of Computation

Basics of Sets and Relation, Countability and Diagonalisation, Principle of mathematical induction,Pigeon- hole principle. Fundamentals of formal languages and grammars, Chomsky hierarchy of languages.

UNIT-II Finite Automata

Deterministic finite automata (DFA), Nondeterministic finite automata (NFA) and equivalence with DFA, Minimization of finite automata, NFA with Epsilon Transitions, Finite Automata with output.

UNIT-III Regular Grammar

Regular expressions and Regular languages, Regular grammars and equivalence with finite automata, properties of regular languages, pumping lemma for regular languages

UNIT-IV Context-Free Grammar

Context-free grammars (CFG) and language(CFL), parse trees, ambiguity in CFG, Reduction of CFGs, Chomsky and Greibach normal forms.

UNIT-V Push Down Automata

Deterministic pushdown automata and Non-Deterministic push down automata, Acceptance by two methods: Empty stack and Final State, Equivalence of PDA with CFG, closure properties of CFLs.

UNIT-VI Turing Machines

The basic model for Turing machines (TM), Turing recognizable recursively

enumerable) and Turing-decidable (recursive) languages, variants of Turing machines, unrestricted grammars and equivalence with Turing machines, TMs as enumerators.

Course Outcomes

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

1. Describe the formal relationships among machines, languages and grammars.
2. Design and Optimize finite automata for given regular languages.
3. Design context free grammar for a given language.
4. Design Pushdown Automata, Turing Machine for given languages.

Text Books

1. John Martin, Introduction to Languages and The Theory of Computation, Tata McGraw Hill

Reference Books

1. John E. Hopcroft, Rajeev Motwani and Jeffrey D. Ullman, Introduction to Automata Theory, Languages, and Computation, Pearson Education Asia.
2. Harry R. Lewis and Christos H. Papadimitriou, Elements of the Theory of Computation, Pearson Education Asia.
3. Dexter C. Kozen, Automata and Computability, Undergraduate Texts in Computer Science, Springer.
4. Michael Sipser, Introduction to the Theory of Computation, PWS Publishing.

Syllabus for Semester III, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDP3003 **Category:** Program Core Course (PCC)

Course: Programming Language Lab **L:** 0Hrs**, T:** 0Hr**, P:** 4Hr**, Per Week, Credits:** 2

Course Objectives:

1. The course focuses on developing the python programming skills to do a variety of programming tasks where the students are encouraged to develop application using python.
2. To cover the basic constructs of python programming, data structures and object-oriented concepts.
3. The course also targets the coverage of important modules and libraries available in python.

Course Outcome

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

1. Design Python Programs using Different Data and Control Structures.
2. Use Functions, Python Files, Modules and Packages to handle complex python programs.
3. Develop Mathematical Models and Scientific Applications in Python using various Libraries.
4. Write Small Python Applications using Web Scrapping and PyGame Modules.

Syllabus

- Introduction to Python: Basic Building Blocks of Python, Installation and Working with Python, Keywords, Variables and Operators.
- Data Types: int , float, complex, User Input, Arithmetic Expressions ,Using Strings and Operations on Strings, List, Slicing List and Strings, Split, Tuples, Dictionary, Sets.
- Flow Control: Conditional blocks: if, else, elseif, for Loops in Python: Loops with range, Strings, List and Dictionaries, While Loop, break, continue and pass statements.
- Python Functions: Library Functions, User Defined Functions, Function Argument Types, Recursion, Returning Multiple Values, Lambda, Map, Filter, Reduce.
- Python OOPS Basics: Classes, Object, Class Variable and Instance Variable.

- Files: Reading and Writing Files in Python, File Operations and Modes.Modules and Packages: To Create and Import Module.
- Open-Source Python library- Pandas, Sci Py, NumPy, Matplotlib and Seaborn.
- Web scrapping: Web scrapping with the help of standard libraries like Requests and Beautiful Soup.
- Python Pygame (Game Development Library)

Practical based on above theory syllabus

Text Books

1. Python Programming Using Problem Solving Approach: Reema Thareja, OxfordUniversity, Press; First edition.
2. Learning Python: Powerful object-oriented programming, Mark Lutz, O'REILLY publications 5th addition.
3. Introduction to Computing & Problem Solving with Python Jeeva Jose and P SojanLal Ascher.
4. Problem Solving with Algorithms and Data Structures using Python by Brad Miller and David Ranum, 2nd addition**.**

Reference Books

1. Allen Downey, Jeffrey Elkner, Chris Meyers, Learning with Python, Dreamtech Press
2. David M. Baezly “Python Cookbook” O'Reilly Media; Third edition, 2013.

Syllabus for Semester III, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDP3004 **Category:** (VSEC)

Course: Software Laboratory - I **L:** 0Hr**, T:** 0Hr**, P:** 2Hrs**, Per Week, Credits:** 1

Course Objectives:

1. Familiarize students with the use and applications of various statistical tools and packages in data science.
2. Provide hands-on experience in statistical programming and data analysis.
3. Equip students with the skills to perform exploratory data analysis, statistical modeling, and inference.

Course Outcome

Syllabus

1. **Data Exploration and Visualization**

Load a dataset containing information about housing prices in a city. Explore the dataset, calculate summary statistics, and visualize the distribution of housing prices using appropriate plots.

