IPS Academy, Institute of Engineering & Science

(A UGC Autonomous Institute, Affiliated to RGPV, Bhopal) Scheme
Based on AICTE Flexible Curriculum

Department of Computer Science & Engineering

Bachelor of Technology (B. Tech.)

[Computer Science & Engineering (Data Science)]

III – Semester (Scheme & Syllabus)

Sr. No.	Course Type	Course Code	Course Name	Teaching Scheme			Credits
				L	T	P	
1.	BSC	MA03	Probability and Statistics	2	1	-	3
2.	PCC	DS01	Data Communication	2	1	-	3
3.	PCC	DS02	Data Structures and Algorithm	2	1	-	3
4.	PCC	DS03	Computer System Organization	2	1	-	3
5.	PCC	DS04	Foundation of Data Science	2	1	-	3
6.	HSMC	HS04	Entrepreneurship Principles Of Management	1	-	-	1
7.	LC	DS02(P)	Data Structures and Algorithm Lab	-	-	4	2
8.	LC	DS04(P)	Foundation of Data Science Lab	-	-	2	1
9.	SBC	DS01(P)	Programming in Python	-	-	4	2
10.	LLC	LLC02	Liberal Learning Course –II	Credits to be added in Fourth Semester			
11.	MLC	MLC01	Professional Laws, Ethics, Gender, Human Values and Harmony	1	-	-	Audit
Total Academic Engagement and Credits				**12**	**5**	**10**	**21**
				27			

PCC- DS01	Data Communication	2L: 1T:0P (3Hrs.)	Credits: 03

Prerequisite: -Communication System, Digital Communications

Course Objective: The course is designed to understand the basic technologies used in data communication like mode of communication, interfacing of devices, medium of communication and detection and correction of errors occurs during data transmission.

Course Contents: **(40 hrs.)**

Module1: (08Hrs.)

Introduction to data communication: Components, data representation, data flow and basic model, data representation, Serial & Parallel transmission, Modes of data transmission, Encoding: Unipolar, Polar, Bipolar line & block codes, Data compression, Frequency dependent codes, Run length encoding, Relative encoding, LZ Compression, Image and multimedia compression. Review of analog & digital transmission methods

Module 2: (08Hrs.)

Multiplexing: FDM, TDM, WDM, Synchronous & Statistical TDM, North American digital multiplexing hierarchy, European TDM, Spread spectrum: Frequency Hopping & Direct Sequence spread spectrum. Terminal handling & polling. Switched Communication Networks: Circuit, Message, Packet & Hybrid Switching, Soft switch Architecture with their comparative study, X.25, ISDN.

Module 3: (08Hrs.)

Physical Layer: Introduction, Interface, Standards, EIA-232-D, RJ-45, RJ-11, BNC connector & EIA-449 digital Interface: Connection, specifications & configuration, X.21 Modem: Types, features, signal constellation, block schematic, limited distance, dial up, baseband, line driver, Group Band and Null modems etc., ITU-T V-series modem standards Connecting Devices: Active and Passive Hubs, Repeaters, Bridges, Two & Three layer switches & Gateway. Study of various types of topology and their comparative study Design of physical LAN Network using routers switches.

Module 4: (06Hrs.)

Transmission Media: Transmission line characteristics, Guided Media: Unguided media, Telephone Network, Digital Subscriber Line: ADSL, HDSL, SDSL, VDSL,Cable TV network for data transfer.

Module 5: (10Hrs.)

Transmission Errors: Content Error , flow integrity error , methods of error control ,Error detection,Error correction ,Bit error rate , Error detection methods: Parity checking , Checksum Error Detection, Cyclic Redundancy Check, Hamming code, Interleaved codes, Block Parity, Convolution code, Hardware Implementation, Checksum .

Course Outcomes:

Students earned credits will develop ability to

1. Illustrate the different modes of data transmission, encoding techniques.
2. Illustrate the different types of multiplexing technique and switching techniques.
3. Illustrate the interfacing and connecting devices and standards used in communication.
4. Summarize the Different types of media of transmission and networks.
5. Analyze the problem of errors in communication and technique of error detection and corrections in transmission.

