Curriculum

for

One Year (Two Semester)

Post Graduate Diploma Course In

DATA SCIENCE AND MACHINE LEARNING

Prepared by:
Institute of Research Development and Training
Vikas Nager, Kanpur

APPROVED IN CDC MEETING OF BOARD OF TECHNICAL EDUCATION,LUCKNOW,U.P DATED 26-05-22

CONTENTS

APPROVED IN CDC MEETING OF BOARD OF TECHNICAL EDUCATION,LUCKNOW,U.P DATED 26-05-22

FIRST SEMESTER

SECOND SEMESTER

APPROVED IN CDC MEETING OF BOARD OF TECHNICAL EDUCATION,LUCKNOW,U.P DATED 26-05-22

PREFACE

Polytechnics play an important role in meeting the requirements of trained technical manpower for industries and field organizations. The initiatives being taken by the State Board of Technical Education, UP and IRDT Kanpur to design and implement the curricula of 4 new age diploma programmes as per the needs of the industry and to equipped the students with strong analytical and programming skills which makes them highly demanding and employable on completion of the program.

Data Science refers to extraction of knowledge from large volumes of data that are structured or unstructured, which is continuation of data mining and predictive analytics. It involves different categories of analytical approaches for modeling various types of business scenarios and arriving at solution and strategies for optimal decision-making in marketing, finance, operations, organizational behavior and other managerial aspects. This new field of study breaks down into a number of different areas, from constructing big data infrastructure and configuring the various server tools that sit on top of the hardware, to performing the analysis and developing the right transformations to generate useful results.

The objective of this program is to make Data Scientists and Data Analysts. There is a huge demand for resources skilled in Data Science. The MCKINSEY Global Institute has predicted that in forthcoming years by 2020, the world will face a shortage of 38-40 million highly skilled manpower with deep analytical skills that can leverage data analysis to make effective decisions for their organizations. So it is quite obvious that existing resources along with new candidates who are interested in perusing career in this field needs to be trained. Our objective is to create a pool of talent who can meet this demand. This course is meant to sensitize students for computational statistics applications and usage as well as provide hands-on experience with solving real world data science issues.

On completion of the Course, the Participants will learn the concept of Data Analytics using open source statistical tools like R, Python and some very good visualization tools and techniques. They will be able to implement industry oriented Data Analytics Project.

In order to meet the requirements of future technical manpower, we will have to revamp our existing technical education system and one of the most important requirements is to develop outcome-based curricula of diploma programmes. The curricula for diploma programmes have been revised by adopting time-tested and nationally acclaimed scientific method, laying emphasis on the identification of learning outcomes of diploma programme.

The real success of the diploma programme depends upon its effective implementation. However best the curriculum document is designed, if that is not implemented properly, the output will not be as expected. In addition to acquisition of appropriate physical resources, the availability of motivated, competent and qualified faculty is essential for effective implementation of the curricula.

It is expected of the polytechnics to carry out job market research on a continuous basis to identify the new skill requirements, reduce or remove outdated and redundant courses, develop innovative methods of course offering and thereby infuse the much needed dynamism in the system.

Manoj Kumar
Director
Institute of Research Development & Training
Kanpur, U.P.

ACKNOWLEDGEMENTS

We gratefully acknowledge the guidance and contribution received from the following persons:

i) Principal Secretary, Technical Education & Vocational Education, Govt. of UP for his exemplary vision and approach.

ii) Special Secretary, Technical Education Department, Govt. of UP for his support and motivation.

iii) Sh. Manoj Kumar, Director, Technical Education, UP and Director, I.R.D.T., Kanpur for continuously motivating, guiding and taking keen interest in the review of this curriculum.

iv) All the participants from industry/field organizations, engineering colleges, polytechnics and other technical institutions for their professional inputs during curriculum workshops.

v) Faculty/Subject Experts from different U.P. Government polytechnics for content updating.

vi) Concerning officers of BTE Lucknow / IRDT Kanpur for their support and assistance in the conduct of Curriculum workshops at different places.

Coordinator
Institute of Research Development & Training,
Kanpur, U.P.