2. **Descriptive Statistics**

Given a dataset of student grades, calculate measures of central tendency (mean, median, mode) and measures of variability (variance, standard deviation). Interpret the results and discuss the implications of different measures.

3. **Probability Distributions**

Simulate a binomial distribution with given parameters (number of trials and probability of success). Plot the distribution and calculate the probability of different events occurring.

4. **Sampling Distributions and Confidence Intervals**

Generate a random sample from a population with known mean and standard deviation. Construct confidence intervals for the sample mean and interpret the results.

5. **Hypothesis Testing**

Analyze a dataset containing the weights of two different groups of individuals. Conduct a two-sample t-test to determine if there is a significant difference in mean weights between the two groups.

6. **Chi-square Test**

Given a dataset of categorical data, perform a chi-square test to determine if there is a significant association between two categorical variables.

7. **Analysis of Variance (ANOVA)**

Analyze a dataset containing crop yields from different fertilizer treatments applied to multiple plots of land. Conduct a one-way ANOVA to determine if there is a significant difference in mean crop yields among the different fertilizer treatments.

8. **Simple Linear Regression&Multiple Linear Regression**

Using a dataset of housing prices and square footage, fit a simple linear regression model to predict housing prices based on square footage. Evaluate the model's performance and interpret the regression coefficients.

Analyze a dataset containing information about cars, including factors such as engine size, fuel efficiency, and price. Fit a multiple linear regression model to predict the price of a car based on various features.

9. **Non-linear Regression**

Given a dataset of population growth over time, fit a non-linear regression model (e.g., exponential or logistic) to model the population growth pattern. Evaluate the model's fit and interpret the results.

10. **Logistic Regression**

Using a dataset containing customer data and purchase history, build a logistic regression model to predict whether a customer will make a purchase based on various factors (e.g., age, income, past purchases).

These problem statements cover a range of scenarios and applications, allowing students to practice data exploration, visualization, descriptive statistics, probability distributions, inferential statistics, regression analysis, and their implementation using statistical programming tools. Additionally, the case study problem statement encourages students to apply their skills to a real-world dataset and communicate their findings effectively.

Practical based on above theory syllabus

Course Outcomes:

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

1. Understand and apply different statistical techniques on given data.
2. Perform and interpret various probability distributions.
3. Carry out hypothesis testing and calculate confidence intervals.
4. Create and evaluate regression models.
5. Implement statistical analysis using programming tools and packages.
6. Analyze and visualize data using statistical programming techniques.
7. Apply statistical methods and programming skills to real-world data science problems.

Reference Books:

1. An Introduction to Statistics with Python: With Applications in the Life Science by Thomas Haslwanter, Springer.
2. Practical Statistics for Data Scientists by Peter Bruce, Andrew Bruce, Peter Gedeck, Reilly Media.

Syllabus for Semester III, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDP3006 **Category:** (HSSM)

Course: Idea Lab **L:** 0Hr**, T:** 0Hr**, P:** 4Hrs**, Per Week, Credits:** 2

Course Objectives

This course provides students with a unique opportunity to engage, explore, experience, express and excel in innovative thinking. The course will accelerate the development of indigenous solutions by inculcating creative skills, problem-solving skills and entrepreneurship skills in students. The ultimate goal is to help students build his/her creative acumen to address real life challenges.

Course Outcomes:

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

1. Identify a real-world challenge or potential opportunities that demand consideration.
2. Design and propose original ideas or innovative solutions to identified real-world challenge.
3. Build a plan to develop proof of concept (POC)/ prototype/minimum viable product (MVP) for the idea.
4. Critically evaluate the feasibility and viability of proposed idea, considering its impact on various stakeholders.

Execution Plan for the Subject:

- Conduction of lectures/guest lectures to familiarize the students with concepts of principles of design thinking, innovative and business oriented solution building.
- Students will identify the potential areas or real-world problems that require innovative solution.
- Students will contribute to the creation and presentation of their solution. The developed solution can be a prototype or an idea or a business plan or a software solution, etc.
- Students will analyze and provide constructive feedback on the ideas and solutions presented by self and peers.

Syllabus for Semester III, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDT3007 **Category:** Value Education Course (VEC)

Course: Cyber Laws and Ethics in IT **L:** 2Hrs, **T:** 0Hr, **P:** 0Hr, **Per Week, Credits:** 2

Course Objectives:

1. To explore ethical principles in the business world and IT organization
2. To understand various intellectual property issues such as copyrights, patents, Trade secrets, and trademarks.
3. To identify emerging plagiarism policies, anonymity issues, identity theft, and consumer profiling.
4. To interpret Indian Information Technology Act 2000 and explore cyber security issues in the society and business world.

Syllabus:

UNIT I

Ethics in business world & IT professional malpractices, Introduction to firewalls, IDS System, Distortion and fabrication of information

UNIT II

Ethics of IT Organization: Contingent Workers H- IB Workers, Whistle- blowing, Protection for Whistle- Blowers, Handling Whistle- blowing situation, Digital divide.

UNIT III

Intellectual Property: Copyrights, Patents, Trade Secret Laws, Key Intellectual property issues, Plagiarism, Privacy: The right of Privacy, Protection, Key Privacy and K-Anonymity issues, Identity Theft, Consumer Profiling,

UNIT IV

Cyber laws and rights in today's digital age, Emergence of Cyberspace, Cyber Jurisprudence, Cyber Crimes against Individuals, Institution and State, Hacking, Digital Forgery, Cyber Stalking/Harassment, Cyber terrorism, cyber tort, Cyber Defamation & hate speech,Competitive Intelligence, Cybersquatting, The IndianInformation Technology Act 2000 IT Act.