List of Text/Reference Book:

1. Behrouz A Forouzan, "Data communication and networking", 4th edition, McGrawHill Education, 2017.
2. Tanenbaum A. S., "Computer Networks", Pearson Education, 5th edition, 2011.
3. William Stallings, "Data & Computer Communication", Pearson Education, 8th edition, 2006.
4. Comer, "Internetworking with TCP/ IP Vol-1", Pearson education, 6th edition, 2015.

PCC-DS02	Data Structure & Algorithm	2L:1T:0P (3hrs.)	Credits: 03

Course Objective:

The objective of this course is to understand different types of data structures and algorithms used in programming**.**

Prerequisite: C/C++ Language

Course Contents: **(48 hrs.)**

Module 1: (12hrs.)

Introduction to Data Structure: Classification of Data structures, Abstract Data Types, Common operations Algorithms: Designing algorithms, Analyzing algorithms, Asymptotic Notations.

Introduction to linear data structures- Arrays: Its characteristics, Representation in memory, Types (One-Dimensional Arrays and Multi-Dimensional Arrays), Common operations with cost estimation and Application.

Linked List**:** Its characteristics, Representation of linked list in memory, Types (Singly Linked List, Circular linked list and Doubly linked list), Common operations with cost estimation and Application.

Module 2: (08hrs.)

Stacks: Stacks as ADT, Different implementation of stack, Multiple stacks. Application of Stack: Conversion of infix to postfix notation using stack, Evaluation of postfix expression, Recursion.

Queues: Queues as ADT, Different implementation of queue, Circular queue, Concept of Dqueue and Priority Queue, Queue simulation, Application of queues.

Module 3: (10hrs.)

Tree: Definitions, Height, Depth, Order, Degree, Various types etc. Binary Search Tree – Various Operations i.e. Insertion, Deletion, Traversal, Searching, etc, AVL Tree, Heap, Introduction to B tree, B+ tree.

Module4: (08hrs.)

Graphs: Introduction, Classification of graph: Directed and Undirected graphs etc, Representation, Graph Traversal: Depth First Search (DFS), Breadth First Search (BFS), Graph algorithm: Minimum Spanning Tree (MST) - Kruskal, Prim’s algorithms. Dijkstra’s shortest path algorithm. Application of graphs.

Module5: (10hrs.)

Sorting: Introduction, Classification of sorting method, Sort methods like: Bubble Sort, Quick sort. Selection sort, Heap sort, Insertion sort, Merge sort and Radix Sort; Comparison of various sorting techniques.Searching: Sequential search, Binary search Hashing: Hash Function, Collision, Collision Resolution Technique

Course Outcome:

1. Understand the concept of algorithmic complexity and basic data structures such as arrays, linked lists.
2. Introduce the concept of data structures through ADT including Stack, Queues.
3. Describe the basic concepts of Tree data structure, its types and various operations
4. Understand the concept of Graph data structure, its various representation, operations and Minimum Spanning Tree (MST)
5. Apply various algorithms for solving problems like sorting, searching and hashing

List of Text / Reference Books:

1. Ellis Horowitz, Sartaj Sahni, “Fundamentals of Data Structures” Computer Science Press.

2. Mark Allen Weiss “Algorithms, Data Structures, and Problem Solving with C++” , Pearson Education (US) 1996.

3. AM Tanenbaum, Y Langsam& MJ Augustein, “Data structure using C and C++”, 2nd Ed., 2006 , Prentice Hall India.

4. Robert Kruse, Bruse Leung, “Data structures & Program Design in C”, 2nd Ed., 1997,Pearson Education.

5. Aho, Hopcroft, Ullman, “Data Structures and Algorithms”, Pearson Education.

6. Coremen Thomas, Leiserson CE, Rivest RL, “Introduction to Algorithms” 3rd Ed, 2009 , PHI.

7. Horowitz & Sahani, “Analysis & Design of Algorithm” Computer Science Press

8. Ullmann, “Design & Analysis of Computer Algorithms” Pearson.

9. NPTEL Course Link: https://nptel.ac.in/courses/106/102/106102064/ 10 NPTEL Course Link:https://nptel.ac.in/courses/106/105/106105157//

PCC-DS03	Computer System Organization	2L : 1T:0P (3 hrs.)	Credits: 03

Prerequisite:

Course Objective:

Students to be familiarize the basic principles of computer architecture, Design and Multi Processing, Types of data transfer, Concept of semi conductor memories which is useful for research work in field Computer System.