1. SALIENT FEATURES OF POST GRADUATE DIPLOMA PROGRAMME IN DATA SCIENCE & MACHINE LEARNING

1)	Name of the Programme	:	Post Graduate Diploma Programme in Data Science and Machine Learning
2)	Duration of the Programme	:	One year (Two Semesters)
3)	Entry Qualification	:	Graduation or equivalent NSQF Level as Prescribed by State Board of Technical Education, UP
4)	Intake	:	60 (or as prescribed by the Board)
5)	Pattern of the Programme	:	Semester Pattern
6)	NSQF Level	:	Level - 8
7)	Ratio between theory and practical	:	45 : 55 (Approx.)

8) Seminar And Case Studies :

The PG Diploma students are to be instructed to prepare the presentation of current research work of the concerned field as suggested by the mentor/supervisor/faculty member. Different parameters could be taken into account in determining the effectiveness of the research presentation. Before the presentation, the topic/content has to be approved by the mentor/supervisor/faculty member.

12) Student Centered Activities:

A provision of 3-6 hrs per week has been made for organizing Student Centred Activities for overall personality development of students. Such activities will comprise of co–curricular activities such as expert lectures, self study, games, hobby classes like photography, painting, singing etc. seminars, declamation contests, educational field visits, NCC, NSS and other cultural activities, disaster management and safety etc.

13) Project work

A project work has been included in the curriculum to enable the student get familiarize with the practices and procedures being followed in the industries and provide an opportunity to work on some live projects in the industry.

Key Features

- Designed for professionals, Science and Engineering Graduates
- Industry driven case studies prepared by esteemed industry experts
- hours of interactive sessions by industry experts
- Overview of the latest technologies used in AI systems
- Industry oriented course curriculum.
- Realization of the concepts through hands on assignments and projects

2. EMPLOYMENT OPPORTUNITIES

Job Roles In Data Science And Machine Learning

- Data Analyst
- Data Engineers
- Database Administrator
- Machine Learning Engineer
- Data Scientist
- Data Architect
- Statistician
- Business Analyst
- Data and Analytics Manager

Many Industrial Sectors rely on Machine Learning.

3. LEARNING OUTCOMES OF POST GRADUATION DIPLOMA IN DATA SCIENCE AND MACHINE LEARNING

SNo	Learning Outcomes
After undergoing this program , Students will be able to	
1	Understand the basic data science concepts and methods to solve problems in real-world contexts
2	Understand the basic data science concepts related to Artificial Intelligence
3	Understand the basic data science concepts related to Machine Learning
4	Learn the concepts of Linear Algebra, probability and Statistics to apply in Data Science and Machine learning
5	Learn the basic programming skills to think the computational concept in Python
6	To learn about language components of Python program
7	Understand the sorting and searching algorithms of Artificial intelligence
8	Knowledge of Git Repository to manage the code
9	Understanding of Excel/Spreadsheet for Data Analytics
10	Develop the knowledge of professional communication skills
11	Relate the real world case studies with concepts of Data Science and Machine Learning
12	Demonstrate skill and knowledge of tools and techniques specific to the professional field of study.
13	Perform mathematical operations by using NumPy
14	Apply Data Cleaning/Filtering operations using Pandas
15	visualize and draw different plots using Seaborn and matplotlib
16	Understand the basic data science concepts and methods of EDA
17	Identify obvious errors, better understand patterns within the data,
18	Understand the fundamental concepts of machine learning algorithms
19	Apply a mathematical approach to solve various Machine Learning and Data science-based problem
20	Develop various models involved in learning from data.
21	Solve real life problems by application of acquired knowledge and skills
22	Prepare details project proposal and report
23	Plan and execute given tasks/project as a team member or leader
24	Manage resources effectively at the workplace

4. DERIVING CURRICULUM AREAS FROM LEARNING OUTCOMES OF THE PROGRAMME

SNo	Learning Outcomes	Curriculum Areas /Subjects
The Following curriculum area subjects have been derived from Learning Outcomes		
1	Understand the basic data science concepts and methods to solve problems in real-world contexts	FUNDAMENTALS OF DATA SCIENCE AND MACHINE LEARNING
2	Understand the basic data science concepts related to Artificial Intelligence	
3	Understand the basic data science concepts related to Machine Learning	
4	Learn the concepts of Linear Algebra, probability and Statistics to apply in Data Science and Machine learning	ESSENTIAL MATHEMATICS FOR DATA SCIENCE AND MACHINE LEARNING
5	Learn the basic programming skills to think the computational concept in Python	PYTHON FOR DATA SCIENCE
6	To learn about language components of Python program	
7	Understand the sorting and searching algorithms of Artificial intelligence	
8	Knowledge of Git Repository to manage the code	DATA SCIENCE - TOOLS AND TECHNIQUES
9	Understanding of Excel/Spreadsheet for Data Analytics	
10	Develop the knowledge of professional communication skills	
11	Relate the real world case studies with concepts of Data Science and Machine Learning	SEMINAR AND CASE STUDIES
12	Demonstrate skill and knowledge of tools and techniques specific to the professional field of study.	
13	Perform mathematical operations by using NumPy	ADVANCED PROGRAMMING FOR DATA SCIENCE
14	Apply Data Cleaning/Filtering operations using Pandas	
15	visualize and draw different plots using Seaborn and matplotlib	
16	Understand the basic data science concepts and methods of EDA	EXPLORATORY DATA ANALYSIS AND VISUALIZATION METHODS
17	Identify obvious errors, better understand patterns within the	