Course Outcomes

On successful completion, of course student will able to learn:

1. To analyze the role of ethics in IT organization.
2. To identify various cyber laws with respect to legal dilemmas in the Information Technology field.

3. To interpret various intellectual property rights, Privacy, Protection issues in Information Technology field.
4. To describe the ways of precaution and prevention of Cyber Crime as well as Human Rights.

Text Books:

1. George Reynolds, "Ethics in information Technology",5th edition, Cengage Learning
2. Hon C Graff, Cryptography and E-Commerce - A Wiley Tech Brief, Wiley Computer Publisher, 2001.

Reference Books:

1. Michael Cross, Norris L Johnson, Tony Piltzecker, Security, Shroff Publishers and Distributors Ltd.
2. Debora Johnson," Computer Ethic s",3/e Pearson Education.
3. Sara Baase, "A Gift of Fire: Social, Legal and Ethical Issues, for Computing and the Internet," PHI Public at ions.
4. Chris Reed & John Angel, Computer Law, OUP, New York, (2007).
5. Dr Pramod Kr.Singh, "Laws on Cyber Crimes [Along with IT Act and Relevant Rules]" Book Enclave Jaipur India.

Syllabus for Semester IV, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDT4001 **Category:** Program Core Course (PCC)

Course: Operating System **L:** 3Hrs, **T:** 0Hr, **P:** 0Hr, **Per Week, Credits:** 3

Course Objectives

The objective of this course is to familiarize the prospective engineers with:

1. To learn how Operating System is Important for Computer System.
2. To make aware of different types of Operating System and their services.
3. To learn different process scheduling algorithms and synchronization techniques to achieve better performance of a computer system.
4. To know memory organization and management concepts.
5. To learn secondary memory, directory and file management

SYLLABUS

Unit I: Introduction:

Concept of Operating Systems, Generations of Operating systems, Types of Operating Systems, OS Services, System Calls, Structure of an OS - Layered, Monolithic, Microkernel Operating Systems, Concept of Virtual Machine, Case study on different Operating Systems.

Unit II: Processes and CPU Scheduling:

Definition, Process Relationship, Different states of a Process, Process State transitions, Process Control Block (PCB), Context switching

Threads:Definition, Various states, Benefits of threads, Types of threads, Concept of multithreads.

ProcessScheduling: Foundation and Scheduling objectives, Types of Schedulers, Scheduling criteria: CPU utilization, Throughput, Turnaround Time, Waiting Time, Response Time; Scheduling algorithms: Pre-emptive and Non pre-emptive, FCFS, SRTF, Priority, RR

Unit III: ProcessSynchronization:

Critical Section, Mutual Exclusion, Peterson's solution, Hardware Solution, Software Solutions, Semaphores, Monitors, Message Passing, Classical IPC Problems: Producer-Consumer Problem, Reader-Writer Problem, Dinning Philosopher Problem etc.

Unit IV: Deadlocks:

Definition, Necessary and sufficient conditions for Deadlock, Starvation, Deadlock Prevention, Deadlock Avoidance: Banker's algorithm, Deadlock detection and Recovery.

Unit V: MemoryManagement:

Basic concept, Logical and Physical address mapping, Memory allocation: Contiguous Memory allocation, Fixed and variable partition, Internal and External fragmentation and Compaction,

Paging: Page allocation, Hardware support for paging, Advantages & Disadvantages of paging. Segmentation, Virtual Memory: Basics of Virtual Memory

Demand Paging : Page Fault Mechanism, Page Replacement, Page Replacement algorithms: First in First Out (FIFO), Least Recently used (LRU), and Optimal.

Unit VI: FileManagement:

Concept of File, Access methods, File types, File operations, Directory structure, File System structure

DiskManagement**:** Disk structure, Disk scheduling - FCFS, SSTF, SCAN, C-SCAN, LOOK, C-LOOK

Course Outcomes

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

1. Understands the different services provided by Operating System at different level.
2. Understanding the performance and design trade-offs in complex software systems
3. Understand the role of various components (process, memory, file systems, etc.) of operating system.
4. Understands the use of different process scheduling algorithm and synchronization techniques to avoid deadlock
5. Analyze and apply resource (CPU, Memory, Disk) management policies.

Text Books

1. Operating System Concepts, 8th Edition by A. Silberschatz, P.Galvin, G. Gagne,Wiley India Edition.
2. Modern Operating Systems, 2nd Edition by Andrew Tanenbaum, PHI.

Reference Books

1. Operating Systems: Internals and Design Principles, 5th Edition, William Stallings,Prentice Hall of India.
2. Operating Systems, W. Stalling, Macmillan.
3. Operating Systems, H. M. Dietel, Addison Wesley Longman.
4. Understanding the Linux Kernel, 3rd Edition, Daniel P. Bovet, Marco Cesati, O'Reilly.
5. Linux Kernel Development, Robert Love, Third Edition, Addison-Wesley, 2010.
6. The design of Unix Operating system, Maurice J. Bach, Pearson Education, India.