Course Contents: (42 hrs.)

Module 1: (10hrs.)

Basic Structure of Computer: Structure of Desktop Computers, CPU: General Register Organization-Memory Register, Instruction Register, Control Word, Stack Organization, Instruction Format, ALU, I/O System, bus, CPU and Memory Program Counter, Bus Structure, Register Transfer Language-Bus and Memory Transfer, addressing modes. Control Unit Organization: Basic Concept of Instruction, Instruction Types, Micro Instruction Formats, Fetch and Execution cycle, Hardwired control unit, Micro- programmed Control unit micro program sequencer Control Memory, Sequencing and Execution of Micro Instruction.

Module 2: (8hrs.)

Computer Arithmetic: Addition and Subtraction, Tools Compliment Representation, Signed Addition and Subtraction, Multiplication and division, Booths Algorithm, Division Operation, Floating Point Arithmetic Operation, Number concept 1's and 2's complement representation, addition and subtraction using 2's complement.

Module 3: (8hrs.)

I/O Organization: I/O Interface–PCI Bus, SCSI Bus, USB, Data Transfer: Serial, Parallel, Synchronous, Asynchronous Modes of Data Transfer, Direct Memory Access (DMA), I/O Processor.

Module 4: (08hrs.)

Memory Organization: Main memory-RAM, ROM, Secondary Memory –Magnetic Tape, Disk, Optical Storage, Cache Memory: Cache Structure and Design, Mapping Scheme, Replacement Algorithm, Improving Cache Performance, Virtual Memory, memory management hardware.

Module 5: (08hrs.)

Multiprocessors: Characteristics of Multiprocessor, Inter-Processor Communication and Synchronization. Memory in Multiprocessor System, Concept of Pipelining, Vector Processing, Array Processing, RISC And CISC, Study of Multicore Processor –Intel, AMD.

Course Outcome:

1. Explain the basic structure & components of the computer system, Microprogrammed Control Unit.
2. Demonstrate the concepts of computer arithmetic.
3. Explain the input output organization of the computer system.
4. Illustrate memory organization and memory management techniques.
5. State the core concepts of multiprocessor and pipelining.

List of Text / Reference Books:

1. Morris Mano , "Computer System Architecture " 3rd Ed., 2007, PHI
2. Alan Clements: "Computer Organization and Architecture", 2012, Cengage Learning
3. Subrata Ghosal: "Computer Architecture and Organization", 2011, Pearson Education
4. William stalling ,"Computer Organization and Architecture" 10th Ed., 2016, Pearson Education
5. M. Usha, T.S. Shrikant: "Computer System Architecture & Organization", 2019, Willey India
6. Chaudhuri, P.Pal: "Computer Organization and Design", 3rd Ed. PHI

PCC-DS04	Foundation of Data Science	2L :1T:0P (3 hrs.)	Credits:03

Prerequisite: Engineering Mathematics.

Course Objective: This course provides a concise introduction to the fundamental concepts of Data Science

Course Contents :(48 hrs)

Module1: (10hrs.)

Introduction: What is Data Science? Big Data and Data Science – Datafication - Current landscape of perspectives - Skill sets needed; Matrices - Matrices to represent relations between data, and necessary linear algebraic operations on matrices -Approximately representing matrices by decompositions (SVD and PCA); Statistics: Descriptive Statistics: distributions and probability - Statistical Inference: Populations and samples - Statistical modeling - probability distributions - fitting a model - Hypothesis Testing

Module 2: (08hrs.)

