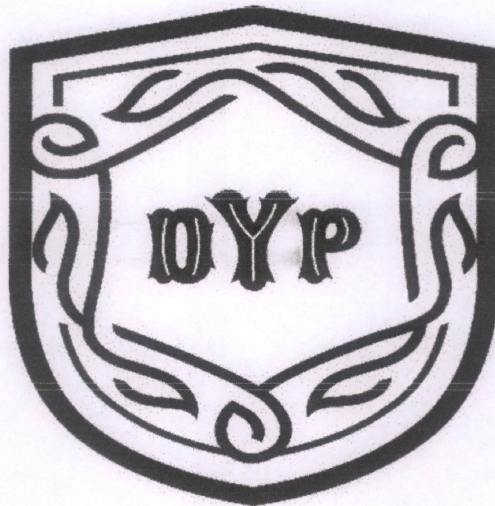# D. Y. Patil College of Engineering & Technology,

Kasaba Bawada, Kolhapur

(An Autonomous Institute)

Accredited by NAAC with 'A' Grade

Structure & Syllabus

for

Second Year B. Tech

of

Computer Science & Engineering (Data Science)

(With effect from 2024-25)

D. Y. PATIL COLLEGE OF ENGINEERING & TECHNOLOGY, KOLHAPUR

Teaching and Evaluation Scheme from Year 2024-25 (as per NEP-2020)

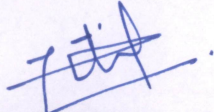## Second Year B. Tech. Computer Science & Engineering (Data Science)

SEMESTER - III

Sr. No	Course Code	Course Type	Course Name	Teaching Scheme				Theory			Practical/ Tutorial		Total Marks
				Credits	Contact Hrs			ISE	MSE	ESE	INT	OE/ PoE	
					L	P	T						
1	231DSPCCL201	PCC	Probability & Statistics	3	3	-	-	20	30	50	-	-	100
2	231DSPCCL202	PCC	Data Structures	3	3	-	-	20	30	50	-	-	100
3	231DSPCCP202	PCC	Data Structures Laboratory	1	-	2	-	-	-	-	25	25	50
4	231DSPCCP203	PCC	Programming Lab - I	3	2	2	-	-	-	-	50	50	100
5	231DSMDML201	MDM-I	Fundamentals of Data Science	2	2	-	-	20	-	30	-	-	50
6	231DSOECL201	OEC-I $	Data Science for Everyone	4	3	2	-	20	30	50	25	-	125
7	231DSFPP201	CEP/FP	Domain Specific Mini Project	2	-	4	-	-	-	-	25	25	50
8	231DSVECL201	VEC	Environmental Studies	2	2	-	-	-	-	50	-	-	50
9	231DSHSSML201	HSSM	Economics and Management for IT	2	2	-	-	20	30	-	-	-	50
10	231DSCCA201	CCA	Liberal Learning	-	-	2*	-	-	-	-	50*	-	Grade
11	231DSMCL201	MC	Finishing School Training - III	-	3*	-	-	-	-	-	50*	-	Grade
			Total	22	17	10	-	100	120	230	125	100	675

$ - Open & Distance Learning   * - Values are not included in total   **Min. Marks for Passing**: 40% of total marks of individual course

Program Coordinator   HOD CSE(Data Science)   Dean Academics   Principal

D. Y. PATIL COLLEGE OF ENGINEERING & TECHNOLOGY, KOLHAPUR

Teaching and Evaluation Scheme from Year 2024-25 (as per NEP-2020)

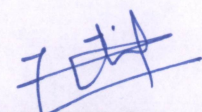## Second Year B. Tech. Computer Science & Engineering (Data Science)

SEMESTER - IV

Sr. No	Course Code	Course Type	Course Name	Teaching Scheme				Theory			Practical/ Tutorial		Total Marks
				Credits	Contact Hrs			ISE	MSE	ESE	INT	OE/ PoE	
					L	P	T						
1	231DSPCCL204	PCC	Discrete Mathematical Structure	3	3	-	-	20	30	50	-	-	100
2	231DSPCCL205	PCC	Design and Analysis of Algorithm	3	3	-	-	20	30	50	-	-	100
3	231DSPCCP206	PCC	Programming Lab - II	4	2	4	-	-	-	-	50	75	125
4	231DSMDML202	MDM-II	Methodology of Data Analytics using Machine Learning	2	2	-	-	20	-	30	-	-	50
5	231DSOECL202	OEC-II	Introduction to Data Engineering	2	2	-	-	-	-	50	-	-	50
6	231DSAECP201	AEC	Soft Skill	2	-	4	-	-	-	-	25	25	50
7	231DSVSECP201	VSEC	Web Application Development - I	2	1	2	-	-	-	-	25	50	75
8	231DSVECL202	VEC	Human Values and Ethics	2	2	-	-	-	-	50	-	-	50
9	231DSHSSML202	HSSM	Programming Ethics	2	2	-	-	-	-	50	-	-	50
10	231DSCCA202	CCA	Liberal Learning	-	-	2*	-	-	-	-	50*	-	Grade
11	231DSMCL202	MC	Finishing School Training - IV	-	3*	-	-	-	-	-	50*	-	Grade
			Total	**22**	**17**	**10**	**-**	**60**	**60**	**280**	**100**	**150**	**650**

* - Values are not included in total   **Min. Marks for Passing**: 40% of total marks of individual course

Program Coordinator   HOD CSE(Data Science)   Dean Academics   Principal

Course Type Abbreviation	Course Type Definition
PCC	Professional Core Course
CEP/FP	Comm. Engg. Project/Field Project
MDM	Multi-Disciplinary Minor
VEC	Value Education Course
OEC	Open Elective Course
HSSM	Humanities Social Science & Management
MC	Mandatory Course
CCA	Co-Curricular Activities
AEC	Ability Enhancement Course
VSEC	Vocational Skills Enhancement Course

Abbreviations:

ISE: In Semester Evaluation

MSE: Mid Semester Evaluation

ESE: End Semester Evaluation

INT: Internal Evaluation

OE: Oral Examination

PoE: Practical Oral Examination

Semester – III

D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – III
Course Code: 231DSPCCL201
Course Name: Probability & Statistics

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 03 Hrs/ Week Tutorials: 00 Hrs/Week Practicals: 00 Hrs/Week	03	ISE: 20 Marks MSE: 30 Marks ESE: 50 Marks

Course Description:

This course plays important role in Data Science. This course provides fundamentals of probability and statistics which required for Data Science. This course focuses on dispersion and measure of central tendency, testing hypothesis, correlation and regression, probability distribution and recurrence relation.

Course Objectives:

❖ To introduce students to understand, explain and apply the fundamental probability and statistical concepts at the core of computer science.
❖ To understand use of concepts of statistics, measures of dispersion.
❖ To learn the fundamental theory of testing hypothesis and sample tests.
❖ To learn the concepts of correlation, regression and curve fitting.
❖ To understand use of recurrence relation.

Course Outcomes:

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

1. Apply the knowledge to study the data given with respect to dispersion and measure of Central tendency.
2. Understand tests for hypothesis and its significance.
3. Describe the statistical data numerically by using correlation and regression.
4. Solve basic problems in probability theory, including problems involving the binomial, Poisson and normal distributions.
5. Apply the recurrence relation to solve the counting problems and program analysis problems.
6. Make use of method of least squares to fit the curves for bivariate data.

Prerequisites:

Basic Probability Theory

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	3	2	-	-	2	-	-	-	-	-	-	-	-	1	3
2	3	2	-	-	-	-	-	-	-	-	-	-	-	1	2

|															
3	3	2	-	-	2	-	-	-	-	-	-	-	-	1	2
4	3	2	-	-	2	-	-	-	-	-	-	-	-	1	3
5	3	2	-	-	2	-	-	-	-	-	-	-	-	1	3
6	3	2	-	-	2	-	-	-	-	-	-	-	-	1	3

Course Contents

Unit 1	Frequency distribution and measure of central Tendency	06 Hours

Frequency distribution, Continuous frequency distribution, Graphical representation of a Frequency distribution- Histogram, frequency polygon, Measure of central tendency- Arithmetic mean, median and mode, Range, Quartile deviation, Mean deviation, Standard deviation

Unit 2	Testing of hypothesis	06 Hours

Introduction, Statistical hypothesis (Simple and Composite), Null hypothesis, Alternative hypothesis, Critical region, Type I and Type II errors, Level of significance, Test for Goodness of fit of chi square distribution and t- distribution.

Unit 3	Correlation and Regression	06 Hours

Introduction, Types of correlation, Karl Pearson's coefficient of correlation, Interpretation of the coefficients of corrections, Computation of coefficient of correlation for ungroup data, Lines of regression , Calculations of equations of the lines of regression

Unit 4	Probability Distribution Functions	06 Hours

Introduction, Elementary theory of probability, Random variables. Discrete probability distribution, Continuous probability distribution, Binomial distribution, Poisson distribution, Normal distribution.

Unit 5	Recurrence Relation	06 Hours

Introduction, Definition of recurrence relation, Linear recurrence relation with constant coefficients, Construction of recurrence relation, Solution of recurrence relation- Homogeneous and non-homogeneous, Solution of homogeneous and non-homogeneous recurrence relation

Unit 6	Curve Fitting	06 Hours

Fitting of curve by method of lest squares, Fitting of straight lines, Fitting of exponential curve, Fitting of second degree parabolic curve

Textbooks:

1. Probability and Statistics for Engineers and Scientists, Walpole, Myers, Myers, Ye, Pearson Education Inc., 8th Edition, 2007, ISBN: 978-81-317-1552-9.
2. Numerical Methods in Engineering and Science, Dr. B.S. Grewal, Khanna Publishers, Delhi.
3. Advanced Engineering Mathematics, H. K. Dass, S.Chand, New Delhi.

Reference Books:

1. Applied statistics and Probability for Engineers, Douglas C Montgomery, George C Runger, Wiley Asia Student Edition, 4th Edition, 2007, ISBN: 978-81-265-2315
2. Statistics for Management, Richard I Levin, David S Rubin, Prentice Hall India, 7th Edition, 1997, ISBN: 9780134762920.
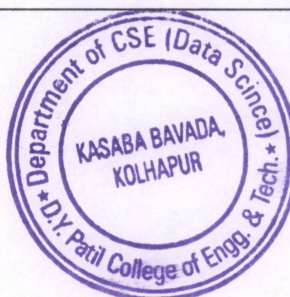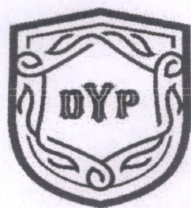
3. Probability and Statistics, Purna Chandra Biswal, PHI Learning Private Limited, Eastern Economy Edition, 2007, ISBN: 978-81-203-3140-2
Useful Links:
1. https://people.iith.ac.in/aravind/Files-DM/LLM-MFCS-2004.pdf

D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – III
Course Code: 231DSPCCL202
Course Name: Data Structures

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 03 Hrs/ Week	03	ISE: 20 Marks
Tutorials: 00 Hrs/Week		MSE: 30 Marks
Practicals: 00 Hrs/Week		ESE: 50 Marks

Course Description:

A course on data structures particularly aims to equip students with a fundamental understanding of various data structures, operations on these data structures, their algorithms, and their applications.

Course Objectives:

❖ To make the students familiar with fundamental concepts of data structures.

❖ To perform the implementation of data structures and conduct various operations on them.

❖ To strengthen the logical thinking ability of students in order to solve real world problems using appropriate data structures.

Course Outcomes:

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

1. Explain the fundamental concepts of data structures.
2. Apply and compare various searching and sorting techniques.
3. Describe various LIFO and FIFO data structures to solve the real world problem.
4. Explain the concept of singly and doubly linked list.
5. Describe the concept of trees, its classification and application.
6. Apply the concept of graph and its application in real world problem solving.
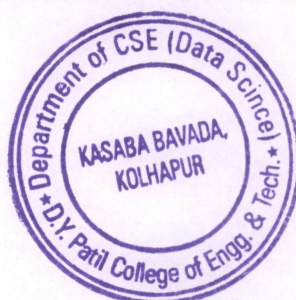
Prerequisites:

Programming Language

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	3	-	-	-	-	-	-	-	-	-	-	-	2	-	2
2	3	3	2	-	-	-	-	-	-	-	-	-	2	-	3
3	3	3	2	-	-	-	-	-	-	-	-	-	2	-	2
4	3	1	1	-	-	-	-	-	-	-	-	-	2	-	2
5	3	1	1	-	-	-	-	-	-	-	-	-	2	-	2

6	3	2	2	-	-	-	-	-	-	-	-	-	2	-	3

Course Contents

Unit 1	Introduction to Data Structures and Algorithms	06 Hours

Basic Terminology, Elementary Data Structure Organization, Classification of Data Structures, Operations on Data Structures, Abstract Data Type, Algorithms, Different Approaches to Designing an Algorithm, Control Structures Used in Algorithms, Time and Space Complexity, Worst-case, Average-case, Best-case, and Amortized Time Complexity, Time–Space Trade-off, Expressing Time and Space Complexity, Algorithm Efficiency, Big O Notation, Omega Notation, Theta Notation.