Syllabus for Semester IV, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDP4001 **Category:** Program Core Course (PCC)

Course: Operating System Lab **L:** 0Hr**, T:** 0Hr**, P:** 2Hrs**, Per Week, Credits:** 1

SYLLABUS

Practical based on CDT4001 Syllabus

1. Linux System Commands
2. Execution of I/O system calls
3. CPU scheduling algorithms
4. Inter process communication
5. Process synchronization
6. Deadlock Detection
7. Memory Management
8. Disk management

Course Outcomes

On Successful completion of course, students will be able to:

1. Implement system commands by making use of LINUX system calls.
2. Implement processes and process schedulers.
3. Design solutions to process synchronization and deadlock handling.
4. Implement Memory management and File management solutions.

Syllabus for Semester IV, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDT4002 **Category:** Program Core Course (PCC)

Course: Artificial Intelligence **L:** 3Hrs**, T:** 0Hr**, P:** 0Hr**, Per Week, Credits:** 3

Course Objectives

The objective of this course is to familiarize the prospective engineers with:

1. To understand challenges involved in designing intelligent systems.
2. To represent given problem using state space representation and solve it by using different search techniques.
3. To understand knowledge representation methods using logic programming.
4. To understand uncertainty theory in designing AI systems.

SYLLABUS

UNIT I:

Introduction: Basics of problem solving, problem representation (toy problems and real world problems); Structure of agent, rational agent, Specifying task environment, Properties of task environment; measuring problem-solving performance

UNIT II:

Uninformed search techniques: Depth, Breadth, Uniform Cost, Depth Limited, Iterative deepening DFS, Bidirectional Search

UNIT III:

Informed search techniques: Heuristic Based Search, Greedy Best First Search, A* Search; Local Search algorithms: Hill-climbing, Simulated Annealing, Genetic Algorithms.

UNIT IV:

Adversarial Search: Two player Games, The min-max algorithm, Alpha-Beta pruning. Constraint Satisfaction Problems: Constraint propagation, backtracking search

UNIT V:

Propositional Logic: Inference, Equivalence, Validity and satisfiability, Resolution, Forward and Backward Chaining, First Order Logic: Syntax and Semantics of FOL, Inference in FOL, Unification, Forward Chaining, Backward Chaining, and Resolution.

UNIT VI:

Uncertainty Knowledge and Reasoning: Probability and Baye's Theorem, Statistical reasoning: Bayesian networks, Naïve bayes algorithm, Fuzzy Logic, Introduction to expert system

Course Outcomes

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

1. Represent given problem using state space representation
2. Apply uninformed and informed search techniques.
3. Solve the fully informed two player games using different AI techniques.
4. Solve the AI problems by using logic programming
5. Apply uncertainty theory based on techniques like probability theory and fuzzy logic.

Text Book:

1. Stuart Russel and Peter Norvig; Artificial Intelligence: A Modern Approach; Third Edition; Pearson Education, 2009.

Reference Books:

1. E.Rich, K. Knight, S. B. Nair; Artificial Intelligence; 3rd Edition; Tata McGraw Hill, 2014.
2. Denis Rothman; Artificial Intelligence By Example: Develop machine intelligence from scratch using real artificial intelligence use cases; Kindle Edition, Packt Publishing Ltd, 2018

Syllabus for Semester IV, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDP4002 **Category:** Program Core Course (PCC)

Course: Artificial Intelligence Lab **L:** 0Hrs**, T:** 0Hr**, P:** 2Hr**, Per Week, Credits:** 1

Course Outcomes:

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

1. Implement different AI toy problems by using search techniques.
2. Design two player games using min-max algorithm with Alpha-Beta pruning.
3. Simulate AI problems using logic programming.
4. Implement algorithms to handle uncertainity.

PRACTICALS BASED ON CDT4002 SYLLABUS

Reference Books

1. Stuart Russel and Peter Norvig; Artificial Intelligence: A Modern Approach; Third Edition; Pearson Education.
2. E.Rich, K. Knight, S. B. Nair; Artificial Intelligence; 3rd Edition; Tata McGraw Hill.
3. Denis Rothman; Artificial Intelligence By Example: Develop machine intelligence from scratch using real artificial intelligence use cases; Kindle Edition, Packt Publishing Ltd, 2018

Syllabus for Semester IV, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDP4003	**Category:** (PCC)
Course: Design and Analysis of Algorithms	**L:** 3Hr**, T:** 0Hr**, P:** 0Hrs**, Per Week, Credits:** 3

Course Objectives:

The objective of this course is to familiarize the prospective engineers with:

1. Techniques of effective problem solving in computing.
2. Analysis of different paradigms of problem solving to solve a given problem in efficient way.

Syllabus

UNIT-I

Algorithm Design Basics and Recurrence Relations

Principles of designing algorithms and complexity calculation, Asymptotic notations for analysis of algorithms, worst case and average case analysis, amortized analysis and it's applications. Recurrence relations and their solutions using substitution, recurrence tree and master theorem methods

UNIT-II

Divide and Conquer technique

Introduction ,Basic strategy, Binary Search, finding Maximum and Minimum elements from array, Quick sort, Merge sort, Strassen's matrix multiplication, Maximum sub-array problem, Closest pair of points problem.

UNIT-III

Greedy technique

Introduction ,Basic strategy, fractional knapsack problem, Minimum cost spanning trees, Huffman Coding , activity selection problem, Scheduling problem with and without deadlines, Find maximum sum possible equal to sum of three stacks, K Centers Problem.