Data preprocessing: Data cleaning - data integration - Data Reduction Data Transformation and Data Discretization. Evaluation of classification methods – Confusion matrix, Students T-tests and ROC curves- Exploratory Data Analysis - Basic tools (plots, graphs and summary statistics) of EDA, Philosophy of EDA. The Data Science Process.

Module 3: (10hrs.)

Introduction to Machine Learning Concepts: Association Rule mining - Linear Regression Logistic Regression- Classifiers - k-Nearest Neighbors (k-NN), k-means -Decision tree - Naive Bayes- Ensemble Methods -Random Forest. Feature Generation and Feature Selection - Feature Selection algorithms

Module 4: (08hrs.)

Clustering: Choosing distance metrics - Different clustering approaches - hierarchical agglomerative clustering, k-means (Lloyd’s algorithm), - DBSCAN - Relative merits of each method - clustering tendency and quality.

Module 5: (06hrs.)

Case Studies/Projects related to data science, Exploratory Data Analysis using Python (NumPy, Pandas, Matplotlib, Seaborn), Interactive Dashboard Creation using Tableau, Data Visualization using BITools.

Course Outcome:

1. State the overview of the Data Science.
2. Explain the process of Data preprocessing.
3. Discuss the various Machine Learning Algorithms.
4. Explain the clustering techniques.
5. Discuss the various Data Science Tool**.**

List of Text / Reference Books:

1. Cathy O'Neil and Rachel Schutt, " Doing Data Science, Straight Talk From The Frontline", O'Reilly, 2014.

2. Jiawei Han, Micheline Kamber and Jian Pei, " Data Mining: Concepts and Techniques", Third Edition. ISBN 0123814790, 2011.

3. Mohammed J. Zaki and Wagner Miera Jr, "Data Mining and Analysis: Fundamental Concepts and Algorithms", Cambridge University Press, 2014.

4. Matt Harrison, "Learning the Pandas Library: Python Tools for Data Munging, Analysis, and Visualization , O'Reilly, 2016.

5. Joel Grus, "Data Science from Scratch: First Principles with Python", O'Reilly Media, 2015.

6. Wes McKinney, "Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython", O'Reilly Media, 2012.

7. NPTEL Course Link: https://nptel.ac.in/courses/106/106/106106212/

8. NPTEL Course Link: https://nptel.ac.in/courses/106/106/1061061

LC-DS02(P)	Data Structure and Algorithm Lab	0L:0T:4P (4 hrs)	Credits: 02

Prerequisite: Knowledge of programming language like C/C++

Course Objective:

The objective of this course is to understand different types of data structures and algorithms used in program.

Course Contents:

Module 1:
Introduction to linear data structures- Arrays, String, representation & Operations.Linked List: Representation of linked list in memory, different implementation of linked list. Circular linked list, doubly linked list, etc.

Module 2:
Stacks: Stacks as ADT, Different implementation of stack, multiple stacks. Application of Stack: Conversion of infix to postfix notation using stack, evaluation of postfix expression, Recursion.

Module 3:
Queues: Queues as ADT, Different implementation of queue, Circular queue, Concept of Dequeue and Priority Queue, Application of queues.

Module 4:
Tree: Definitions - Height, depth, order, degree etc. Binary Search Tree - Operations, Traversal, Searching.
Graphs: Introduction, Classification of graph: Directed and Undirected graphs, etc, Representation, Graph Traversal: Depth First Search (DFS), Breadth First Search (BFS)

Module 5:
Sorting: Introduction, Classification of sorting method, Sort methods like: Bubble Sort, Quick sort. Selection sort, Heap sort, Insertion sort Searching: Basic Search Techniques: Sequential search, Binary search, Comparison of search methods.

Course Outcome:

1. Understand basic data structures such as arrays, linked lists.
2. Introduce the concept of data structures through ADT including stack.
3. Understand the basic operations of Queues.
4. Understand the basic concept of Tree and Graph and their operations.
5. Demonstrate and implement searching and sorting algorithms.