Unit 2	Searching and Sorting	06 Hours

Introduction to Searching- Linear Search, Binary Search
Introduction to Sorting -Bubble Sort, Insertion Sort, Selection Sort, Merge Sort, Quick Sort, Radix Sort Heap Sort.

Unit 3	Stacks and Queues	06 Hours

Stacks - Introduction to Stacks, Array Representation of Stacks, Operations on a Stack, Applications of Stacks.
Queues - Introduction to Queues, Array Representation of Queues, Types of Queues - Circular Queues, Deques, Priority Queues, Applications of Queues.

Unit 4	Linked List	06 Hours

Introduction, Basic Terminologies, Memory Allocation and De-allocation for a Linked List.
Singly Linked Lists -Traversing a Linked List, searching for a Value in a Linked List, inserting a New Node in a Linked List, Deleting a Node from a Linked List.
Doubly Linked Lists - Inserting a New Node in a Doubly Linked List, deleting a Node from a Circular Doubly Linked List, Applications of Linked Lists.

Unit 5	Trees	06 Hours

Introduction, Basic Terminology, Types of Trees- General Trees, Forests, Binary Trees, Binary Search Tree and Operations, Traversing a Binary Tree -Pre-Order Traversal, In-order Traversal, Post-order Traversal, Hashing and Collision, Heaps- Operations and their applications.

Unit 6	Graphs	06 Hours

Introduction, Graph Terminology, Directed Graphs, Terminology of a Directed Graph, Representation of Graphs - Adjacency Matrix Representation, Graph Traversal Algorithms - Breadth-First Search Algorithm, Depth-first Search Algorithm.

Textbooks:

1. Data Structures Using C, Reema Thareja, Oxford university Press Second Edition. [Unit 1-6]

Reference Books:

1. An introduction to data structures with Applications, Jean-Paul Tremblay, Paul. G. Soresan, Tata Mc-Graw Hill International Editions, 2nd edition.
2. Data Structures- A Pseudo Code Approach with C, Richard F. Gilberg and Behrouz A., 2nd Edition [15 Nov. 2007].
3. Data Structure using C, A. M. Tanenbaum, Y. Langsam, M. J. Augenstein, PHI publication, 2nd Edition, [1996].
4. Schaum's Outlines Data Structures, Seymour Lipschutz, Tata McGraw Hill Education.

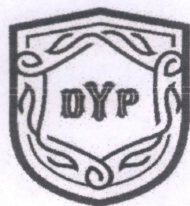| |
|---|
| 5. Data Structures- A Pseudocode Approach with C, Richard F. Gilberg and Behrouz A. Forouzon 2nd Edition. |
| **Useful Links:** |
| 1. https://nptel.ac.in/courses/106102064 |
| 2. https://archive.nptel.ac.in/courses/106/106/106106127/ |

D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – III
Course Code: 231DSPCCP202
Course Name: Data Structures Laboratory

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 00 Hrs/ Week	01	ISE: 25 Marks
Tutorials: 00 Hrs/Week		ESE-POE: 25 Marks
Practicals: 02 Hrs/Week		

Course Description:

A course on data structures particularly aims to equip students with a fundamental understanding of various data structures, operations on these data structures, their algorithms, and their applications.

Course Objectives:

❖ To make the students familiar with fundamental concepts of data structures.

❖ To perform the implementation of data structures and conduct various operations on them.

❖ To strengthen the logical thinking ability of students in order to solve real world problems using appropriate data structures.

Course Outcomes:

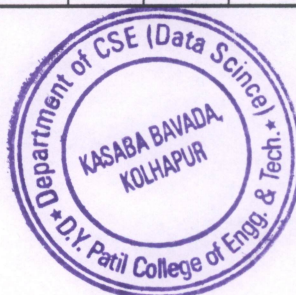On completion of the course, students will be able to –

1. Apply the fundamental concepts of data structures.
2. Apply and compare various searching and sorting techniques.
3. Solve the real-world problems using various LIFO and FIFO data structures.
4. Implement the concept of singly and doubly linked list.
5. Implement the concept of various types of trees.
6. Apply the concept of graph and its application in real world problem solving.

Prerequisites:

Programming Language

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	3	2	1	-	-	-	-	-	1	1	-	1	2	-	3
2	3	3	2	1	-	-	-	-	1	1	-	1	2	-	3
3	3	3	2	-	-	-	-	-	1	1	-	1	2	-	3
4	3	1	1	-	-	-	-	-	1	1	-	1	2	-	3
5	3	1	1	-	-	-	-	-	1	1	-	1	2	-	3

6	3	2	2	1	-	-	-	-	1	1	-	1	2	-	3

Experiment List:

Expt. No.	Experiment	S/O	Hours
1	Write a program to perform different operations on one-dimensional & two-dimensional arrays.	O	2
2	Write a program to perform matrix operations on two-dimensional arrays.	O	2
3	Write a program to implement linear search & binary search.	O	2
4	Write a program to implement bubble sort & insertion sort.	O	2
5	Write a program to implement selection sort.	O	2
6	Write a program to implement merge sort.	O	2
7	Write a program to implement quick sort technique.	O	2
8	Write a program to implement different operations on stack using array.	O	2
9	Write a program to implement different operations on queue using array.	O	2
10	Write a program to implement dequeue and circular queue using array.	O	2
11	Write a program to implement singly linked list.	O	2
12	Write a program to implement doubly linked list.	O	2
13	Write a program to implement binary search tree & demonstrate tree traversal techniques.	O	2
14	Write a program to implement heap & its operations.	O	2
15	Write a program to implement graph.	O	2

D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – III
Course Code: 231DSPCCP203
Course Name: Programming Lab - I

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 02 Hrs/ Week	03	ISE: 50 Marks
Tutorials: 00 Hrs/Week		ESE-POE: 50 Marks
Practicals: 02 Hrs/Week		

Course Description:

This course is about the understanding the fundamentals of object oriented programming & its applications in real world problem solving.

Course Objectives:

❖ To make students to learn difference between procedural and object oriented programming.
❖ To strengthen fundamentals of object oriented programming and the skills required to solve the real world problems.
❖ To make the use of features of object oriented programming to design and solve the complex problems.

Course Outcomes:

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

1. Implement the concept of pointer in procedural language (C language).
2. Explain the fundamentals of C++ programming.
3. Apply the fundamentals of object oriented programming to design the solution for real world problems.
4. Make the use of inheritance to solve the problems.
5. Implement & understand the importance of polymorphism in object oriented programming.
6. Understand & apply the concepts of exception & file handling in real world problem solving.

Prerequisites:

Procedural Programming Language (C Language)

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	3	3	3	-	-	-	-	-	1	1	-	1	2	1	3
2	3	2	2	-	-	-	-	-	1	1	-	1	2	1	2
3	3	3	3	-	-	-	-	-	1	1	-	1	2	1	3
4	3	3	3	-	-	-	-	-	1	1	-	1	2	1	3

5	3	3	3	-	-	-	-	-	1	1	-	1	2	1	3
6	3	3	3	-	-	-	-	-	1	1	-	1	2	1	3

Course Contents

Unit 1	Introduction to Pointers & Structures in C	04 Hours

Pointers: Definition, Pointer Variables, The Pointer Operators, Pointer Expressions, Pointers and Arrays, Array of Pointers, Initializing Pointers, Pointers to Functions and structures, C's Dynamic Allocation Functions.

Structures: Structures, Arrays of Structures, Passing Structures to Functions, Structure Pointers, Unions, Macro expansion.

Unit 2	Fundamentals of C++ Programming	04 Hours

C++ Program Structure, variables, operators, control structure – if, if-else, switch, looping while, do-while, for, C++ keywords. Input/output – I/O streams and standard I/O devices, cin and associated functions, cout and formatted output. User Defined function - declaration, definition & calling function, function call stack and activation records, storage classes, scope rules, function - default arguments. Reference and reference arguments to the function. Pointer variables, new and delete operator, dynamic arrays.

Unit 3	Class & Object	04 Hours

Object Oriented fundamentals, Class and object - concept and need, Class declaration, Class members - member variables and functions, access specifiers, UML notations for class, implementation of member functions. Object Declaration, Accessing class members, class scope, accessor and mutator functions, order of public and private members of the class. Constructors, invoking a constructor, constructors and default parameters, array of objects and constructor, destructor. this pointer, static members, constant objects and member function, Data abstraction, structure and class, information hiding

Unit 4	Inheritance	04 Hours

Inheritance: concept, implementation, base classes and derived classes, members in base classes and derived classes, overriding base class members, UML notations for inheritance, constructors of derived and base classes, destructor in derived class, Inheritance as public, protected and private Composition (Aggregation) and association – concept, implementation and UML Notation

Unit 5	Polymorphism	04 Hours

Polymorphism: Need, concept, implementation using function overloading, Multiple Inheritance, function overriding, virtual function, pure virtual function, abstract classes, Friend function and friend classes, accessing base class functions from derived class objects, accessing derived class functions from base class objects. Operator overloading: fundamentals of operator overloading, overloading binary operators, overloading unary operator

Unit 6	Exceptions Handling and File Processing	04 Hours

Exception handling: Introduction, Handling exceptions within program, C++ mechanism of exception handling, throwing an exception, order of catch blocks, creating exception classes, rethrowing exceptions, exception handling techniques, Standard Library Exception Hierarchy.

File Processing: Introduction, Files and streams, creating and opening a file, file opening modes, Reading data from file, updating file, Random access file – creating and opening a random access file, reading and writing to a random access file, object serialization.

Page | 10
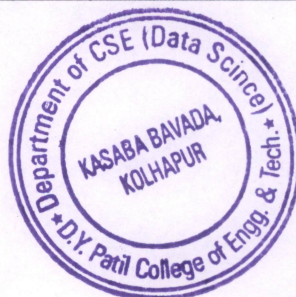
Textbooks:
1. C++ Programming: From Problem Analysis To Program Design, D.S. MALIK, Cengage Learning (5th Edition).
2. C++ How To Program, Paul deitel, Harvey deitel, Pearson Publication 8th Edition.

| Reference Books: |
|---|
| 1. C++ Programming with language, Bjarne Stroustrup, AT & T. |
| 2. Object oriented Programming in C++, R.Lafore, Galgotia Publications 3rd Edition. |
| 3. C++programming, John Thomas Berry, PHI. |
| 4. Object –Oriented Analysis & Design: Understanding System Development with UML 2.0 , Docherty, Wiley India Ltd. |

Useful Links:
1. http://www.spoken-tutorial.org/ NMEICT Project of Govt. Of India
2. C++ API Documentation: https://devdocs.io/cpp/
3. C++ API Reference : https://en.cppreference.com/w/
4. http://www.spoken-tutorial.org/ NMEICT Project of Govt. Of India.

Experiment List:

Expt. No.	Name of Experiment	S/O	Hours
1	Write a program to implement dynamic memory allocation using C.	O	2
2	Write a program to implement Matrix class and its operations. (Addition, Subtraction, Multiplication, Division) (basic C++ program and class)	O	2
3	Write a program to implement BMI (Body Mass Index) calculator. (Class & Object)	O	2
4	Write a program to implement Single & Multilevel Inheritance.	O	2
5	Write a program to implement the Multiple Inheritance.	O	2
6	Write a program to overload insertion and extraction operators for the complex class using friend function. (Friend Function)	O	2
7	Write a program to overload the basic arithmetic operators for the class complex numbers, also modify the show method to include I/O manipulators. (Operator Overloading)	O	2
8	Write a program to calculate mod for a given set of number. Numbers can be integers, floating point numbers and double precision numbers. (Function Overloading)	O	2
9	Write a program to implement the has-a-relationship between given entities. (Composition)	O	2
10	Write a program to implement exception handling.	O	2
11	Write a program to implement file handling.	O	2
12	Students are supposed to develop a mini project using all features of object oriented programming.	O	6
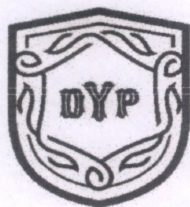
D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – III
Course Code: 231DSMDML201
Course Name: Fundamentals of Data Science

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 02 Hrs/ Week	02	ISE: 20 Marks
Tutorials: 00 Hrs/Week		MSE: --
Practicals: 00 Hrs/Week		ESE: 30 Marks

Course Description:

The aim of the course is to get basic knowledge about data science and its processes. This course also aims to visualize the complex data using different data visualization tools. It also provides different statistical methods to perform data analysis.

Course Objectives:

❖ To provide the students with the basic knowledge of Data Science and its processes.

❖ To make the students to visualize the data using Data visualization tools.

❖ To make the students to understand different statistical methods for data analysis.