UNIT-IV

Dynamic Programming technique

Introduction ,Basic strategy, difference between Dynamic and Greedy approach, Bellmen ford algorithm, all pairs shortest path algorithm, multistage graphs, optimal binary search trees, traveling salesman problem, String Editing, Longest Common Subsequence problem and its variations.

UNIT-V

Backtracking and Branch and bound techniques

Introduction ,Basic strategy ,N-Queen's problem, Vertex coloring problem, Hamiltonian circuit problem, sum of subset problem, Branch and bound general technique, applications, travelling sales person problem

UNIT-VI

P and NP class

Basic concept, non-deterministic algorithms NP-hard and NP-complete problems, Cook's Theorem, decision and optimization problems, polynomial reduction, , Introduction to Approximation algorithm. ,vertex cover problem, clique cover problem

Course Outcomes

After successful completion of this course, the student will be able to,

1. Use mathematical formulation, complexity analysis and methodologies to solve the recurrence relations for algorithms.
2. Apply Greedy and Divide and Conquer algorithms and their usage in real life examples.
3. Apply Dynamic programming and Backtracking and branch and bound Paradigms to solve the real life problems.
4. Analyze P and NP class problems and formulate solutions using standard approaches.

Text Books

1. Thomas H. Cormen et.al; "Introduction to Algorithms"; 3 Edition; Prentice Hall, 2009.
2. Horowitz, Sahani and Rajasekaram; "Computer Algorithms", Silicon Press, 2008.
3. Brassard and Bratley; "Fundamentals of Algorithms", 1 Edition; Prentice Hall, 1995.
4. Richard Johnsonbaugh, "Algorithms", Pearson Publication, 2003.

Reference books

1. Parag Himanshu Dave, Balchandra Dave, "Design and Analysis of Algorithms" Pearson Education, O'relly publication
2. Richard Johnsonbaugh, "Algorithms", Pearson Publication, 2003.

Syllabus for Semester IV, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDP4005 **Category:** (VSEC)

Course: Software Laboratory - II **L:** 0Hr**, T:** 0Hr**, P:** 2Hrs**, Per Week, Credits:** 1

Course Objectives:

The objective of this course is to impart necessary and practical knowledge of recent Software strategies based on frameworks and development skills required to design real-life web based projects by using advanced object oriented programming by :

1. Designing Enterprise applications by encapsulating an application's business logic.
2. Designing and developing multi-tier web applications
3. Designing applications using existing frameworks like Spring and Hibernate.

Experiments based on:

JDBC Java Database Connectivity (JDBC): The Design of JDBC, Basics of Structured

Query Language, JDBC Configuration, Executing a basic SQL Statement.

Servlet: Handling the Client Request, Generating HTML response, Reading form Data from Servlets: Reading Three Parameters, managing a session.

Java Server Pages (JSP): Invoking Java Code with JSP Scripting Elements, JSP directives and actions, Integrating Servlets and JSP

Spring Framework: Spring Bean Life Cycle, Spring Bean Scope, Basic Bean Wiring

Hibernate: Hibernate Configuration, Hibernate Sessions, Collections Mappings, Hibernate Query Language

Course Outcomes

On successful completion of the course, the student will be able to:

1. Implement Software based database application.
2. Demonstrate server and client side programming using servlets and Java server pages.
3. Perform Dependency Injection using Spring, and create mappings in Hibernate using HQL.

Text Books:

1. M. Deitel, P. J. Deitel, S. E. Santry; Advanced Java 2 Platform HOW TO PROGRAM; Prentice Hall.
2. Cay Horstman, Gary Cornell; Core JAVA Volume-II Advanced Features; 8th Edition.

3. Craig Walls; Spring In Action; 2nd Edition
4. Marty Hall, Larry Brown; Core Servlets and Java Server Pages Volume-1: Core Technologies; 2nd Edition

Reference Books:

1. Jim Keogh; "J2EE:The Complete Reference"; McGraw Hill; Fifth Edition.
2. Spring Framework Documentation https://spring.io/
3. Hibernate Framework architecture Documentation https://hibernate.org/

Syllabus for Semester IV, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDP4006 **Category:** (CEP)

Course: Community Engagement Project **L:** 0Hrs**, T:** 0Hr**, P:** 4Hr**, Per Week, Credits:** 2

Course Objectives

The objective of Community Engagement Project is to instill a sense of social responsibility amongst the students, empowering them to apply their knowledge and skills to positively impact and contribute to the society.

Execution Plan for the Subject:

- The students will impart their knowledge and skills in the society by identifying the potential needs or identify a society need and address it by building a technical solution.

Course Outcomes

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

1. Propose a community engagement project tailored to address society needs by devising a strategy or solution to address it.
2. Apply technical knowledge or skills towards execution of the proposed solution.
3. Evaluate the effectiveness of the project in addressing community needs.
4. Demonstrate ethical principles, project management skills, team work and communication skills for project completion within the confines of a deadline.