	data,	
18	Understand the fundamental concepts of machine learning algorithms	MACHINE LEARNING ALGORITHMS
19	Apply a mathematical approach to solve various Machine Learning and Data science-based problem	
20	Develop various models involved in learning from data.	
21	Solve real life problems by application of acquired knowledge and skills	PROJECT WORK
22	Prepare details project proposal and report	
23	Plan and execute given tasks/project as a team member or leader	
24	Manage resources effectively at the workplace	

5. ABSTRACT OF CURRICULUM AREAS

a) Applied Mathematics

Essential Mathematics for Data Science and Machine Learning

b) Applied Course in Engineering/Technology

Fundamentals of Data Science & Machine Learning
Python for Data Science
Data Science - Tools And Techniques
Seminar and Case Studies
Advanced Programming For Data Science
Exploratory Data Analysis & Visualization Methods
Machine Learning Algorithms
Project work

6. HORIZONTAL AND VERTICAL ORGANIZATION OF THE SUBJECTS

S No	Subjects	Distribution in Periods per week in Various Semesters	
		I	II
1.	Fundamentals of Data Science & Machine Learning	6	-
2.	Essential Mathematics for Data Science and Machine Learning	8	-
3.	Python for Data Science	14	-
4.	Data Science Tools and Techniques	6	-
5.	Seminar and Case Studies	6	-
6.	Advanced Programming For Data Science	-	12
7	Exploratory Data Analysis & Visualization Methods	-	12
8	Machine Learning Algorithms	-	12
9	Project Work	-	8
10	Student Centered Activities	2	2

STUDY AND EVALUATION SCHEME FOR
Post Graduate Diploma Course In
DATA SCIENCE AND MACHINE LEARNING

FIRST SEMESTER:

Sr. No.	SUBJECTS	STUDY SCHEME Periods/Week			Credits	MARKS IN EVALUATION SCHEME								Total Marks of Internal & External
						INTERNAL ASSESSMENT			EXTERNAL ASSESSMENT					
		L	T	P		Th	Pr	Tot	Th	Hrs	Pr	Hrs	Tot	
1.1	Fundamentals of Data Science & Machine Learning	6	-	-	5	20	-	20	50	2 ½	-	-	50	70
1.2	Essential Mathematics for Data Science and Machine Learning	8	-	-	6	20	-	20	50	2 ½	-	-	50	70
1.3	Python for Data Science	6	-	8	8	20	20	40	50	2 ½	50	3	100	140
1.4	Data Science - Tools and Techniques	2	-	4	4	-	20	20	-	-	50	3	50	70
1.5	Seminar and Case Studies	-	-	6	3	-	20	20	-	-	-	-	-	20
1.6	#Student Centered Activities	-	-	2	1	-	30	30	-	-	-	-	-	30
TOTAL		**22**	**-**	**20**	**27**	**60**	**90**	**150**	**150**	**-**	**100**	**-**	**250**	**400**

Student Centered Activities will consist of co-curricular activities like extension lectures, games, hobby clubs e.g. photography etc., seminars, declamation contests, educational field visits, N.C.C., NSS, Cultural Activities and self study etc.