Course Outcomes:

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

1. Summarize the basic concepts in data science.
2. Identify the data science process for the problem solving.
3. Choose the appropriate data visualization technique for the given problem.
4. Use different statistical methods for data analysis.
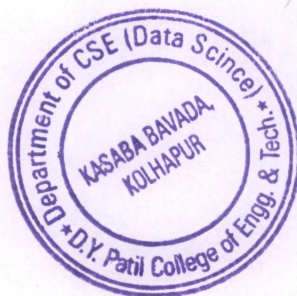
Prerequisites:

Basic knowledge of computer, Basic knowledge of Mathematics

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):
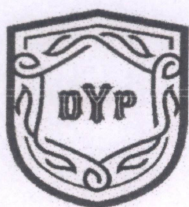
CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	2	-	-	-	-	-	-	-	-	-	-	-	2	-	2
2	2	2	-	-	-	-	-	-	-	-	-	-	2	-	2
3	2	2	-	-	2	-	-	-	-	-	-	-	2	-	3
4	2	2	-	-	-	-	-	-	-	-	-	-	2	-	3

Course Contents		
Unit 1	**Data Science and its scope**	**06 Hours**
What is Data Science, A Brief History, Difference between Data Science and Data Analytics, Knowledge and Skills for Data Science Professionals, Some Technologies used in Data Science, Benefits and uses of Data Science, Facets of Data.		
Unit 2	**Data Science Process**	**06 Hours**
Overview, Defining research goals and creating a project charter, Retrieving data, Cleansing, integrating, and transforming data, Exploratory data analysis, Build the models, Presenting findings and building applications on top of them.		
Unit 3	**Data Visualization**	**06 Hours**
Introduction to data visualization, Visual encoding, Data visualization software, Data visualization libraries, Basic data visualization tools, Specialized data visualization tools, Advanced data visualization tools, Visualization of geospatial data, Data visualization types		
Unit 4	**Statistical Data Analysis**	**06 Hours**
Role of statistics in data science, Kinds of statistics - Descriptive statistics, Inferential statistics, Probability theory - Random variables, Independence, Sample space, Odds and risks, Expected values, Standard errors, Bayesian probability, Probability distribution		

Textbooks:

1. Introducing Data Science, Davy Cielen, Arno D. B. Meysman, Mohamed Ali, Manning Publications. [Unit 1 & 2]
2. Data Science Fundamentals and Practical Approaches, DR. Gypsi Nandi, DR. Rupam Kumar Sharma, BPB Publications, India , ISBN 978-93-89845-662 [Unit 3 & 4]

Reference Books:

1. Essentials of Data Science and Analytics, DR. Amar Sahay, O'REILLY Publication.
2. Fundamentals of Data Science, Sanjeev Wagh, Manisha Bhende & Anuradha Thakare, CRC Press.

Useful Links:

1. **https://onlinecourses.nptel.ac.in/noc21_cs23/preview**

D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – III
Course Code: 231DSOECL201
Course Name: Data Science for Everyone

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 03 Hrs/ Week	03	ISE: 20 Marks
Tutorials: 00 Hrs/Week		MSE: 30 Marks
Practicals: 00 Hrs/Week		ESE: 50 Marks

Course Description:

Data science for Everyone encompasses statistics, machine learning, data visualization, and data engineering, among other areas. The basic aim of this course is to understand foundational concepts of data science and practical skills to analyze and interpret data effectively.

Course Objectives:

❖ To provide strong foundation for data science and application area related to it.

❖ To understand the underlying core concepts and emerging technologies in data science.

Course Outcomes:

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

1. Understand the fundamental principles of data science and introduction to modern tools used in the era of data science.
2. Describe the basic data manipulation and analysis techniques.
3. Explain the importance of data science tools and programming languages.
4. Explain the use of probability and statistics in the field of data science.
5. Understand basic machine learning algorithms and its use in real-world problem solving.

Prerequisites:

Mathematics, MS-Excel

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	3	-	-	-	1	-	-	-	-	-	-	-	3	-	2
2	2	-	-	-	1	-	-	-	-	-	-	-	2	-	2
3	3	1	1	-	2	-	-	-	-	-	-	-	2	-	2
4	3	2	-	-	-	-	-	-	-	-	-	-	2	2	2
5	2	2	-	1	-	-	-	-	-	-	-	-	2	2	2

Course Contents

Unit 1	Introduction to Data Science	06 Hours

Unit 1	Introduction to Data Science	06 Hours
Overview of Data Science- Definition and significance, Applications in various industries, The Data Science Process, Steps in a data science project: data collection, cleaning, analysis, and interpretation, Tools and Technologies- Introduction to tools like Python, R, Jupyter Notebook.		
Unit 2	**Data Collection and Data Wrangling**	**06 Hours**
Types of Data- Structured vs. unstructured data, Databases and data sources, Data Collection Methods, APIs, web scraping, databases, Data Cleaning, Handling missing values, Data formatting and transformation.		
Unit 3	**Exploratory Data Analysis (EDA)**	**06 Hours**
Descriptive Statistics- Mean, median, mode, variance, standard deviation, Data Visualization, Introduction to Matplotlib, Seaborn, and Plotly, creating basic plots: histograms, bar charts, scatter plots, Identifying Patterns and Outliers.		
Unit 4	**Introduction to Probability and Statistics**	**06 Hours**
Basic Probability Concepts- Probability distributions (normal, binomial, etc.), Statistical Inference- Hypothesis testing, Confidence intervals.		
Unit 5	**Introduction to Machine Learning**	**06 Hours**
Supervised vs. Unsupervised Learning, Basic Algorithms- Linear Regression, Classification (e.g., Decision Trees, k-NN), Clustering (e.g., K-means).		
Unit 6	**Case Studies**	**06 Hours**
Tableau, Microsoft Power BI Case Study 1: Urban planning and smart cities Case Study 2: E-commerce personalization and recommendation systems		

Textbooks:

1. Doing Data Science, Cathy O'Neil and Rachel Schutt, O'Reilly Media, 1st Edition
2. The Elements of Statistical Learning, Trevor Hastie, Robert Tibshirani, and Jerome Friedman, Springer 2nd Edition.
3. Data Science from Scratch: First Principles with Python, Joel Grus, O'Reilly, 1st edition, 2015.

Reference Books:

1. Doing Data Science, Straight Talk from the Frontline, Cathy O'Neil, Rachel Schutt, O' Reilly, 1st edition, 2013.
2. Mining of Massive Datasets, Jure Leskovec, Anand Rajaraman, Jeffrey David Ullman, Cambridge University Press, 2nd edition, 2014.
3. Pattern Recognition and Machine Learning, Christopher M. Bishop, Springer Publication.

Useful Links:

1. https://www.youtube.com/watch?v=XohgKT13FKY&list=PLqICp9VkFcbEWeZ0Q-6gsHCRaqe5eyf
2. https://www.youtube.com/watch?v=fn1rKKNLuzk&list=PL15FRvx6P0OWTlNBS_93NHG2hIn9cy

D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – III
Course Code: 231DSOECL201
Course Name: Data Science for Everyone Laboratory

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 00 Hrs/ Week Tutorials: 00 Hrs/Week Practicals: 02 Hrs/Week	01	ISE: 25 Marks

Course Description:

Data science for Everyone encompasses statistics, machine learning, data visualization, and data engineering, among other areas. The basic aim of this course is to understand foundational concepts of data science and practical skills to analyze and interpret data effectively.

Course Objectives:

❖ To provide strong foundation for data science and application area related to it.
❖ To understand the underlying core concepts and emerging technologies in data science.

Course Outcomes:

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

1. Understand the fundamental principles of data science and introduction to modern tools used in the era of data science.
2. Apply the basic data manipulation and analysis techniques.
3. Make use of data science tools and programming languages effectively.
4. Implement the fundamental concepts of probability and statistics in the field of data science.
5. Apply machine learning algorithms to solve real-world problems.

Prerequisites:

Mathematics, MS-Excel

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	3	-	-	-	1	-	-	-	-	-	-	-	3	-	2
2	2	2	2	-	2	-	-	-	-	-	-	-	2	-	3
3	3	2	2	-	2	-	-	-	-	-	-	-	2	-	3
4	3	2	1	1	-	-	-	-	-	-	-	-	2	2	3
5	3	2	2	1	2	-	-	-	-	-	-	-	2	2	3

Experiment List:

Expt. No.	Experiment	S/O	Hours
1	Installation of R studio & environment configuration.	S	2
2	Write a program to define variables and datatypes of R.	O	2
3	Write a program to implement loading and exploring dataset.	O	2
4	Write a program to perform cleaning and pre-processing of data.	O	2
5.	Write a program to perform data visualization using different types of techniques.	O	2
6	Write a program to implement simple linear regression.	O	2
7	To study the structure and main characteristics of a dataset.	S	2
8	Write a program to find subset of dataset by using subset (), aggregate () functions on iris dataset	O	2
9	To study the Tableau tool	S	2
10	To study the PowerBI tool	S	2

D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – III
Course Code: 231DSFPP201
Course Name: Domain Specific Mini Project

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 00 Hrs/ Week	02	ISE: 25 Marks
Tutorials: 00 Hrs/Week		ESE-OE: 25 Marks
Practicals: 04 Hrs/Week		

Course Description:

This course emphasis on a problem-based learning approach. It is a group activity where students have to present an idea / solution for the problem chosen. Then requirement analysis and design specification of the system is to be developed by the students. This is followed by software design, implementation, testing and finally demonstrate the results obtained. This course helps the students to learn how to analyze the demands of a customer and represent them in the form of software requirements specification (SRS) document including quality requirements. Ultimately this course enhances students programming skills and enable them to learn how to perform requirement analysis, system designing, coding, testing and report writing.

Course Objectives:

- ❖ To formulate the problem statement.
- ❖ To follow the SDLC model for development of project.
- ❖ To enhance team work including skills of team building and intra personal communication.
- ❖ To develop the logical skills and use of appropriate data structures for solving the engineering problems.

Course Outcomes:

On completion of the course, students will be able to –
1. Define appropriate problem statement for real world problems.
2. Organize an effective project plan with clear objectives and prepare a synopsis.
3. Design the various modules of the project to provide a solution to the problem with the help of various design tools.
4. Develop the proposed system using suitable development platform.
5. Able to present their work and prepare technical project report.

Prerequisites:

Mathematics, Data Structures, Software Engineering and knowledge of Programming language.

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	3	3	-	1	-	1	1	1	3	3	2	2	2	-	3

| 2	3	3	-	1	1	1	1	1	3	3	2	2	2	-	3
3	3	3	2	1	1	1	1	1	3	3	2	2	2	1	3
4	3	3	2	1	2	1	1	1	3	3	2	2	2	2	3
5	-	-	-	-	2	-	-	2	3	3	2	2	2	-	3

Course Contents

The Project should be undertaken preferably by a group of 3-4 students who will jointly work and implement the project. The group will select a project with the approval from the domain expert panel and submit the name of the project with a synopsis. The Project should consist of defining the problem and analyzing it, designing the solution and implementing it using a suitable programming language. A presentation and demonstration based on the above work is to be given by the group for ISE. The work will be jointly assessed twice in a semester by an internal domain expert panel. A hard copy of project report of the work done is to be submitted along with the softcopy of the project during ESE.

Project topics may be selected from following domains:

1) Real world applications in Data Analytics
2) Probability and Statistics
3) Data Preprocessing
4) Web Page design
5) Web Scrapping
6) Healthcare Analytics
7) Analytics using modern tools & techniques.
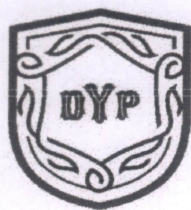
D. Y. Patil College of Engineering & Technology, Kolhapur

(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – III
Course Code: 231DSVECL201
Course Name: Environmental Studies

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 02 Hrs/ Week Tutorials: 00 Hrs/Week Practicals: 00 Hrs/Week	02	ISE: -- MSE: -- ESE: 50 Marks

Course Description:

The main objective of course is to create awareness among students regarding environmental issues and its impact on society. Knowledge regarding environmental components, its degradation and protection of environment is need for sustainable future ahead.

Course Objectives:

❖ Understand the scope and importance of Environmental Studies and sustainable development.
❖ Understand connection between environmental health and developmental activities.
❖ Understand the importance of Environmental Management for its protection through technical and legislative point of view.
❖ Acquire problem solving skills through visits to different locations, identifying the Environmental problems and proposing solution for societal benefits.

Course Outcomes:
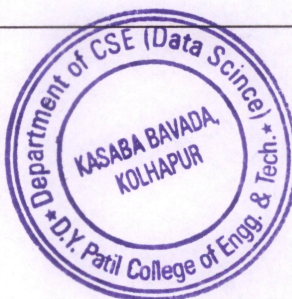
On completion of the course, students will be able to –

1. Understand the scope and importance of Environmental awareness and Sustainable development
2. Understand various Environmental issues due to development
3. Understand various modes of Environmental management through technology and legislation
4. Acquire problem solving attitude through actual field experience, reporting it in the form of Field project work.
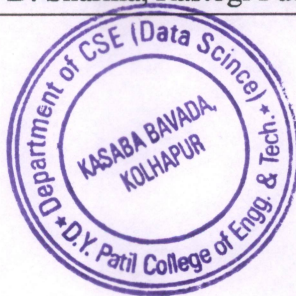
Prerequisites:

Understanding of Environmental Education course

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

| CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	-	-	-	-	-	1	3	2	-	-	-	2	-	-	2
2	-	-	-	-	-	1	3	1	-	-	-	2	-	-	2
3	-	-	-	-	-	1	3	1	-	-	-	2	-	-	2
4	-	-	-	-	-	2	3	1	-	1	1	2	-	-	3

Course Contents

Unit 1	Our Environment	04 Hours

Introduction to Environment, Scope of Environmental studies, importance of environmental awareness.
Concept of sustainability, Sustainable Development- history and Goals, environmental ethics, Sustainability ethics.
Population growth of world and reduced health content of the environment.