List of Text / Reference Books:

1. Ellis Horowitz, Sartaj Sahni, "Fundamentals of Data Structures" Computer Science Press.
2. Mark Allen Weiss "Algorithms, Data Structures, and Problem Solving with C++" , Pearson Education (US) 1996
3. R. G. Dromey "How to Solve it by Computer", 2nd Impression by, PHI
4. AM Tanenbaum, Y Langsam& MJ Augustein, "Data structure using C and C++", 2nd Ed., 2006 , Prentice Hall India.
5. Robert Kruse, Bruse Leung, "Data structures & Program Design in C", 2nd Ed., 1997, Pearson Education.
6. Aho, Hopcroft, Ullman, "Data Structures and Algorithms", Pearson Education.
7. Richard, GilbergBehrouz, Forouzan ,"Data structure – A Pseudocode Approach with C", 2nd Ed., Thomson press.

List of Experiments:

1. To perform insertion and deletion operations on array. (CO1)
2. To perform multiplication operation on matrix. (CO1)
3. To implement single linked list. (CO1)
4. To calculate factorial of number using recursion. (CO2)
5. To demonstrate implementation of Stack. (CO2)
6. To demonstrate implementation of Linear Queue. (CO3)
7. To implement Binary Search Tree. (CO4)
8. To perform BFS and DFS operations on graph. (CO4)
9. To perform Binary Search Operation. (CO5)
10. To perform sorting operation using Bubble Sort. (CO5)
11. To perform sorting operation using Insertion Sort. (CO5)

LC-DS04	Foundation of Data Science Lab	0L:0T:2P (2Hrs)	Credits: 01

Prerequisite:

Course Objective:

The objective of this course is to understand different types of data structures and algorithms used in program.

Course Contents:

Module 1:

Introduction Introduction to Data Science – Evolution of Data Science – Data Science Roles – Stages in a Data Science Project – Applications of Data Science in various fields – Data Security Issues.

Module 2:

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

Module 3:

Exploratory Data Analytics Descriptive Statistics – Mean, Standard Deviation, Skewness and Kurtosis – Box Plots – Pivot Table – Heat Map – Correlation Statistics – ANOVA.

Module 4:

Model Development Simple and Multiple Regression – Model Evaluation using Visualization – Residual Plot – Distribution Plot – Polynomial Regression and Pipelines – Measures for In-sample Evaluation – Prediction and Decision Making.

Module 5:

Model Evaluation Generalization Error – Out-of-Sample Evaluation Metrics – Cross Validation – Overfitting – Under Fitting and Model Selection – Prediction by using Ridge Regression – Testing Multiple Parameters by using Grid Search.

Course Outcome:

1. Understand the basic concept of Dataset.
2. Understand the process of Data preprocessing.
3. Demonstrate and implement the various Machine Learning Algorithms.
4. Understand the basic concept of the clustering techniques.
5. Implement the Dataset & various Algorithm Data Science Tool**.**

List of Text / Reference Books:

1. JojoMoolayil, "Smarter Decisions: The Intersection of IoT and Data Science",PACKT, 2016.
2. Cathy O'Neil and Rachel Schutt, "Doing Data Science", O'Reilly, 2015.
3. David Dietrich, Barry Heller, Beibei Yang, "Data Science and Big data Analytics", EMC 2013
4. Raj, Pethuru, "Handbook of Research on Cloud Infrastructures for Big DataAnalytics", IGI Global.

List of Experiments:

1. **READING AND WRITING DIFFERENT TYPES OF DATASETS USING PYTHON**

- Reading different types of data sets (.txt, .csv) from web and disk and writing in file in specific disk location.
- Reading Excel data sheet in python.
- Reading XML dataset in python.

2. VISUALIZATIONS:

- Find the data distributions using box and scatter plot.
- Find the outliers using plot.
- Plot the histogram, bar chart and pie chart on sample data

3. EXPLORATORY DATA ANALYSIS (EDA):

- Perform EDA on Credit Card Fraud Detection Dataset (open source dataset) for analyzing the data.