B. Tech. Computer Science and Engineering (Data Science) [2023-24 NEP]
Honors and Minors Scheme

Following courses are proposed under Honors in CSE(Data Science)

Sr. No.	Category	Course code	Course Name	Hours/ week			Credits	Maximum marks			ESE Duration (Hrs)
				L	T	P		Continuous Evaluation	End Sem Exam	Total	
1	Honor	CDTH	Data Science Essentials	3	0	0	3	50	50	100	-
2	Honor	CDTH	System Architecture Analysis	3	0	0	3	50	50	100	-
3	Honor	CDTH	Data Engineering	4	0	0	4	50	50	100	-
4	Honor	CDTH	Business and Web Analytics	4	0	0	4	50	50	100	-
5	Honor	CDTH	Project	4	0	0	4	50	50	100	-
			TOTAL	**18**	**0**	**0**	**18**	**250**	**250**	**500**	

Note : Marks division for all the courses mentioned above (CSE(DS) Honors) will be of 50 marks for Continuous Evaluation and 50 marks for End Sem Exam. Evaluation pattern for Continuous Evaluation and ESE can be lab conduction, viva voce, execution based test and MCQ.

Scheme for Minors

Following courses are proposed under Minors in CSE(Data Science)

Sr. No.	Category	Course code	Course Name	Hours/ week			Credits	Maximum marks			ESE Duration (Hrs)
				L	T	P		Continuous Evaluation	End Sem Exam	Total	
1	Minor	CDTM	Foundation of Data Science	3	0	0	3	50	50	100	3
2	Minor	CDTM	Statistics for Data Science	3	0	0	3	50	50	100	3
3	Minor	CDTM	Exploratory Data Analysis	4	0	0	4	50	50	100	3
4	Minor	CDTM	AI tools for Data Science	4	0	0	4	50	50	100	3
8	Minor	CDTM	Project	4	0	0	4	50	50	100	-
			TOTAL	**18**	**0**	**0**	**18**	**250**	**250**	**500**	

Syllabus for Semester I, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDTM **Category:** Minor

Course: Foundation of Data Science **L:** 3Hrs**, T:** 0Hr**, P:** 0Hr**, Per Week, Credits:** 3

Course Objectives

The objective of this course is to familiarize the prospective engineers with:

1. The fundamentals of data science and the basics of Python programming constructs.
2. Representation of complex data using python data structures - lists, tuples, dictionaries.
3. Manipulation of Numpy arrays and Pandas series and data frames
4. Data Wrangling techniques for Data Analysis

SYLLABUS

UNIT-I

Introduction

Introduction to Data Science, Benefits and uses, Facets of data, Evolution of Data Science – Data Science Roles – Data science process – Applications of Data Science in various fields – Data Security Issues.

UNIT-II

Python Basics

Basics of Python including data types, operators, variables, expressions, control structures using sample dataset, objects and functions, Class – Constructors – Object Creation – Inheritance – Overloading. Text Files and Binary Files – Reading and Writing, Decorators

UNIT-III

Data Structures

Python sequence data structures including – creation and manipulation String, Array, List, Tuple, Set, and Dictionary. Introduction to various python libraries for data science – Numpy, Pandas, seaborn, matplotlib, SciPy.

Unit – IV: Numpy and Pandas Packages

NumPy - NumPy ndarray:- Vectorization Operation - Array Indexing and Slicing - Transposing Array and Swapping Axes - Saving and Loading Array - Universal Functions - Mathematical and Statistical Functions in Numpy.

Series and DataFrame in pandas – Creating, manipulating and reading DataFrames, Panda's Index Objects - Reindexing Series and DataFrames - Dropping entries from Series and Data Frames - Indexing, Selection and Filtering in Series and Data Frames - Arithmetic Operations between Data Frames and Series - Function Application and Mapping.

UNIT-V

Data Wrangling – Clean, Transfer, Merge and Reshape

Combining and Merging Data Sets – Reshaping and Pivoting – Data Transformation – String manipulations – Regular Expressions.

Unit – VI: Data Aggregation and Group Operations

GroupBy Mechanics – Data Aggregation – GroupWise Operations and Transformations – Pivot Tables and Cross Tabulations – Date and Time data types.

Course Outcomes

After successful completion of this course, the student will be able to,

1. Identify the need for data science and solve basic problems using Python conditionals, loops and functions.
2. Determine the methods to create and manipulate Python programs by utilizing the data structures like Array, lists, tuples, set and dictionaries.
3. Apply the efficient storage and data operations using NumPy arrays. and perform powerful data manipulations using Pandas
4. Interpret the data wrangling process including cleaning, transferring, merging and and reshape
5. Use the aggregations and group operations for data analysis in python.

Text Books

1. Davy Cielen, Arno D. B. Meysman, Mohamed Ali, "Introducing Data Science", manning publications 2016.
2. Wes Mckinney. Python for Data Analysis. O'Reilly Media, 2013.
3. Grus, Joel. Data Science from Scratch: First Principles with Python. O'Reilly Media, 2015.

Reference books

1. Allen B. Downey, ``Think Python: How to Think Like a Computer Scientist'', 2nd edition,Updated for Python 3, Shroff/O'Reilly Publishers, 2016
2. David J. Pine, Introduction to Python for Science and Engineering, CRC Press, 2019.
3. Jake vander Plas, Python Data Science Handbook – Essential Tools for Working with Data, O'Really Media, 2017

Syllabus for Semester I, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDTM **Category:** Minor

Course: Statistics for Data Science **L:** 3Hrs, **T:** 0Hr, **P:** 0Hr, **Per Week, Credits:** 3

Course Objectives

The primary objective of this course is to provide students with a comprehensive understanding of statistical concepts, methods, and techniques essential for data analysis and decision-making processes. Students will learn how to collect, organize, analyze, and interpret data using various statistical tools and methodologies.