Unit 2	Development and Environmental Health	10 Hours

Natural resources - Types (renewable and non-renewable), developmental benefits, Forest- Benefits, problems (Deforestation), Biodiversity - importance, threats, conservation, Ecosystems- importance, problem associated with major ecosystems, ecological restoration, Air- Benefits, problems (Pollution, climate change), Water- Benefits, problems (Depletion, pollution). Soil/ Land- Benefits, problems (Degradation, loss of fertility, desertification), Mineral- Benefits, problems (Mining, over exploitation, depletion, pollution), Energy resources- Benefits, problems (depletion, energy crisis)

Urbanization and Environmental health - Urban problems, Solid waste- Effects of MSW, Plastic waste, Hazardous waste, E- waste

Unit 3	Environmental Management	10 Hours

Renewable energy technologies - current, new (Bio gas, Bio fuel, hydrogen, etc.), Pollution abetment – 5R, ZLD, carbon credit, bio remedies, Soil/ land reclamation, Sustainable agriculture, Concept of EIA, Environmental audit, ISO certification (ISO 14001), Role of CPCB and MPCB in Environmental protection of India.

Emerging technologies for environmental management- GIS, Remote sensing, Smart bin, IoT integration, Waste-to-Energy Technologies, Recycling Automation, Advanced Data Analytics, Circular Economy Practices, Sustainable Packaging Solutions, Community Engagement and Education, Decentralized Waste Treatment, Zero-Waste Initiatives, Legislative and Regulatory Changes.

Environmental legislation - Environmental Protection Act, Air Act, Water Act, Solid waste Management Act, Hazardous waste Management Rule, E- Waste (Management) Rules, 2022

Unit 4	Field Project Work	--

Case studies based on site visit (Each candidate has to go for field visit and complete a project work on Environmental issues and probable solutions)

Textbooks:

1. Handbook of Environmental Studies, Dr. G. R. Parihar, Satyam Publishers and Distributors (1 January 2013), ISBN-10 : 9382664408, ISBN-13: 978-9382664406.
2. Environmental Studies, Anubha Kaushik, New Age International Private Limited (1 January 2007), ISBN-10: 8122422403, ISBN-13: 978-8122422405.
3. Introduction to Environmental Engineering and Science, Masters, Pearson Education India; 3rd edition (1 January 2015), ISBN-10: 9332549761, ISBN-13: 978-9332549760.
4. Solid Waste Management in developing countries, Bhide A. D. and Sundersen B. B., Indian National Scientific Documentation Centre, New Delhi.

Reference Books:

1. Handbook of Environmental Laws, Rules, Guidelines, Compliances and Standards, Trivedi R.K., Vol. I &II, Environmental Media.
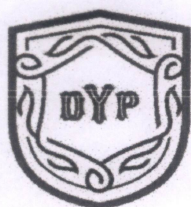2. Ecology And Environment Pb, P. D. Sharma, Rastogi Publications (1 January 2011).

Useful Links:

1. Environment English Book 1-3-2022 Final Corrected copy_compressed.pdf
2. Manual on Municipal Solid Waste Management- Ministry of Urban Development, Govt. of India
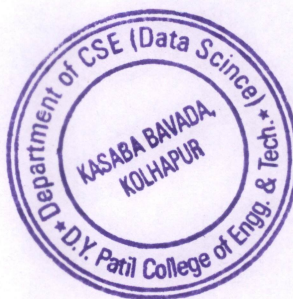
	D. Y. Patil College of Engineering & Technology, Kolhapur
![DYP logo]	(An Autonomous Institute)
	Second Year B. Tech CSE (Data Science)
	Semester – III
	Course Code: 231DSHSSML201
	Course Name: Economics and Management for IT

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 02 Hrs/ Week	02	ISE: 20 Marks
Tutorials: 00 Hrs/Week		MSE: 30 Marks
Practicals: 00 Hrs/Week		ESE: --

Course Description:

The course is intended to provide basic understanding of Economics and Management to engineering students with following aspects –

❖ To impart knowledge, with respect to concepts of management information system.
❖ To expose the students to the characteristic and applications of Decision Support Systems.
❖ To help the students to understand different trends in current information system technology and also IT Tools & Techniques for Business operations.

Course Objectives:

❖ To get the overview of system development management life cycle.
❖ To understand scope and objective of management information system.
❖ To enhance management decision making.
❖ To make the engineering students aware about different corporate case studies.

Course Outcomes:

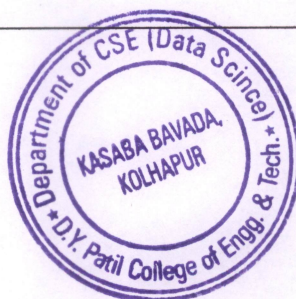On completion of the course, students will be able to –

1. Explain the concepts of system development management life cycle.
2. Describe scope and objective of management information system.
3. Develop the decision making skills and practices.
4. Elaborate the different corporate case studies.

Prerequisites:

Basics knowledge of Computer

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	2	-	-	-	-	-	-	-	-	-	2	1	1	-	2
2	2	-	-	-	-	-	-	-	-	-	2	1	1	-	2
3	2	-	-	-	-	-	-	-	-	-	2	1	1	-	2
4	2	-	-	-	-	-	-	-	-	-	1	1	1	-	2

Course Contents		
Unit 1	**Management Information System**	**06 Hours**
Conceptual foundations of information systems; Information theory; Information resource management; Types of information systems; Systems development - Overview of systems and design; System development management life-cycle, designing for online and distributed environments; Implementation and control of project.		
Unit 2	**Scope and Objectives of MIS**	**06 Hours**
MIS meaning and role, MIS concepts, Management science structure, Information flow in management, MIS for management support, Planning with MIS, control with MIS. Problem solving & decision making, Development of MIS, strategic & project planning for MIS.		
Unit 3	**Enhancing Management Decision Making**	**06 Hours**
Decision support systems (DSS) – understanding DSS, characteristics components, major DSS applications. Group decision support systems (GDSS), - elements, characteristics, how GDSS can enhance group decision - making? Executive support systems (ESS) – role of ESS in the organization, developing ESS, benefits of ESS.		
Unit 4	**Case Studies**	**06 Hours**
Web Publishing: types of websites, Web surfing, E- commerce, B2B, B2C, C2C, E-commerce security issues, Ethical issues.		

Textbooks:

1. Management of Information systems, Gordon B. Davis & Margreth H. Olson, Pearson Edition.

Reference Books:

1. MIS Concepts & Design, Robert C. Murdik, PHI 2nd Edition
2. Information system, H.F. & Abraham, S.
3. Engineering Economics, R.Paneerselvam, PHI publication
4. Modern Economic Theory, Dr. K. K. Dewett & M. H. Navalur, S. Chand Publications

Useful Links:

1. https://onlinecourses.swayam2.ac.in/cec21_ge05/preview
2. https://www.coursera.org/courses?query=management%20information%20systems
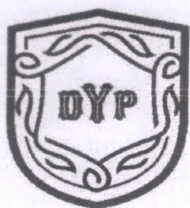
	D. Y. Patil College of Engineering & Technology, Kolhapur	
	(An Autonomous Institute)	
	Second Year B. Tech CSE (Data Science)	
	Semester – III	
	Course Code: 231DSCCA201	
	Course Name: Liberal Learning	

Teaching Scheme:	**Audit Course**	**Evaluation Scheme:**
Lectures: 00 Hrs/ Week		Grade Based Evaluation
Tutorials: 00 Hrs/Week		
Practicals: 02 Hrs/Week		

List of Clubs Available at Department:

1. Data Analytics Club:

Aim:

The primary aim of a Data Analytics Club is to create a collaborative and dynamic environment where students can enhance their knowledge, skills, and practical experience in data analytics. This involves fostering a community that promotes learning, innovation, and professional growth in the field of data analytics. The club seeks to bridge the gap between academic knowledge and real-world application, preparing members for successful careers in data analytics by providing resources, networking opportunities, and hands-on experiences.

Objectives:

- **Skill Development:**
  - Technical Skills: Provide training and workshops on data analytics tools and technologies such as Python, R, SQL, Tableau, Power BI, and machine learning.
  - Soft Skills: Enhance communication, teamwork, problem-solving, and project management skills through collaborative projects and presentations.
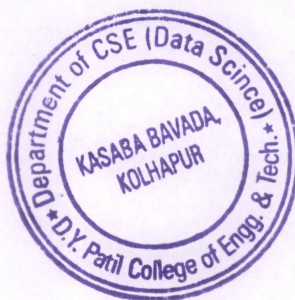- **Knowledge Expansion:**
  - Guest Lectures: Invite industry professionals to speak about current trends, best practices, and real-world applications of data analytics.
  - Industry Trends: Keep members informed about the latest developments in data analytics, including new tools, techniques, and methodologies.
- **Hands-on Experience:**
  - Projects: Offer opportunities to work on real-world data projects, either individually or in teams, to apply theoretical knowledge in practical scenarios.
  - Competitions: Organize and participate in data analytics competitions and hackathons to foster a competitive and innovative spirit.
- **Networking:**
  - Industry Connections: Facilitate connections with professionals and alumni working in the field of data analytics.
  - Peer Networking: Create a community where students can collaborate, share knowledge, and support each other's learning journeys.

- **Career Preparation:**
  - Internship Opportunities: Provide information about internships and job openings in the field of data analytics.
  - Resume Building: Offer workshops on resume writing, LinkedIn profile optimization, and interview preparation specific to data analytics roles.
- **Research & Innovation:**
  - Research Projects: Encourage and support members in conducting research projects and publishing their findings in academic or industry journals.
  - Innovation: Foster a culture of innovation by encouraging members to explore new ideas and approaches in data analytics.

Outcomes:
- **Enhanced Skill Set:** Members will develop a strong foundation in data analytics tools and techniques, making them more competitive in the job market.
- **Practical Experience:** Participation in real-world projects and competitions will provide practical experience, helping members to apply theoretical knowledge in practical scenarios.
- **Professional Growth:** Networking opportunities with industry professionals and alumni will help members build valuable connections and gain insights into the industry.
- **Career Readiness:** Members will be better prepared for careers in data analytics through exposure to job opportunities, resume workshops, and interview preparation sessions.
- **Research Contributions:** Members will have opportunities to contribute to the field through research projects and publications.
- **Community Building:** The club will create a supportive community where students can share knowledge, collaborate on projects, and help each other grow.
- **Innovation and Creativity:** Members will be encouraged to think creatively and innovate, leading to new ideas and approaches in data analytics.
- **Increased Confidence:** Through presentations, workshops, and networking events, members will gain confidence in their abilities to communicate and apply data analytics concepts.

2. Open Source Software Developers Club:

Aim:

The primary aim of an Open Source Student Club is to foster a community of students who are passionate about open source software and collaboration. The club seeks to promote the use, development, and contribution to open source projects, encouraging members to learn, innovate, and share their knowledge with the wider community. By doing so, the club aims to empower students with the skills and mindset needed to contribute meaningfully to the open source ecosystem and to prepare them for careers in technology and software development.

Objectives:
- **Education & Skill Development:**
  - Technical Workshops: Provide training on various open source technologies, programming languages, and development tools.

- Best Practices: Teach best practices for contributing to open source projects, including version control, code reviews, and collaboration techniques.
- **Project Involvement:**
  - Contributions: Encourage and guide members to contribute to existing open source projects.
  - Initiate Projects: Support members in starting and maintaining their own open source projects.
- **Community Engagement:**
  - Collaboration: Foster a collaborative environment where members can work together on projects and share knowledge.
    Outreach: Engage with the wider open source community through events, meetups, and online platforms.
- **Networking:**
  - Industry Connections: Facilitate connections with professionals, open source contributors, and organizations in the open source community.
  - Peer Networking: Create opportunities for members to network with each other and build lasting professional relationships.
- **Career Preparation:**
  - Internship and Job Opportunities: Provide information about internships, job openings, and career paths in open source development.
    Professional Development: Offer workshops on resume building, portfolio creation, and interview preparation specific to open source careers.

Outcomes:
- **Enhanced Technical Skills:** Members will develop proficiency in open source tools and technologies, improving their coding and development skills.
- **Practical Experience:** Participation in real-world open source projects will provide hands-on experience and a deeper understanding of software development.
- **Community Contribution:** Members will make meaningful contributions to open source projects, helping to advance the open source ecosystem.
- **Professional Growth:** Networking opportunities with industry professionals and active contributors will help members build valuable connections and gain insights into the industry.
- **Career Readiness:** Members will be better prepared for careers in open source development through exposure to job opportunities, portfolio building, and interview preparation sessions.
- **Innovation & Creativity:** Members will be encouraged to innovate and explore new ideas, leading to the development of new open source projects and solutions.
- **Increased Confidence:** Through workshops, presentations, and collaboration, members will gain confidence in their abilities to contribute to open source projects and communities.
- **Community Building:** The club will create a supportive and collaborative community where students can share knowledge, work on projects together, and help each other grow.