SECOND SEMESTER:

Sr. No.	SUBJECTS	STUDY SCHEME Periods/Week			Credits	MARKS IN EVALUATION SCHEME								Total Marks of Internal & External
						INTERNAL ASSESSMENT			EXTERNAL ASSESSMENT					
		L	T	P		Th	Pr	Tot	Th	Hrs	Pr	Hrs	Tot	
2.1	Advanced Programming For Data Science	6	-	6	8	20	20	40	50	2 ½	50	3	100	140
2.2	Exploratory Data Analysis & Visualization Methods	6	-	6	8	20	20	40	50	2 ½	50	3	100	140
2.3	Machine Learning Algorithms	6	-	6	8	20	20	40	50	2 ½	50	3	100	140
2.4	Project work	-	-	8	4	-	50	50	-	-	100	3	100	150
2.5	#Student Centered Activities	-	-	2	1	-	30	30	-	-	-	-	-	30
TOTAL		**18**	-	**28**	**29**	**60**	**140**	**200**	**150**		**250**		**400**	**600**

Student Centred Activities will comprise of co-curricular activities like extension lectures, games, hobby clubs e.g. photography etc., seminars, declamation contests, educational field visits, N.C.C., NSS, Cultural Activities and self study etc.

8. GUIDELINES FOR ASSESSMENT OF STUDENT CENTERED ACTIVITIES (SCA)

It was discussed and decided that the maximum marks for SCA should be 30 as it involves a lot of subjectivity in the evaluation. The marks may be distributed as follows:

i. 10 Marks for general behavior and discipline

(by HODs in consultation with all the teachers of the department)

ii. 5 Marks for attendance as per following:

(by HODs in consultation with all the teachers of the department)

a)	75-80%	2 Marks
b)	80 - 85%	4 Marks
c)	Above 85%	5 Marks

iii. 15 Marks maximum for Seminars/Sports/NCC/Cultural/Co-curricular/ NSS activities as per following:
(by In-charge Sports/NCC/Cultural/Co-curricular/NSS)\

a)	15 -	State/National Level participation
b)	10 -	Participation in two of above activities
c)	5 -	Inter-Polytechnic level participation

Note: There should be no marks for attendance in the internal sessional of different subjects.

1.1 FUNDAMENTALS OF DATA SCIENCE AND MACHINE LEARNING

L T P
6 - -

RATIONALE

In today's world, data is becoming so vast. Every Company requires data to work, grow, and improve their businesses. Now, handling such huge amount of data is a challenging task for every organization. So to handle, process, analyze and apply this data we required some complex, powerful, and efficient algorithms and technology, and that technology came into existence as Data Science and Machine Learning.

LEARNING OUTCOMES

After completion of this course, the students will be able to:

- Understand the basic data science concepts and methods to solve problems in real-world contexts.
- Asking the correct questions and analyzing the raw data.
- Visualizing the data to get a better perspective.
- Understanding the processing of data to make decisions and finding the result.
- Understand the basic data science concepts related to Artificial Intelligence.
- Understand the basic data science concepts related to Machine Learning.

.

DETAILED CONTENTS

1. INTRODUCTION TO DATA AND DATA SCIENCE (20 Periods)

Data definition, various forms of data, categories of data, data quality and issues, data architecture.

Definition, Need and importance of Data science. Applications and job opportunities in data science. What do data scientists do?

Data science life cycle. Components of data science.

2. INTRODUCTION TO DATA ANALYTICS AND BIG DATA (20 Periods)

Overview of data analysis. Types of data analytics; - descriptive, predictive and perspective, inferential with examples,

Big data overview, characteristics of big data. Characteristics of Big Data sources of Big data, different sectors using Big Data. Requirements and challenges of Big Data.

3. DATA PRE-PROCESSING AND DATA VISUALIZATION (15 Periods)

What is data preprocessing, data preprocessing importance, major task in data preprocessing as data cleaning, data integration, data reduction and data transformation

What is data visualization? Data visualization tools

4. INTRODUCTION TO MACHINE LEARNING (18 Periods)

Machine learning definition, methods And Approaches, Supervised and Unsupervised learning, Decision Tree Learning, Deep learning concept related to data science.

5. INTRODUCTION TO ARTIFICIAL INTELLIGENCE (11 Periods)

Artificial Intelligence and its importance. Need and job roles of AI professional. Goals of Artificial intelligence.Components of AI. Advantages and shortcomings of artificial intelligence. Types of artificial intelligence. Concept of agents in artificial intelligence

INSTRUCTIONAL STRATEGY

The subject is conceptual. Students should be given clear idea about the basic concepts of Data Science and Machine Learning. Student should be asked to explain the algorithm and its applications in Data Science.