4. LINEAR REGRESSION MODEL FOR PREDICTION:

- Apply Regression Model techniques to predict the future values of data on the open source available datasets.

5. LOGISTIC REGRESSION MODEL:

- Import the Red-Wine dataset from the UCI Machine Learning Repository having three qualities of wines. Apply logistic regression model for multi-class classification of the wine categories.

6. MODEL EVALUATION USING RESIDUAL PLOT:

- Plotting Accuracy and Error Metrics against number of iterations for evaluation of model performance.

7. EVALUATING UNDER-FITTING AND OVER-FITTING:

- Plotting Learning curves for model evaluation for Under-fitting and Over-fitting

8. APPLY PRE-PROCESSING TECHNIQUES ON IPL:

- Dataset using python with various your own desired operations by Collab online tools.
- **APPLY PRE-PROCSSSING TECHNIQUES ON WALMART SALES FORECASTING:** Dataset using python with operation use Kaggle Online tool

9. APPLY VARIOUS OPERATIONS ON TWITTER:

- Dataset using python with using clustering algorithm.

SBC-DS01	Programming in PYTHON	0L:0T:4P (4Hrs)	Credits: 02

Prerequisite: Nil

Course Objective: The course is designed to provide Basic knowledge of Python. Python programming is intended for software engineers, system analysts, program managers and user support personnel who wish to learn the Python programming language. Learning Outcomes: Problem solving and programming capability

Module1:
Introduction, History, Features, Python –Environment Setup Local Environment Setup, Getting Python, Installation of Python, Use of IDE

Module 2:
Python –Basic Syntax Python Identifiers, Reserved Words, Lines & Indentation, Multiline Statements, Quotation in Python, Comments & other useful constructs, Python –Variables Assigning Values to Variables, Multiple Assignment, Standard Data Types

Module3:
Python –Variables, Assigning Values to Variables, Multiple Assignment, Standard Data Types; Python Numbers, Python Strings, Python Lists, Python Tuples, Dictionary, Data Type Conversion

Module 4:
Python –Basic Operators, Types of Operators, Arithmetic Operators, Comparison Operators, Assignment Operators, Bitwise Operators, Logical Operators, Operator Precedence, Python – Decision Making & Loops, Flowchart, If statement Syntax

Module 5:
Python-Functions, Syntax for defining a function, Calling a Function, Function Arguments, Anonymous Functions Python-Applications & Further Extensions, Data analysis packages.

Course Outcome:

1. Install Python and have knowledge of syntax of Python.

2. Describe the Numbers, Math functions, Strings, List, Tuples and Dictionaries in Python.

3.Express different Decision Making statements and Functions.

4. Develop code in Python using functions, loops etc.

5. Design GUI Applications in Python and evaluate different database operations.

List of Text Books / Reference Books:

1. Eric Matthes, "Python Crash Course: A Hands-On, Project-Based Introduction to Programming", No Starch Press.
2. ZedA. Shaw, "Learn Python the Hard Way" (3rdEdition), Addison Wesley.
3. Paul Barry, "Head-First Python", O"Reilly. 4. John Zelle, Franklin,"Python Programming", Beedle & Associates Inc.

List of Experiments:

2. To write a Python program to print Hello. Also, find area of a triangle.
3. To write a program to find the square root and exponentiation of a number.
4. To write a Python program to find if a number is even or odd.
   To write a Python program to find middle number in a group of three numbers
5. To write a Python Program to find the maximum from a list of numbers.
6. To write a Python Program to find prime numbers in an interval.
7. To write a Python Program to Display Multiplication table of a number.
8. To write a Python Program to find Sum of natural numbers using recursive function.
9. Python Program to Make a Simple Calculator.
10. Write a Python function that takes two lists and returns True if they are equal otherwise false.
11. Write a python program to display a particular month of a year using calendar module.
12. To demonstrate data analysis packages using python like Pandas, Filtering.
13. To demonstrate data analysis packages using python like math.libs, numpy.