3. Rational Programmers Club:

Aim:

The primary aim of a Rational Programmers Club within the Data Science Department is to cultivate a community where students can enhance their coding skills, apply their knowledge to solve real-world problems, and prepare for careers in data science and related fields. The club seeks to create an environment that encourages continuous learning, collaboration, and innovation in coding, particularly as it applies to data science.

Objectives:

- **Skill Development:**
  - Programming Efficiency: Provide training in various programming languages commonly used in data science, such as Python, R, SQL, and Java.
  - Data Science Tools: Teach members how to use data science tools and libraries, such as pandas, NumPy, scikit-learn, TensorFlow, and more.
- **Practical Applications:**
  - Projects: Encourage members to participate in or initiate coding projects that solve real-world data science problems.
  - Hackathons: Organize and participate in hackathons and coding competitions to foster a spirit of innovation and problem-solving.
- **Knowledge Sharing:**
  - Workshops & Seminars: Conduct workshops, seminars, and coding boot camps to share knowledge and best practices in coding and data science.
  - Peer Learning: Promote peer-to-peer learning through study groups, code reviews, and collaborative projects.
- **Career Preparation:**
  - Portfolio Building: Help members build a portfolio of coding projects that demonstrate their skills and knowledge in data science.
  - Career Guidance: Provide guidance on career paths in data science, including resume building, interview preparation, and internship/job search strategies.
- **Networking:**
  - Industry Connections: Facilitate connections with data science professionals, alumni, and industry partners.
  - Community Building: Create a supportive community where members can network, collaborate, and share resources.
- **Innovation & Research:**
  - Cutting-edge Topics: Explore and work on cutting-edge topics in data science, such as machine learning, artificial intelligence, big data analytics, and more.
  Research Projects: Encourage members to participate in or initiate research projects and publish their findings.

Outcomes:

- **Enhanced Coding Skills:** Members will develop strong programming skills and a deep understanding of data science tools and techniques.
- **Practical Experience:** Participation in real-world projects and hackathons will provide hands-on experience, enabling members to apply their knowledge practically.
- **Professional Growth:** Networking with industry professionals and peers will help members build valuable connections and gain insights into the data science industry.
- **Career Readiness:** Members will be better prepared for careers in data science through portfolio building, resume workshops, and interview preparation sessions.
- **Community Contributions:** The club will create a collaborative community where members can share knowledge, support each other's learning, and contribute to the field of data science.
- **Innovation & Research Contributions:** Members will have opportunities to work on innovative projects and research, contributing to advancements in data science.
- **Increased Confidence:** Through coding challenges, presentations, and collaborative projects, members will gain confidence in their abilities to code and solve complex data science problems.
- **Lifelong Learning:** The club will instill a mindset of continuous learning and curiosity, encouraging members to stay updated with the latest developments in coding and data science.

Evaluation:

Evaluation of individual student will be carried out based on following criteria -

1. Knowledge & Understanding
2. Critical Thinking
3. Communication Skills
4. Ethical Reasoning
5. Cultural Awareness and Diversity
6. Interdisciplinary Connections
7. Creativity and Innovation
8. Collaboration and Teamwork

Program Coordinator   HOD CSE(Data Science)   Dean Academics   Principal

Semester – IV

D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – IV
Course Code: 231DSPCCL204
Course Name: Discrete Mathematical Structure

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 03 Hrs/ Week	03	ISE: 20 Marks
Tutorials: 00 Hrs/Week		MSE: 30 Marks
Practicals: 00 Hrs/Week		ESE: 50 Marks

Course Description:

This course is about to introduce topics in Discrete Mathematics relevant to Data Analysis. The objective is to teach students how to think logically and mathematically. This Course consists of concepts of discrete mathematical structures such as Set Theory, Algebraic systems, Lattices.

Course Objectives:

❖ Apply basic concepts of set theory, logic, proof techniques, graphs and trees.
❖ Analyze the basic concepts of relations and functions.
❖ Learn the concepts of algebraic system & groups.

Course Outcomes:

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

1. Apply the logical reasoning to solve the problems.
2. Solve the problems based on binary relations.
3. Describe the different terminologies in algebraic structures.
4. Evaluate the problems based on Boolean functions.
5. Solve different graph problems like PERT graph, tree traversal.
6. Solve different problems on counting principle, permutation & combinations.

Prerequisites:

Basic Knowledge of Mathematics

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	3	2	-	-	-	-	-	-	-	-	-	-	1	-	3
2	3	2	-	-	-	-	-	-	-	-	-	-	1	-	3
3	2	-	-	-	-	-	-	-	-	-	-	-	1	-	2
4	3	2	-	-	-	-	-	-	-	-	-	-	1	-	3
5	3	2	-	-	-	-	-	-	-	-	-	-	1	-	3
6	3	2	-	-	-	-	-	-	-	-	-	-	1	-	3

Course Contents		
Unit 1	**Mathematical Logic**	**08 Hours**
Introduction, statements and notations, connectives, Statement formulas and truth tables, well-formed formulas, Tautologies, Equivalence of formulas, Duality law, Tautological implications, other connectives, Normal Forms, Theory of Inference for statement calculus – validity using truth table, rules of inference.		
Unit 2	**Set Theory & Binary Relations**	**07 Hours**
Basic concepts of set theory, types of operations on sets, ordered pairs, Cartesian Product, relation, properties of binary relations, matrix and graph representation, partition and covering of set, equivalence relation, composition, POSET and Hasse diagram, Function – types, composition of functions.		
Unit 3	**Algebraic Structures**	**04 Hours**
Algebraic Systems, Semigroups and Monoids, Homomorphism, Groups: Definition and examples, subgroups and homomorphism.		
Unit 4	**Lattices & Boolean Algebra**	**06 Hours**
Lattice as POSETs, definition, examples and properties, Lattice as algebraic systems, Special lattices, Boolean algebra definition and examples, Boolean functions, representation and minimization of Boolean functions.		
Unit 5	**Graph Theory**	**06 Hours**
Basic concepts of graph theory, Complete, Regular and Bipartite Graphs, Graph Coloring, Storage representation and manipulation of Graphs, PERT and related techniques.		
Unit 6	**Permutation & Combination**	**05 Hours**
Basics of Counting, Counting Principle, Rule of Sum & Product, Pigeonhole Principle, Permutations and Combinations, Inclusion-Exclusion Principle, Discrete Probability Theory.		
Textbooks:		
1. Discrete Mathematical Structures with Application to Computer Science, J. P. Tremblay & R. Manohar, MGH International. [Unit 1-5] 2. Elements of Discrete Mathematics, C. L. Liu and D. P. Mohapatra, Tata McGraw-Hill. [Unit 6]		
Reference Books:		
1. Discrete Mathematics and its Applications, Kenneth H. Rosen, AT&T Bell Labs. 2. Discrete Mathematics, Semyour Lipschutz, MarcLipson, MGH, Schaum's outlines.		
Useful Links:		
1. https://nptel.ac.in/courses/106106183 2. https://nptel.ac.in/courses/106106094 3. https://nptel.ac.in/courses/111107058 4. https://www.geeksforgeeks.org/mathematics-combinatorics-basics		
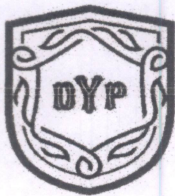
D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – IV
Course Code: 231DSPCCL205
Course Name: Design and Analysis of Algorithm

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 03 Hrs/ Week	03	ISE: 20 Marks
Tutorials: 00 Hrs/Week		MSE: 30 Marks
Practicals: 00 Hrs/Week		ESE: 50 Marks

Course Description:

This course introduces basic methods for the design and analysis of efficient algorithms. Different algorithms for a given computational task are presented and their relative merits evaluated based on performance measures. It introduces the fundamental techniques for designing and analyzing algorithms, including asymptotic analysis, divide-and-conquer algorithms, greedy algorithms, dynamic programming, traversal methods and even backtracking approach. It also provides introduction to NP-completeness.

Course Objectives:

 ❖ To introduce algorithm design methods / techniques with analysis.
 ❖ To devise algorithm for given problem statement and compute its complexity.
 ❖ To introduce complex computational problems.

Course Outcomes:

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

1. Devise algorithm and analyze its space and time complexity by using recurrence relation.
2. Apply divide & conquer techniques to solve real time problems.
3. Solve & analyze problems using the greedy techniques.
4. Apply appropriate dynamic-programming algorithm to solve real world problem.
5. Solve the problems using the principles of backtracking.
6. Describe the concepts of NP hard and NP complete.

Prerequisites:

 Data structures

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	3	3	3	1	-	-	-	-	-	-	-	-	1	2	6
2	3	2	2	2	-	-	-	-	-	-	-	2	1	2	3
3	3	2	2	2	-	-	-	-	-	-	-	2	1	2	3
4	3	2	2	2	-	-	-	-	-	-	-	2	1	2	3
5	3	2	2	2	-	-	-	-	-	-	-	2	1	2	3
6	2	2	-	-	-	-	-	-	-	-	-	-	1	-	2

Course Contents

Unit 1	Introduction to Algorithms	06 Hours
Definition, Properties of Algorithms, Expressing Algorithm, Flowchart, Algorithm Design Techniques, Performance Analysis of Algorithms, Types of Algorithm's Analysis, Order of Growth, Asymptotic Notations, Recursion, Recurrences Relation		
Unit 2	Divide and Conquer	06 Hours
The general method, Binary search, Finding the maximum and minimum, Merge sort, Quick sort, Strassen's Matrix Multiplication, analysis of Divide and Conquer algorithms.		
Unit 3	Greedy Algorithms	06 Hours
Introduction to Greedy Technique, Greedy Method, Optimal Merge Patterns, Huffman Coding, Knapsack Problem, Activity Selection Problem, Job Sequencing with Deadline, Minimum Spanning Tree, Prim's Algorithm, Kruskal's algorithm, Single-Source Shortest Path Algorithm		
Unit 4	Dynamic Programming	07 Hours
The general method, Comparison of Divide-and-Conquer and Dynamic Programming Techniques, Longest Common Sub-sequence, matrix multiplication, Bellman Ford, Floyd Warshall, Multistage graphs, All pair shortest paths, 0/1 knapsack, Traveling Sales person problem.		
Unit 5	Backtracking	06 Hours
Backtracking Concept, N–Queens Problem, Four–Queens Problem, Eight–Queen Problem, Hamiltonian Cycle, Sum of Subsets Problem, Graph Coloring Problem, Branch and Bound: Introduction, Travelling Salesperson Problem, 15-Puzzle Problem, Comparisons between Backtracking and Branch and Bound.		
Unit 6	NP Hard and NP Complete Problems	05 Hours
Introduction, the Complexity Class P, the Complexity Class NP, Polynomial-Time Reduction, the Complexity Class NP-Complete.		

Textbooks:

1. Fundamentals of Computer Algorithms, Ellis Horowitz, Satraj Sahani, Saguthevar Rajasejaran, Universities Press, Second Edition.[Unit 1-6]

Reference Books:

1. Fundamentals of Algorithmics, Gilles Brassard, Paul Bratley, Pearson Education.
2. Mastering Algorithms with C, Kyle Loudon, SPD O'Reilly.
3. Computer Algorithms- Introduction to Design and Analysis, Allen Van Gelder , Sara Baase, Pearson Education.

Useful Links:

1. http://personal.kent.edu/~rmuhamma/Algorithms/algorithm.html
2. https://www.ics.uci.edu/~goodrich/teach/cs260P/notes/

D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – IV
Course Code: 231DSPCCP206
Course Name: Programming Lab - II

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 02 Hrs/ Week	04	ISE: 50 Marks
Tutorials: 00 Hrs/Week		ESE-POE: 75 Marks
Practicals: 04 Hrs/Week		

Course Description:

Python is a high-level programming language which is used for developing a wide range of applications in different domains. It has a strong community around machine learning, data modeling, data analysis and artificial intelligence (AI), with extensive resources and libraries built for the same purpose.

Course Objectives:

❖ To make the students to learn basics of python programming language.

❖ To expose the students to various data structures.

❖ To make the students aware of various Object Oriented concepts.

❖ To expose the students to advanced concepts in Python.