MEANS OF ASSESSMENT

- Assignments and quiz/class tests, mid-term and end-term written tests
- Viva-voce

RECOMMENDED BOOKS

1. Data Science and Analytics with Python Programming by V. K. Jain(Author) , Khanna Book Publication.

2. Data Science and Machine Learning Interview Questions Using Python a Complete Question Bank to Crack Your Interview by Narayanan Vishwanathan.

3. Foundations of Artificial Intelligence and Expert Systems by V.S.Jankiraman, K.Sarukesi, P.Gopalakrishnan.

4. Data Science and Machine Learning by N. Meenakshi K. E. Rajakumari S. Hariharasitaraman (Author)

5. The Data Science Handbook: Advice and Insights from 25 Amazing Data Scientists Paperback by Carl Shan (Author), William Chen (Author)

REFERENCE WEBSITES

1. https://javatpoint.com
2. https://www.analyticsvidhya.com
3. https://www.upgrad.com
4. https://iiitl.ac.in

SUGGESTED DISTRIBUTION OF MARKS

Topic No.	Time Allotted (Periods)	Marks Allotted (%)
1.	20	20
2.	20	20
3.	15	20
4.	18	20
5.	11	20
Total	**84**	**100**

1.2 ESSENTIAL MATHEMATICS FOR DATA SCIENCE AND MACHINE LEARNING

L	T	P
8	-	-

RATIONALE

Machine learning (ML) is one of the most popular topics. This particular topic is having applications in all the areas of engineering and sciences. Various tools of machine learning are having a rich mathematical theory.So it is necessary to have knowledge of all such mathematical concepts. In this subject we will learn basic mathematical concepts required for Data Science and Machine learning

LEARNING OUTCOMES

After undergoing the course, the students will be able to:

- To understand Matrix operations and uses of matrix in different problems.
- Apply elementary row and column operations in finding the inverse of a matrix.
- Find Eigenvalues, Eigenvectors of a matrix and their different properties.
- Understand concept of probability distribution and their applications
- Understand the concept of statistics and their application.

DETAILED CONTENTS

1. Linear Algebra (35 Periods)

Vector : Vectors Introduction, Dot and Cross product, Scaler and vector triple product, projection.

Matrix : Introduction to Matrix, Algebra of Matrices, Inverse Addition, Multiplication of matrices, Null matrix and a unit matrix, Square matrix, Symmetric, Skew symmetric, Hermitian, Skew hermitian, Orthogonal, Unitary, diagonal and Triangular matrix. Transpose of a matrix and properties.

Introduction of Determinant and properties, Cramer`s Rule, Determinant of a matrix. Definition and Computation of inverse of a matrix, Trace of a matrix.
Elementary Row/Column Transformation Meaning and use in computing inverse and rank of a matrix, Gauss's Elimination Method
Linear Dependence, Rank of a Matrix Linear dependence/independence of vectors, Definition and computation of rank of matrix. Computing rank through determinants, Elementary row transformation and through the concept of a set of independent vectors, Consistency of equations.

Eigen Pairs, Cayley-Hamilton Theorem Definition and evaluation of eign values and eign vectors of a matrix of order two and three, Cayley-Hamilton theorem (without Proof)and its verification, Use in finding inverse and powers of a matrix.
Introduction to Vector Space- Vector Space, Basis, Subspace, Linear independence and Dimension, Changing basis, Doing a transformation in a changed basis, Changing the eigenbasis, Linear Transformations, Norms and Spaces, Orthogonal Complement and Projection Mapping

2. Probability Theory (25 Periods)

Introduction to probability, Addition and Multiplication theorem and simple problem, Random Variable and Continuous Random Variable, Conditional Probability with Bayes` Theorem, Probability Density Function, Normal Distribution, Central Limit Theorem, Binomial Distribution, Conjugate distribution, Poisson Distribution, Bell Curve and sampling distribution, Discriminant Analysis.

3. Statistics (26 Periods)

Statistics Introduction, Types of Statistics (Descriptive, Prescriptive, Predictive), Types of Statistical Measures : Measures of Frequency, Measures of central Tendency (Mean, Median, Mode), Measures of Spread (Standard Variance and Standard Deviation, Quartiles), Measures of Position , Correlation, Covariance

4. Calculus (26 Periods)

The Fundamental Theorem of Calculus and Integration Products, Quotients, and Chains: Simple Rules for Calculus Differential Calculus, Understand the Definition of a Derivative, How to Find the Derivative of a Function How to Take Partial Derivatives The Jacobian Matrix: Introducing Vector Calculus Lagrange Multipliers: An Introduction to Constrained Optimization

INSTRUCTIONAL STRATEGY

Teachers should lay emphasis on mathematical concepts with its applications , so that students can relate the Data Science domain with Mathematical Concepts. Experts from industries may be invited to deliver lectures and share experiences with the students.