SYLLABUS

Unit 1:

Introduction, Variables and Types of data, Data collection and sampling techniques, uses and misuses of statistics, Organizing data, histograms, frequency polygons, ogives, other types of graphs.

Unit 2:

Measures of central tendency, measures of variation, measures of position, five-number summary, boxplots, etc.

Unit 3:

Probability distributions, normal distributions, binomial distribution, other types of distributions, central limit theorem.

Unit 4:

Confidence interval and sample size, Hypothesis testing.

Unit 5:

Testing the Difference Between Two Means, Two Proportions, and Two Variances, Correlation and Covariance.

Unit 6:

Simple linear regression, multiple linear regression, logistic regression.

Course Outcomes

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

1. Apply sampling techniques to generate appropriate samples.
2. Apply descriptive statistical techniques for data analysis.
3. Interpret data to perform hypothesis testing.
4. Apply regression models.

Text Books:

1. Elementary Statistics A Step by Step Approach by Allan G. Bluman , McGraw Hill Publications, Seventh Edition.

2. Practical Statistics for Data Scientists by Peter Bruce and Andrew Bruce, O`Reilly Publications.

B. Tech. Computer Science and Engineering (Data Science) [2023-24 NEP]
Teaching & Evaluation Scheme

OPEN ELECTIVE SCHEME

Sr. No.	Category	Course code	Course Name	Hours/ week			Credits	Maximum marks			ESE Duration (Hrs)
				L	T	P		Continuous Evaluation	End Sem Exam	Total	
1	OE	CDT3005	Introduction to Data Science	4	0	0	4	50	50	100	2
2	OE	CDT4004	Tools for Data Science	2	0	0	2	50	50	100	2
3	OE	CDT5006	Data Analytics Visualization	2	0	0	2	50	50	100	2
			TOTAL	**08**	**0**	**0**	**08**	**150**	**150**	**300**	

Syllabus for Semester III, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDT3005 **Category:** OPEN ELECTIVE (OE)

Course: Introduction to Data Science **L:** 4Hrs**, T:** 0Hr**, P:** 0Hr**, Per Week, Credits:** 4

Course Objectives

The objective of this course is to familiarize the prospective engineers with:

1. The fundamentals of data science, including its benefits, uses, and various facets.
2. The proficiency in data cleaning, integration, transformation, reduction, and discretization techniques.
3. The concept of a random sample and different sampling techniques., explore statistics and sampling distributions and Hypothesis and Inference.
4. Descriptive statistics such as mean, standard deviation, skewness, and kurtosis for data exploration.
5. The fundamentals of conditional probability, random variables and distribution and regression analysis

SYLLABUS

UNIT-I

Introduction

Introduction to Data Science, Benefits and uses, Facets of data , Evolution of Data Science – Data Science Roles – Data science process – Applications of Data Science in various fields – Data Security Issues.

UNIT-II

Data Collection and Data Pre-Processing

Data Collection Strategies – Data Pre-Processing Overview – Data Cleaning – Data Integration and Transformation – Data Reduction – Data Discretization

UNIT-III

Probability and Random Variables

Basic concepts of probability, Conditional probability, total probability, independent events, Bayes' theorem, Random variables, Moments and moment generating functions, Useful distributions, Joint distribution, conditional distribution, Transformation of random variables, Covariance

UNIT-IV

Statistics and Sampling

Inferential statistics, Random sampling and sampling techniques, Statistics and sampling distributions, Mixture models, Statistical hypothesis testing, p-values, confidence intervals, p-hacking, Bayesian inference

UNIT-V

Exploratory Data Analytics

Descriptive Statistics –

Measure of Central Tendency - Mean, Standard Deviation, Measures of Dispersion, Skewness &

Kurtosis, Mean square deviation, Correlation

UNIT-VI

Advanced Techniques and Applications

Regression Analysis Introduction, Linear Regression - Simple Linear Regression, Multiple Linear Regression and Polynomial Regression, Sparse Model, Logistic Regression, Case studies and real world examples

Course Outcomes

After successful completion of this course, the student will be able to,

1. Describe the concepts of data science including evaluation of data science, data collection data pre-processing techniques.
2. Calculate and interpret probabilities, random variable and random processes.
3. Recognize how descriptive Statistics, inferential statistics, modelling and statistical computing can be utilized in an integrated capacity.
4. Understand basics of Regression Analysis technique in data science.

Text Books

1. Davy Cielen, Arno D. B. Meysman, Mohamed Ali, "Introducing Data Science", manning publications 2016.
2. Joel Grus, "Data Science from scratch", O'Reilly, 2015.
3. Laura Igual , Santi Seguí, "Introduction to Data Science - A Python Approach to Concepts, Techniques and Applications , Springer, 2017
4. The Elements of Data Analytic Style by Jeff Leek
5. Exploratory Data Analysis with R, by Roger Peng

Reference books

6. Murtaza Haider, "Getting Started with Data Science – Making Sense of Data with Analytics", IBM press, E-book.
7. Lillian Pierson, "Data Science for Dummies", 2017, 2nd Edition.
8. Roger Peng, "The Art of Data Science", lulu.com 2016

Syllabus for Semester IV, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDT4004 **Category:** OPEN ELECTIVE (OE)

Course: Tools for Data Science **L:** 2Hrs, **T:** 0Hr, **P:** 0Hr, **Per Week, Credits:** 2

Course Objectives:

In this course, students will learn about

1. Handling Data with Python Libraries
2. Extracting and Analyzing data from different resources
3. Data Analysis with Pandas, Visualization using matplotlib, Seaborn.
4. Learn powerful R tools for solving data problems with greater clarity and ease.
5. Examine your data, generate hypotheses and test them and transform it into convenient for analysis.