Course Outcomes:

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

1. Describe the basic concepts in python.
2. Make use of arrays and functions.
3. Identify and apply the data structures to solve a problem.
4. Demonstrate the use of Object Oriented concepts in problem solving.
5. Perform pattern matching using regular expression, exception and file handling
6. Make use of different python libraries for data processing

Prerequisites:

Programming Language, Data Structures

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	2	2	-	-	-	-	-	1	2	2	-	2	2	-	2
2	2	2	2	-	-	-	-	1	2	2	-	2	2	-	3
3	2	2	2	-	-	-	-	1	2	2	-	2	2	-	3
4	2	2	2	-	-	-	-	1	2	2	-	2	2	-	3
5	2	2	2	-	-	-	-	1	2	2		2	2	-	3
6	2	2	2	-	3	-	-	1	2	2		2	2	-	3

Course Contents

Unit 1	Getting Started with Python	04 Hours

Python overview, Features, Python installation and Working of it, Advantages, Applications, Python Code Execution, Variables, Namespaces, Statement, Indentation, Comments, Input and Output, Python blocks, Control statements, Branching statements.

Unit 2	Array, Function, Modules & Packages	04 Hours

Array: When to use array, how to use array, defining array, length of array, array indexing, searching in array, loop through array, array slice, operations on array

Functions: Defining Functions, Calling Functions, Passing Arguments, Keyword Arguments, Default Arguments, Scope and Lifetime of Variables in a Function, Anonymous / Lambda function, map() function, reduce() function, filter() function,

Modules: Creating modules, import/export modules

Packages: Introduction to PIP, Installing Packages via PIP, Using Python Packages

Unit 3	Data types in Python	04 Hours

String: Creating string, accessing elements of string, string length, concatenation, String formatting operator, built-in string methods and function, slice operation.

List: Creating list, access and update values in lists, nested and cloning lists, basic list operations, List methods, list comprehensions, looping in lists.

Tuple: Creating tuple, accessing values in a tuple, updating tuple, deleting elements in tuple, basic tuple operations

Dictionary: Creating a dictionary, accessing values, add, modify, delete, sort items in a dictionary, looping over a dictionary

Set: Creating a Set and set operations

Unit 4	Object Oriented Programming using Python	04 Hours

Procedural and Object-Oriented Programming, Objects, class, Method overloading, Polymorphism, Inheritance

Unit 5	File Handling & Exception Handling	04 Hours

File Handling: Files and File Paths, The os.path Module, The File Reading/Writing Process, shutil Module, Directories, Compressing Files with the zipfile Module

Regular Expression: Finding Patterns of Text without Regular Expressions, Finding Patterns of Text with Regular Expressions, More Pattern Matching with Regular Expressions, match(), search(), findall() function.

Exception Handling: What is Exception?, Handling an exception, try….except…else, try-finally clause, Argument of an Exception, Python Standard Exceptions, Raising an exceptions, User-Defined Exceptions

Unit 6	Python Libraries - Numpy, Pandas, Matplotlib	04 Hours

Numpy: Introduction to numpy, creating arrays, Using arrays and Scalars, Indexing Arrays, Array Transposition, Universal Array Function, Array Processing, Array Input and Output.

Pandas: Introduction of pandas, Series in pandas, Index objects, Reindex, Drop Entry, Selecting Entries, Data Alignment, Rank and Sort, Summary Statics, Missing Data, Index Hierarchy.

Matplotlib: Simple Line Plots, Simple Scatter Plots, Visualizing Errors, Histogram.

Textbooks:

1. Beginning Python: Using Python 2.6 and Python 3.1., James Payne, Wrox Publication.
2. Python Programming, Anurag Gupta, G. P. Biswas, McGraw-Hill
3. Introduction to computing and problem-solving using python, E. Balagurusamy, McGraw Hill Education
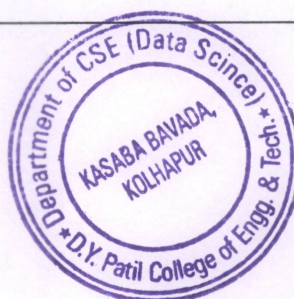
Reference Books:
1. Learn Python the Hard Way, Zed Shaw's Hard Way Series, 3rd Edition.
2. Python Projects, Laura Cassell, Alan Gauld, Wrox Publication.

Useful Links:
1. Virtual Lab - https://python-iitk.vlabs.ac.in/
2. NPTEL Videos - https://nptel.ac.in/courses/106/106/106106182/

Experiment List:

Expt. No.	Experiment	S/O	Hours
1	To study the installation of Python 3.0	S	2
2	Write a program to implement control and branching statements.	O	2
3	Write a program to create and manipulate arrays.	O	2
4	Write a program to create and use functions.	O	2
5	Write a program to create anonymous functions and make use of map(), reduce() & filter() functions.	O	2
6	Write a program to create and use the python modules	O	2
7	Write a program to import modules and packages from standard libraries and third-party repositories.	O	2
8	Write a program to implement and manipulate string & lists.	O	2
9	Write a program to implement and manipulate tuples & dictionaries.	O	2
10	Write a program to perform different set operations	O	2
11	Write a program to create classes and objects using python.	O	2
12	Write a program to implement inheritance and polymorphism.	O	2
13	Write a program to read from and write to text files.	O	2
14	Write a program to find the patterns in text file using regular expression	O	2
15	Write a program to implement the exception handling.	O	2
16	Write a program to implement array operations using Numpy.	O	2
17	Write a program to implement universal function in Numpy.	O	2
18	Write a program to implement different data operations in Pandas.	O	2
19	Write a program to implement different charts and graphs in Matplotlib.	O	2
20	Students are supposed to develop a mini project using all features of python programming.	O	4
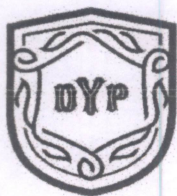
D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – IV
Course Code: 231DSMDML202
Course Name: Methodology of Data Analytics using Machine Learning

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 02 Hrs/ Week	02	ISE: 20 Marks
Tutorials: 00 Hrs/Week		MSE: --
Practicals: 00 Hrs/Week		ESE: 30 Marks

Course Description:

This course is an introduction to the theoretical aspects of design and implementation of algorithms that enable machines to "learn" from experience. Course will provide in-depth knowledge to the areas of Supervised and Unsupervised Machine Learning.

Course Objectives:

❖ To understand classification algorithms to classify multivariate data.

❖ To understand machine learning algorithms & concepts.

Course Outcomes:

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

1. Describe the different machine learning terminologies.
2. Explain the different linear regression concepts.
3. Explain the different logistic regression concepts.
4. Describe different supervised and unsupervised algorithms.

Prerequisites:

Probability Theory, Statistics

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	2	2	-	-	-	-	-	-	-	-	-	-	-	2	2
2	2	2	-	-	-	-	-	-	-	-	-	-	-	2	2
3	2	2	-	-	-	-	-	-	-	-	-	-	-	2	2
4	2	2	-	-	-	-	-	-	-	-	-	-	-	2	2

Course Contents

Unit 1	Introduction to Machine Learning	06 Hours

History and Evolution, Artificial Intelligence Evolution, Different Forms, Machine Learning Categories, Frameworks for Building Machine Learning Systems, Machine Learning Python Packages, Data Analysis Packages, Machine Learning Core Libraries.

Unit 2	Fundamentals of Machine Learning	06 Hours
Machine Learning Perspective of Data, Scales of Measurement, Feature Engineering, Exploratory Data Analysis (EDA), Supervised Learning– Regression		

| Unit 3	Supervised Learning	06 Hours
Regression: Correlation and Causation, Fitting a Slope, How Good Is Your Model?, Polynomial Regression, Multivariate Regression, Multicollinearity and Variation Inflation Factor (VIF), Interpreting the OLS Regression Results, Regression Diagnosis, Regularization, Nonlinear Regression. **Classification:** Logistic Regression, Evaluating a Classification Model Performance, ROC Curve, Fitting Line, Stochastic Gradient Descent, Regularization, Multiclass Logistic Regression, Generalized Linear Models, Decision Trees, Support Vector Machine (SVM), k Nearest Neighbors (kNN), Time-Series Forecasting.		

Unit 4	Unsupervised Learning	06 Hours
Unsupervised Learning Process Flow – Clustering, K-means, Finding Value of k, Hierarchical Clustering, Principal Component Analysis (PCA)		

Textbooks:
1. Mastering Machine Learning with Python in Six Steps, Manohar Swamynathan, Apress Publication. [Unit 1-4]

Reference Books:
1. Machine Learning by Tom M. Mitchell, International Edition 1997, McGraw Hill Education.
2. Machine Learning, Anuradha Srinivasara ghavan, and Vincy Joseph, Kindle Edition, 2020, WILEY.
3. Introduction to Machine Learning, Ethem Alpaydin, Second Edition, 2010, Prentice Hall of India.
4. Practical Machine Learning Sunila Gollapudi Packt Publishing Ltd.

Useful Links:
1. https://www.javatpoint.com/machine-learning
2. https://onlinecourses.nptel.ac.in/noc23_cs18/preview
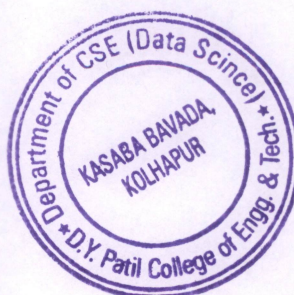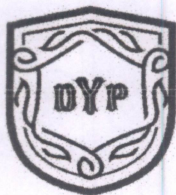
	D. Y. Patil College of Engineering & Technology, Kolhapur
	(An Autonomous Institute)
	Second Year B. Tech CSE (Data Science)
	Semester – IV
	Course Code: 231DSOECL202
	Course Name: Introduction to Data Engineering

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 02 Hrs/ Week	02	ISE: --
Tutorials: 00 Hrs/Week		MSE: --
Practicals: 00 Hrs/Week		ESE: 50 Marks

Course Description:

This course is about the understanding of fundamental techniques involved in the data engineering and will provide understanding of data engineering life cycle. Also, includes topics which focus on source systems of data engineering, storage, ingestion, Security, data Management, Data modelling and Design. They are used in a variety of applications today including Business Intelligence and Analytics, smart cities, healthcare, fraud detection.

Course Objectives:

❖ To make students to learn the basic principles, foundation and building blocks of Data Engineering.
❖ To understand the data engineering lifecycle and ETL model.
❖ To understand the need of basic architecture in data engineering.

Course Outcomes:

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

1. Describe the basic principles, foundation and building blocks of Data Engineering.
2. Define the data engineering lifecycle and ETL model.
3. Explain the need of basic architecture in data engineering.
4. Summarize the technologies used for implementation of data engineering lifecycle.

Prerequisites:

Fundamental of Data Science

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	3	2	1	-	-	-	-	-	-	-	-	-	1	-	2
2	2	2	1	-	-	-	-	-	-	-	-	-	1	-	2
3	2	2	1	-	-	-	-	-	-	-	-	-	1	-	2
4	2	1	1	-	-	-	-	-	-	-	-	-	1	-	2

Course Contents

Unit 1	Foundation and Building Blocks of Data Engineering	06 Hours

What is Data Engineering - Data Engineering Defined, Data Engineering Lifecycle, Evolution of the Data Engineer, Data Engineering and Data Science.

Data Engineering Skills and Activities - Data Maturity and the Data Engineer, The Background and Skills of a Data Engineer, Business Responsibilities, Technical Responsibilities.

Data Engineers Inside an Organization - Internal-Facing Versus External-Facing Data Engineers, Data Engineers and Other Technical Roles, Data Engineers and Business Leadership.

Unit 2	The Data Engineering Life Cycle	06 Hours

What is data engineering life cycle - The data lifecycle vs the data engineering lifecycle, source systems, storage, ingestion, Batch vs streaming, push vs pull, Transformation, serving Data, Analytics, Machine Learning, Reverse ETL.

Major undercurrents across the Data Engineering Lifecycle - Security, data Management, Data modelling and Design, Data Lineage, Data Integration and interoperability, Data Lifecycle management, DataOps

Unit 3	Designing good data architecture	06 Hours

What is data architecture, enterprise architecture, Good data architecture, principles of good data architecture, Major architecture concepts, tight vs loose coupling, examples and types of Data architecture

Unit 4	Choosing technologies across Data Engineering Lifecycle	06 Hours

Team size and capabilities, Speed to market, Interoperability, Cost optimization and business value, Today versus the future: immutable versus transitory technologies, Location (cloud, on premises, hybrid cloud, multi cloud), Build versus buy, Monolith versus modular, Serverless versus servers, Optimization, performance and the benchmark wars, The undercurrents of the data engineering lifecycle

Textbooks:
1. Fundamentals of Data Engineering, Joe Reis &Matt Housley, O'REILLY. [Unit 1-4]

Reference Books:
1. Designing Data-Intensive Applications, Martin Kleppmann, O'REILLY
2. Data Engineering with Python, Paul Crickard, Packt Publication

Useful Links:
1. https://elearn.nptel.ac.in/shop/iit-workshops/completed/introduction-to-data-engineering-using-azure/?v=c86ee0d9d7ed
2. https://www.coursera.org/specializations/data-engineering-foundations
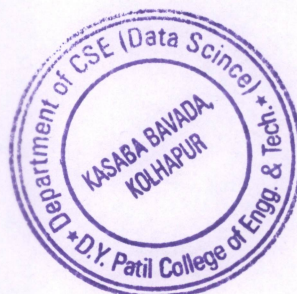
D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – IV
Course Code: 231DSAECP201
Course Name: Soft Skill

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 00 Hr./ Week Tutorials: 00 Hrs/Week Practicals: 04 Hrs/Week	02	ISE: 25 Marks ESE-POE: 25 Marks

Course Description:

Soft skills are character traits and interpersonal skills that characterize a person's relationships with other people. This course includes Communication skills, Writing skills, Techniques for self- development, Teamwork and group discussions, Time and stress management, Professional skills for overall development of an Engineer.