MEANS OF ASSESSMENT

- Assignments and quiz/class tests, mid-term and end-term written tests
- Practice exercises
- Viva-voce

RECOMMENDED BOOKS

1. Introduction to Linear Algebra, Fifth Edition, Gilbert Strang, 2016
2. Linear Algebra and Learning from Data, Gilbert Strang, Wellesley-Cambridge Press,U.S.
3. Mathematics for Machine Learning by Marc Peter Deisenroth Cambridge University Press
4. APPLIED STATISTICS AND PROBABILITY FOR ENGINEERS – BY DOUGLAS MONTGOMERY

5. e-books/e-tools/relevant software to be used as recommended by BTE/NITTTR, Chandigarh.

Websites for Reference:

http://swayam.gov.in

SUGGESTED DISTRIBUTION OF MARKS

Topic No.	Time Allotted (Periods)	Marks Allotted (%)
1.	35	30
2.	25	30
3.	26	20
4.	26	20
Total	**112**	**100**

1.3 PYTHON FOR DATA SCIENCE

L	T	P
6	-	8

RATIONALE

This course introduces to the students the Python language. Upon completion of this course, the student will be able to write non trivial Python programs dealing with a wide variety of subject matter domains. Topics include language components, the IDLE/IDE environment, control flow constructs, strings, I/O, collections, classes, modules, and regular expressions.

LEARNING OUTCOMES

After undergoing the course, the students will be able :

- To understand why Python is a useful scripting language for developers.
- To learn how to execute Python code in a variety of environments
- To learn how to design and program Python applications.
- To learn how to use lists, tuples, and dictionaries in Python programs.
- To learn how to identify Python object types.
- To learn how to use indexing and slicing to access data in Python programs.
- To define the structure and components of a Python program.
- To learn how to write loops and decision statements in Python.
- To learn how to write functions and pass arguments in Python.
- To learn how to build and package Python modules for reusability.
- To learn how to read and write files in Python.
- To learn how to design object-oriented programs with Python classes.
- To learn how to use class inheritance in Python for reusability.
- To learn how to use exception handling in Python applications for error handling.
- To learn how to use the Python Regular Expression capabilities for data verification

DETAILED CONTENTS

1. Introduction to Python (06 Periods)

Brief History, Features, Python Installation using anaconda, Environment Variables, Python Virtual Machine (PVM), Program Execution Steps and Memory Management in Python, First Program Execution using Python`s command line window, using Python`s IDLE graphics window, using system prompt.

2. Basic Language Constructs (10 Periods)

Basic Syntax, Comments, Identifiers and Keywords, Variables, Python Data Types, Python casting, Operators (Arithmetic, Assignment, Comparison, Identity, Relational,Logical, Bit Wise), Operator Precedence and Associativity, Strings, String manipulation, built-in functions (input,print,abs,pow,min, max etc)

Indenting Requirements , Decision Control Instruction (if, else, elif), while loop, for loop and range () function , nested loops, else block of loop, break, continue

3. Functions and Modules (08 Periods)

Introduction, Defining Your Own Functions,Parameters, Function Documentation, Types of arguments, Scope of functions, Recursion, Functions - "First Class Citizens", Passing Functions to a Function, map, filter ,Mapping Functions in a Dictionary, Lambda, Inner Functions, Closures

Modules introduction, built-in modules : sys, math,cmath, random, decimal, date time.

4. Object Oriented Programming (08 Periods)

Introduction to Object Oriented Programming, Procedural vs Object Oriented Programming, Classes and Objects, Creating Classes, Self Variable, Class Variables and Methods, Constructor, Inner Classes, Inheritance, super method, Polymorphism, Overriding and Overloading, Abstract Classes, Interface

5. Collections in Python (16 Periods)

Arrays : Introduction, creation, accessing array elements, built-in methods

Lists : Introduction, creating list using range(), accessing and updating list elements, Looping in Lists, List in a List, Operations on a list: Slicing, Splicing, Sub-setting, Indexing and Slicing, Condition (true/false) on a List, Applying functions on a List, List methods - append, count, extend, List methods - index, insert, pop, remove, reverse, sort List comprehension

Tuples : Introduction, accessing elements, Looping in tuples, Basic Operations on Tuples, built-in functions and methods for tuples

Sets : Introduction, accessing Set element, looping in sets, basic set operations, using built-in functions on sets and set methods, mathematical set operations.