Syllabus:

Unit 1: Introduction to Python for Data Science

Python installation and setup, Python syntax and data types, Data structures (lists, tuples, dictionaries, sets), Control flow statements (if-else, loops), Functions and modules, File handling

Unit 2: Data Manipulation with Python

Introduction to NumPy, NumPy arrays and operations, Data manipulation with Pandas, Reading and writing data in different formats, Data cleaning and preprocessing

Unit 3: Data Visualization and Analysis with Python

Introduction to Matplotlib, Basic plotting (line plots, scatter plots, bar charts), Advanced visualization with Seaborn, Descriptive statistics with Pandas, Introduction to scikit-learn for machine learning.

Unit 4: Introduction to R for Data Science

R installation and RStudio, R syntax and data structures (vectors, matrices, lists, data frames), Control flow statements and functions, Working with packages and libraries.

Unit 5: Data Manipulation with R

Reading and writing data in different formats, Data manipulation with dplyr and

tidyr, Data cleaning and preprocessing, Working with dates and times.

Unit 6: Data Visualization and Analysis with R

Introduction to ggplot2, Basic plotting techniques (scatter plots, line plots, bar charts),

Advanced visualization with ggplot2 extensions, Descriptive statistics and hypothesis testing, Linear regression and introduction to machine learning.

Course Outcomes:
On completion of the course the student will be able to

1.Apply programming language concept such as control structures and functions.

2.Use of libraries to analyze and manipulate data.

3.Demonstrate different approaches to data visualisation data structure.

Text Book:

1.Data Visualization with Python: Create an impact with meaningful data insights using interactive and engaging visuals by Tim Gromann, Packt Publishing.

2.Practical Python Data Visualization, by Ashwin Pajankar, Apress.

3. R Programming for Data Science, Roger D. Peng, Lean Publishing.

4. R for Data Science, Hadley Wickham & Garrett Grolemund, O'Reilly Publishing.

Reference books:

1. Python Programming Using Problem Solving Approach: Reema Thareja, Oxford University Press; First edition

2. Data Visualization and Exploration with R, by Eric Pimpler, Geospatial Training Services.

Syllabus for Semester V, B. Tech. Computer Science & Engineering (Data Science)

Course Code: CDT5006 **Category:** OPEN ELECTIVE (OE)

Course: Data Analytics Visualization **L:** 3Hrs**, T:** 0Hr**, P:** 0Hr**, Per Week, Credits:** 3

Course Objectives

The objective of this course is to familiarize the prospective engineers with:

1. To understand data analytics life cycle for solving challenging business problems.
2. To adopt appropriate statistical procedures for analysis based on goals and nature of data.
3. To employ best practices in data visualization to develop charts, maps tables and other visual representations of data.

SYLLABUS

Unit I: Introduction and Overview

Importance of analytics and visualization, data pre-processing, Data Analytics Lifecycle and Different Phases.

Unit II: Basics of Statistics

Variables and Types of data, Data collection and sampling techniques, uses and misuses of statistics, Organizing data, histograms, frequency polygons, other types of graphs.

Unit III: Probability and distribution

Basic concepts of probability, random variables, probability distributions, sampling and estimation, statistical inference

Unit IV: Basic Analysis Techniques

Basic Analysis Techniques: Statistical hypothesis generation and testing, Chi-Square test, t-Test, Analysis of variance, Correlation analysis, Maximum likelihood test.

Unit V: Data Visualization

Understanding Data Visualization Principles, Mapping Data onto Aesthetics, Visualizing - Distributions, Proportions, Time Series, Trends and Uncertainty

Unit VI: Creating Stories with Data

Why Planning?, Creating Interesting Stories with Data – Reader-driven Narratives, Author-driven Narratives; Perceptions and Presentation Methods, Best Practices in Visualization.

Course Outcomes

After successful completion of this course, the student will be able to,

1. Apply data pre-processing and basic analysis techniques.
2. Conduct data analytics using scientific methods.
3. Interpret data to perform hypothesis testing
4. Create presentations and visualizations.

Text Books

1. David Dietrich, Barry Heller and Beibel Yang, - Data Science and Big Data Analytics – Discovering, Analyzing, Visualizing, and Presenting Data, John Wiley and Sons [EMC Education Services], 2015.
2. Elementary Statistics A Step by Step Approach by Allan G. Bluman , McGraw Hill Publications, Seventh Edition. 2. Practical Statistics for Data Scientists by Peter Bruce and Andrew Bruce, O`Reilly Publications.
3. Claus O. Wilke, - Fundamentals of Data Visualization – A Primer on Making Informative and Compelling Figures, O'Reilly, 2019.

Reference books

1. Jiawei Han, Micheline Kamber and Jian Pei, - Data Mining Concepts and Techniques, 3rdedition; Morgan Kaufmann Publishers, 2011.