Course Objectives:

❖ To equip students with professional soft skills like communication, interviews, group discussion, presentation etc.

❖ To enable the students to learn interpersonal skills, work culture and effective management of time and stress.

Course Outcomes:

On completion of the course, students will be able to –
1. Effectively use the principles of communication.
2. Make appropriate use of interviews techniques.
3. Develop skills to conduct meetings & conferences.
4. Make effective presentations & technical report writing.
5. Actively participate in group discussion by following its etiquettes.
6. Effectively manage time and stress.

Prerequisites:

Basic English Knowledge

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	-	-	-	-	-	1	-	-	3	3	-	3	2	-	3
2	-	-	-	-	-	1	-	-	3	3	-	3	2	-	3
3	-	-	-	-	-	1	-	-	3	3	-	3	2	-	3
4	-	-	-	-	2	1	-	-	3	3	-	3	2	-	3
5	-	-	-	-	-	1	-	-	3	3	-	3	2	-	3

6	-	-	-	-	2	1	-	-	3	3	-	3	2	-	3

Course Contents

Unit 1	Communication Skills	06 Hours

Process of communication, Flows of Communication in organization, Barriers to communication (Formal Flow – Upward, Downward, lateral and diagonal, Strategies to improve Organizational Communication, Effectiveness in Managerial Communication, and importance of technical communication, Nonverbal communication.

Unit 2	Interviews Skills	07 Hours

Types of interview, General preparation for interview, gathering information about the company, knowing about the role/job position, Types of interviewing questions, Non-verbal communication to win the interview.

Unit 3	Meeting & Conferences	07 Hours

Planning a meeting (Agenda and notice), Conducting a meeting, Post meeting actions (Minutes), Planning & Conducting a Conference (anchoring and Report writing), and Video/web conferences, Identifying Strengths and Weakness.

Unit 4	Presentation Skills	07 Hours

Effective Presentation strategies: Purpose, analyzing the audience and locale, organizing the content Oral presentation, Graphic presentation, Presentation aids, Personality Development. Newsletters, technical article and business letters.

Technical Reports, characteristics, Importance, objectives, categories of report, format structure of reports, types of reports

Unit 5	Group Discussion	08 Hours

Qualities needed for effective group discussion. Email etiquettes, Telephone Etiquettes, Role and responsibility of engineer, Work culture in jobs. Work place, rights and responsibilities.

Unit 6	Time and Stress Management	07 Hours

Concept & Importance of Time Management, Techniques of Time Management, and Concept & Importance of Stress Management, Techniques of Stress Management, and Overcoming Stage fear and Interpersonal Relationships.

Textbooks:

1. Business Communication and Soft Skill, G.S.B.K Babu Rao, Himalaya Publishing house (1st Edition)
2. Pocket Style Manual, Diane Hacker, Bedford Publication, New York, 2003. (ISBN 0312406843)
3. You Can Win, Shiv Khera, Macmillan Books, New York, 2003.

Reference Books:

1. Technical Communications, Raman Sharma, Oxford Publication, London, 2004.
2. Ethics in Engineering practice and research, Caroline Whit beck, Cambridge 2nd Edition.
3. Business Correspondence and Report Writing, Sharma, R. and Mohan, K., TMH New Delhi 2002.

Useful Links:

1. https://onlinecourses.nptel.ac.in/noc21_hs76/preview
2. https://mindremakeproject.org/2018/07/21/free-online-assessment-and-screening-tools-for-mental-health/

Activity List:

Activity No.	Activities	S/O	Hours
1	Self-assessment and development - SWOT analysis	O	2
2	Resume writing	O	2
3	Verbal and Non Verbal Communication through various means.	O	2
4	Personal Interview	O	2
5	Group Discussion	O	2
6	Learning about different Stress Management techniques.	O	2
7	Prepare and present visual presentation with charts, PowerPoint presentations etc.	O	2
8	Verbal presentation on video clips	O	2
9	Learning about listening and reading comprehensions.	O	2
10	Problem solving and decision-making	O	2
11	Confidence building and corporate laws.	O	2
12	Technical Report writing	O	2
13	Email writing etiquettes	O	2
14	Letter writing	O	2
15	Social Media Platforms Interaction	O	2
16	Role-play activity	O	2

D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – IV
Course Code: 231DSVSECP201
Course Name: Web Application Development – I

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 01 Hr./ Week Tutorials: 00 Hrs/Week Practicals: 02 Hrs/Week	02	ISE: 25 Marks ESE-POE: 50 Marks

Course Description:

This course is about the understanding and application development using the front end technologies. This aims to equip the students with different front end technologies needed to design and develop the applications of different problems related to UI interface.

Course Objectives:

❖ To expose students to emerging web technologies and related tools.
❖ To introduce client-side technologies required for development of web applications.

Course Outcomes:

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

1. Develop static web pages as per the requirement.
2. Create responsive web pages as per the requirement.
3. Write functionalities to make dynamic web pages.

Prerequisites:

Basic Knowledge of Computer

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	2	2	2	-	2	-	-	-	-	-	-	-	2	-	6
2	2	2	2	-	2	-	-	-	-	-	-	-	2	-	6
3	2	2	2	-	2	-	-	-	-	-	-	-	2	-	6

Course Contents

Unit 1	HTML & CSS	05 Hours

HTML: HTML Structure, Block Elements, Inline Elements, Class and ID Attributes, HTML Whitespaces.

CSS SELECTOR: Type, Class and ID Selector, Position and Group Selectors, Attribute Selectors, Pseudo-element Selectors, Pseudoclass Selectors.

Box Model: Display, Box Model, Inline Box, Inline-Block Box.

Unit 2	Responsive Web Designing	05 Hours

Responsive Web Designing: Introduction, Viewport, Grid View, Image, Video, Media Queries, RWD frameworks.

Twitter Bootstrap : Grid Basics, Typography, Tables, Images, Alerts, Button, Button Group, Borders, Labels, Progress bar, Pagination, Tabs, Navbar, Forms, Inputs, Input sizing, Carousel, Scrollspy.

Unit 3	JavaScript	03 Hours

Introduction, Data types and Variables, Operators, Expressions and Statements, Functions and Scope, Document Object Model, Event Handling, Form handling and validations.

Unit 4	jQuery	03 Hours

Introducing jQuery, jQuery selector, Animation effects, Event handling, DOM, jQuery DOM traversing, DOM manipulation.

Textbooks:

1. Pro HTML5 and CSS3 Design Patterns, Michael Bowers, Dionysios Synodinos and Victor Sumner, Apress edition. [Unit 1]
2. Twitter Bootstrap Development How to, David Cochran, Packt Publication. [Unit 2]
3. JavaScript: The Definitive Guide, David Flanagan, O'Reilly Media. [Unit 3]
4. jQuery in Action, Bear Bibeault, Manning Publication. [Unit 4]

Reference Books:

1. Beginning with HTML5 and CSS3 The Web Evolved, Murphy, Apress.
2. JavaScript: The Complete Reference, Thomas A Powell, Fritz Schneider, Tata McGraw Hill.
3. Head First jQuery, Ryan Benedetti, O'reilly Publication.

Useful Links:

1. https://www.w3schools.com/html/default.asp
2. https://www.w3schools.com/css/default.asp
3. https://www.w3schools.com/js/default.asp
4. https://www.w3schools.com/jquery/default.asp

Experiment List:

Expt. No.	Experiment	S/O	Hours
1	Create Web Page structure using HTML5.	O	2
2	Create Web Pages with Class and ID attributes using HTML5.	O	2
3	Apply CSS to web pages created after developing the HTML5 pages.	O	2
4	Apply different CSS selectors to HTML5 web pages.	O	2
5	Create a responsive web page using media queries.	O	2
6	Create a responsive web page using bootstrap.	O	2
7	Write a JavaScript to compute mathematical operations on client side.	O	2
8	Write a JavaScript to handle event generated by client.	O	2
9	Write a JavaScript to perform form validation.	O	2
10	Write a jQuery script to provide animations effects in web pages.	O	2

11	Write a jQuery script to handle event generated by client.	O	2
12	Write a jQuery script to manipulate DOM	O	2

D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – IV
Course Code: 231DSVECL202
Course Name: Human Values & Ethics

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 02 Hrs/ Week	02	ISE: --
Tutorials: 00 Hrs/Week		MSE: --
Practicals: 00 Hrs/Week		ESE: 50 Marks

Course Description:

The methodology of this course is universally adaptable, involving a systematic and rational study of the human being vis-à-vis the rest of existence. It is free from any dogma or value prescriptions. This process of self-exploration takes the form of a dialogue between the teacher and the students to begin with and within the student himself/herself finally.

Course Objectives:

❖ To create an awareness on Engineering Ethics and Human Values.
❖ To understand social responsibility of an engineer.
❖ To appreciate ethical dilemma while discharging duties in professional life.

Course Outcomes:

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

1. Understand the significance of value inputs in a classroom and start applying them in their life and profession.
2. Distinguish between values and skills, happiness and accumulation of physical facilities, the Self and the Body, Intention and Competence of an individual, etc.
3. Understand the role of a human being in ensuring harmony in society and nature.
4. Distinguish between ethical and unethical practices, and start working out the strategy to actualize a harmonious environment wherever they work.

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	-	-	-	-	-	1	-	3	2	2	-	2	1	1	2
2	-	-	-	-	-	2	-	3	2	2	-	2	1	1	2
3	-	-	-	-	-	2	2	3	2	2	-	2	1	1	2
4	-	-	-	-	-	2	2	3	2	2	2	2	1	1	2

Course Contents

Unit 1	Introduction to Value Education	06 Hours

Value Education - Definition, Concept and Need for Value Education.

The Content and Process of Value Education - Basic Guidelines for Value Education, Self-exploration as a means of Value Education, Happiness and Prosperity as parts of Value Education.

Unit 2	Harmony in the Human Being	06 Hours

Human Being is more than just the Body. Harmony of the Self ('I') with the Body. Understanding Myself as Co-existence of the Self and the Body. Understanding Needs of the Self and the needs of the Body. Understanding the activities in the Self and the activities in the Body.

Unit 3	Harmony in the Family, Society and in the Nature	06 Hours

Family as a basic unit of Human Interaction and Values in Relationships. The Basics for Respect and today's Crisis: Affection, e, Guidance, Reverence, Glory, Gratitude and Love. Comprehensive Human Goal: The Five Dimensions of Human Endeavour. Harmony in Nature: The Four Orders in Nature. The Holistic Perception of Harmony in Existence.

Unit 4	Social & Professional Ethics	06 Hours

The Basics for Ethical Human Conduct. Defects in Ethical Human Conduct. Holistic Alternative and Universal Order. Universal Human Order and Ethical Conduct. Human Rights violation and Social Disparities.

Value based Life and Profession. Professional Ethics and Right Understanding. Competence in Professional Ethics. Issues in Professional Ethics – The Current Scenario. Vision for Holistic Technologies, Production System and Management Models.

Textbooks:

1. Human Values, A. N. Tripathy, New Age International Publishers, 2003 (3rd Edition).
2. Indian Ethos and Modern Management, Bajpai. B. L., New Royal Book Co, Lucknow, Reprinted, 2004.
3. Human Society in Ethics & Politics, Bertrand Russell, Routledge London

Reference Books:

1. The Philosophy of Humanism, Corliss Lamont, Humanist Press
2. A Foundation Course in Value Education, Gaur. R.R. ,Sangal. R, Bagaria. G.P, Excel Books, 2009
3. Teachers Manual, Gaur. R.R. ,Sangal. R ,Bagaria. G.P, Excel Books, 2009.
4. Ethical Philosophy of India, I.C. Sharma, Nagin & co Julundhar
5. Introduction to Ethic, William Lilly, Allied Publisher

Useful Links:

1. **https://fdp-si.aicte-india.org/SIPTeachMaterialM1.php**

D. Y. Patil College of Engineering & Technology, Kolhapur
(An Autonomous Institute)
Second Year B. Tech CSE (Data Science)
Semester – IV
Course Code: 231DSHSSML202
Course Name: Programming Ethics

Teaching Scheme:	Credits	Evaluation Scheme:
Lectures: 02 Hrs/ Week	02	ISE: --
Tutorials: 00 Hrs/Week		MSE: --
Practicals: 00 Hrs/Week		ESE: 50 Marks

Course Description:

This course explores the ethical dimensions of programming, focusing on the responsibilities of programmers in creating technology that impacts individuals, societies, and the environment. Students will examine ethical theories and principles, apply them to real-world programming scenarios, and develop strategies for ethical decision-making and advocacy.

Course Objectives:

❖ To understand the ethical theories & principles of programming.
❖ To analyze ethical issues of programming.
❖ To develop ethical responsibilities in project.
❖ To apply ethical frameworks in development.

Course Outcomes:

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

1. Understand ethical theories and principles relevant to programming.
2. Analyze ethical issues in programming, including privacy, security, bias, and social impact.
3. Develop critical thinking skills to evaluate the ethical implications of programming decisions.
4. Apply ethical frameworks to programming practices and project development.
5. Foster a culture of ethical responsibility and accountability in programming communities.

Course Articulation Matrix: Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO's	PO's												PSO's		BTL
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	
1	-	-	-	-	-	2	1	3	1	2	-	2	1	-	2
2	-	-	-	-	-	2	1	3	1	2	-	2	1	-	3
3	-	-	-	-	-	2	1	3	1	2	-	2	1	-	3
4	-	-	-	-	-	2	1	3	1	2	-	2	1	-	3
5	-	-	-	-	-	2	1	3	1	2	-	2	1		3

Course Contents

Unit 1	**Introduction to Programming Ethics**	**4 Hours**

Overview of ethical theories and principles, Ethical considerations in programming: privacy, security, fairness, and transparency.

Case studies: Ethical dilemmas in software development

Unit 2	**Privacy and Security with Bias and Fairness**	**4 Hours**

Ethical implications of data collection and surveillance, protecting user privacy in software design, Addressing security vulnerabilities. Understanding bias in algorithms and data Mitigating bias in machine learning and AI systems. Ensuring fairness and equity in programming practice.

Unit 3	**Intellectual Property, Open Source and Professional Responsibility**	**5 Hours**

Copyright, patents, and licensing in software development, Ethical considerations in open source projects, Balancing innovation and intellectual property rights, Codes of ethics and conduct for programmers, Ethical decision-making frameworks, Ethical responsibilities in project management and team collaboration.

Unit 4	**Social Impact of Technology**	**5 Hours**

Technology and inequality: Digital divide, access, and exclusion.

Environmental impact of technology: Sustainability and green computing, Ethical considerations in technology for social good and humanitarian projects.

Unit 5	**Emerging Technologies, Ethical Leadership and Advocacy**	**6 Hours**

Ethical challenges in emerging technologies: AI, Blockchain, IoT, etc. Regulation and governance of emerging technologies, Ethical innovation and responsible technology development. Promoting ethical behavior in programming communities and workplaces, Ethical leadership and decision-making in technology companies, Advocating for ethical technology policies and practices.

Textbooks:

1. Ethics in Computing: A Concise Module, John R. Magnifico.
2. Ethics for the information edge, Michael J. Quinn, 8th Edition.

Reference Books:

1. Ethics and Technology, Herman Tavani, Wiley Publication.
2. Professional Codes of Conduct and Computer Ethics Education, Martin C and Martin D Computers and Society, Vol 20, No 1, June 1990.

Useful Links:

1. https://onlinecourses.nptel.ac.in/noc22_mg54/preview
2. http://www.cwru.affil/wwwethics/home.html
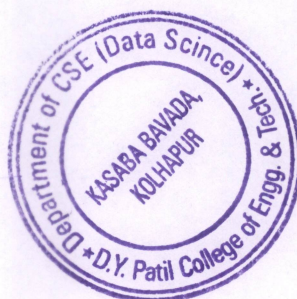
	D. Y. Patil College of Engineering & Technology, Kolhapur
	(An Autonomous Institute)
	Second Year B. Tech CSE (Data Science)
	Semester – IV
	Course Code: 231DSCCA202
	Course Name: Liberal Learning

Teaching Scheme:	Audit Course	Evaluation Scheme:
Lectures: 00 Hrs/ Week		Grade Based Evaluation
Tutorials: 00 Hrs/Week		
Practicals: 02 Hrs/Week		

List of Clubs Available at Department:

1. Data Analytics Club:

Aim:

The primary aim of a Data Analytics Club is to create a collaborative and dynamic environment where students can enhance their knowledge, skills, and practical experience in data analytics. This involves fostering a community that promotes learning, innovation, and professional growth in the field of data analytics. The club seeks to bridge the gap between academic knowledge and real-world application, preparing members for successful careers in data analytics by providing resources, networking opportunities, and hands-on experiences.

Objectives:

- **Skill Development:**
  - Technical Skills: Provide training and workshops on data analytics tools and technologies such as Python, R, SQL, Tableau, Power BI, and machine learning.
  - Soft Skills: Enhance communication, teamwork, problem-solving, and project management skills through collaborative projects and presentations.
- **Knowledge Expansion:**
  - Guest Lectures: Invite industry professionals to speak about current trends, best practices, and real-world applications of data analytics.
  - Industry Trends: Keep members informed about the latest developments in data analytics, including new tools, techniques, and methodologies.
- **Hands-on Experience:**
  - Projects: Offer opportunities to work on real-world data projects, either individually or in teams, to apply theoretical knowledge in practical scenarios.
  - Competitions: Organize and participate in data analytics competitions and hackathons to foster a competitive and innovative spirit.
- **Networking:**
  - Industry Connections: Facilitate connections with professionals and alumni working in the field of data analytics.
  - Peer Networking: Create a community where students can collaborate, share knowledge, and support each other's learning journeys.
- **Career Preparation:**
  - Internship Opportunities: Provide information about internships and job openings in the field of data analytics.

- Resume Building: Offer workshops on resume writing, LinkedIn profile optimization, and interview preparation specific to data analytics roles.
- **Research & Innovation:**
  - Research Projects: Encourage and support members in conducting research projects and publishing their findings in academic or industry journals.
  - Innovation: Foster a culture of innovation by encouraging members to explore new ideas and approaches in data analytics.

Outcomes:
- **Enhanced Skill Set:** Members will develop a strong foundation in data analytics tools and techniques, making them more competitive in the job market.
- **Practical Experience:** Participation in real-world projects and competitions will provide practical experience, helping members to apply theoretical knowledge in practical scenarios.
- **Professional Growth:** Networking opportunities with industry professionals and alumni will help members build valuable connections and gain insights into the industry.
- **Career Readiness:** Members will be better prepared for careers in data analytics through exposure to job opportunities, resume workshops, and interview preparation sessions.
- **Research Contributions:** Members will have opportunities to contribute to the field through research projects and publications.
- **Community Building:** The club will create a supportive community where students can share knowledge, collaborate on projects, and help each other grow.
- **Innovation and Creativity:** Members will be encouraged to think creatively and innovate, leading to new ideas and approaches in data analytics.
- **Increased Confidence:** Through presentations, workshops, and networking events, members will gain confidence in their abilities to communicate and apply data analytics concepts.
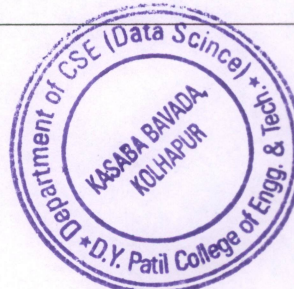
2. Open Source Software Developers Club:

Aim:

The primary aim of an Open Source Student Club is to foster a community of students who are passionate about open source software and collaboration. The club seeks to promote the use, development, and contribution to open source projects, encouraging members to learn, innovate, and share their knowledge with the wider community. By doing so, the club aims to empower students with the skills and mindset needed to contribute meaningfully to the open source ecosystem and to prepare them for careers in technology and software development.

Objectives:
- **Education & Skill Development:**
  - Technical Workshops: Provide training on various open source technologies, programming languages, and development tools.
  - Best Practices: Teach best practices for contributing to open source projects, including version control, code reviews, and collaboration techniques.
- **Project Involvement:**
  - Contributions: Encourage and guide members to contribute to existing open source projects.
  - Initiate Projects: Support members in starting and maintaining their own open source projects.
- **Community Engagement:**

- Collaboration: Foster a collaborative environment where members can work together on projects and share knowledge.
  Outreach: Engage with the wider open source community through events, meetups, and online platforms.
- **Networking:**
  - Industry Connections: Facilitate connections with professionals, open source contributors, and organizations in the open source community.
  - Peer Networking: Create opportunities for members to network with each other and build lasting professional relationships.
- **Career Preparation:**
  - Internship and Job Opportunities: Provide information about internships, job openings, and career paths in open source development.
  Professional Development: Offer workshops on resume building, portfolio creation, and interview preparation specific to open source careers.

Outcomes:
- **Enhanced Technical Skills:** Members will develop proficiency in open source tools and technologies, improving their coding and development skills.
- **Practical Experience:** Participation in real-world open source projects will provide hands-on experience and a deeper understanding of software development.
- **Community Contribution:** Members will make meaningful contributions to open source projects, helping to advance the open source ecosystem.
- **Professional Growth:** Networking opportunities with industry professionals and active contributors will help members build valuable connections and gain insights into the industry.
- **Career Readiness:** Members will be better prepared for careers in open source development through exposure to job opportunities, portfolio building, and interview preparation sessions.
- **Innovation & Creativity:** Members will be encouraged to innovate and explore new ideas, leading to the development of new open source projects and solutions.
- **Increased Confidence:** Through workshops, presentations, and collaboration, members will gain confidence in their abilities to contribute to open source projects and communities.
- **Community Building:** The club will create a supportive and collaborative community where students can share knowledge, work on projects together, and help each other grow.
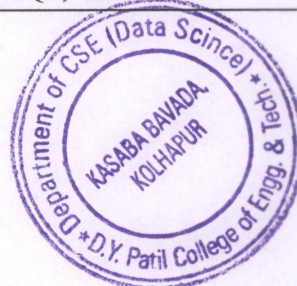
3. Rational Programmers Club:

Aim:

The primary aim of a Rational Programmers Club within the Data Science Department is to cultivate a community where students can enhance their coding skills, apply their knowledge to solve real-world problems, and prepare for careers in data science and related fields. The club seeks to create an environment that encourages continuous learning, collaboration, and innovation in coding, particularly as it applies to data science.

Objectives:
- **Skill Development:**
  - Programming Efficiency: Provide training in various programming languages commonly used in data science, such as Python, R, SQL, and Java.
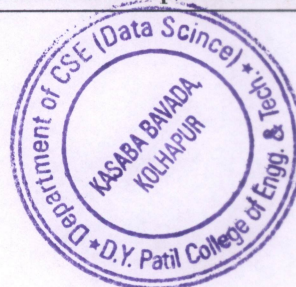
- Data Science Tools: Teach members how to use data science tools and libraries, such as pandas, NumPy, scikit-learn, TensorFlow, and more.
- **Practical Applications:**
  - Projects: Encourage members to participate in or initiate coding projects that solve real-world data science problems.
  - Hackathons: Organize and participate in hackathons and coding competitions to foster a spirit of innovation and problem-solving.
- **Knowledge Sharing:**
  - Workshops & Seminars: Conduct workshops, seminars, and coding boot camps to share knowledge and best practices in coding and data science.
  - Peer Learning: Promote peer-to-peer learning through study groups, code reviews, and collaborative projects.
- **Career Preparation:**
  - Portfolio Building: Help members build a portfolio of coding projects that demonstrate their skills and knowledge in data science.
  - Career Guidance: Provide guidance on career paths in data science, including resume building, interview preparation, and internship/job search strategies.
- **Networking:**
  - Industry Connections: Facilitate connections with data science professionals, alumni, and industry partners.
  - Community Building: Create a supportive community where members can network, collaborate, and share resources.
- **Innovation & Research:**
  - Cutting-edge Topics: Explore and work on cutting-edge topics in data science, such as machine learning, artificial intelligence, big data analytics, and more.
  - Research Projects: Encourage members to participate in or initiate research projects and publish their findings.

Outcomes:
- **Enhanced Coding Skills:** Members will develop strong programming skills and a deep understanding of data science tools and techniques.
- **Practical Experience:** Participation in real-world projects and hackathons will provide hands-on experience, enabling members to apply their knowledge practically.
- **Professional Growth:** Networking with industry professionals and peers will help members build valuable connections and gain insights into the data science industry.
- **Career Readiness:** Members will be better prepared for careers in data science through portfolio building, resume workshops, and interview preparation sessions.
- **Community Contributions:** The club will create a collaborative community where members can share knowledge, support each other's learning, and contribute to the field of data science.
- **Innovation & Research Contributions:** Members will have opportunities to work on innovative projects and research, contributing to advancements in data science.
- **Increased Confidence:** Through coding challenges, presentations, and collaborative projects, members will gain confidence in their abilities to code and solve complex data science problems.
- **Lifelong Learning:** The club will instill a mindset of continuous learning and curiosity, encouraging members to stay updated with the latest developments in coding and data science.
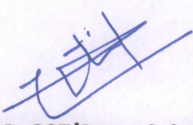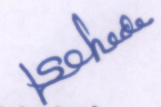
Evaluation:
Evaluation of individual student will be carried out based on following criteria -   1. Knowledge & Understanding   2. Critical Thinking   3. Communication Skills   4. Ethical Reasoning   5. Cultural Awareness and Diversity   6. Interdisciplinary Connections   7. Creativity and Innovation   8. Collaboration and Teamwork

Program Coordinator   **HOD CSE(Data Science)**   **Dean Academics**   **Principal**