Dictionaries : Introduction, accessing the Dictionary element, looping in Dictionaries, basic Dictionaries operations, using built-in functions on Dictionary and Dictionary methods, Iteration over keys, values and items, viewing keys, values and items, Sorting Dictionaries

Passing Collections to a Function : Unpacking arguments

6. Exception Handling (08 Periods)

Errors, Runtime Errors, The Exception Model, Exception Hierarchy, Exception Handling (try, except, finally, else block), Handling Multiple Exceptions, User Defined Exception , Raise, assert

7. File Handling (10 Periods)

Introduction, Types of Files in Python, Opening a File, Access Modes, Reading Data From a File, Writing Data to a File, Additional File Methods, Closing a File, Handling IO Exceptions

8. Regular Expressions (10 Periods)

Introduction, Simple Character Matches, Special Characters, Character Classes, Quantifiers, The Dot Character, Greedy Matches, Grouping, Matching at Beginning or End, Match Objects, Substituting, Splitting a String, Compiling Regular Expression, Using Regular Expressions on Files.

9. Search and Sort Algorithms (08 Periods)

Sorting : Selection Sort, Bubble Sort, Insertion Sort, Merge Sort, Quick Sort
Searching : Linear Search, Binary Search, Jump Search, Interpolation Search, Recursive program to linearly search an element in a given array

LIST OF PRACTICALS

SNo	Description
Module -1 : Introduction to Python	
1.1	Download and Install Anaconda on Local Machine and verify if Python is working or not along with it`s version
1.2	Write down a Python program to print "Welcome to Python World" and execute the program using Python`s Command Line Window
1.3	Write down a Python program to print "Welcome to Python World" and execute the program using System`s Prompt
1.4	Write down a Python program to print "Welcome to Python World" and execute the program using IDLE Graphics Window
Module - 2 : Basic Language Constructs	
2.1	Write a Python program to declare two variables and print the sum of variables
2.2	Find the maximum among two numbers using if else statement and inbuilt function
2.3	Write down a Python program to Calculate the Simple Interest

2.4	Write down a Python program to calculate the factorial of a number
2.5	Write a Python Program to print all Prime Numbers in a Interval
2.6	Write a Python Program to Swap the values of two variables 1) without using third variable 2) Using third variable 3) Using simple python syntax
2.7	Python Program for Sum of squares of first n natural numbers
2.8	Python program to check if a string is palindrome or not
2.9	Reverse words in a given String in Python
2.10.	To remove i'th character from string in Python
2.11	Extract frequency of all the words in string : "Python is good for all programmers, so all programmers should learn python"
2.12	Given a pair of non-empty strings. str1 = 'abcdef' str2 = 'defghia' Count the number of matching characters in those strings
2.13	Find all duplicate characters in string
2.14	Replace all occurrences of a substring in a string
2.15	Write a Program to check whether a triangle is valid or not, when the three angles of the triangle are entered through the keyboard
2.16	Print first 25 numbers using range ()
2.17	Print the multiplication table of the number entered by the user. The table should get displayed in the following form : 23 * 1 = 23 23 * 2 = 46
2.18	Program to compute the GCD of two numbers
Module 3 : Functions and Module	
3.1	Write a program to receive three integers from the keyboard and get their sum and product calculated through a user-defined function calculate_sum_prod()
3.2	Write a program that defines a function count_lower_upper() that accepts a string and calculates the number of uppercase and lowercase alphabets in it.
3.3	If a positive integer is entered through the keyboard, write a recursive function to obtain the prime factors to the number

3.4	A positive integer is entered through the keyboard, write a recursive function to calculate sum of digits of the the 8 digit number
Module 4 : Object Oriented Programming	
4.1	Write a program to create a class called as Car with attributes such as company_name, model_no, color. Create multiple objects of this class. Report how many objects have been created from this class
4.2	Write a program to create classes that can calculate the perimeter and area of a Circle and Rectangle
Module - 5 Collections in Python	
5.1	Write a program to find sum of array
5.2	Write a program to find the largest element in an array
5.3	Write a program to Split the array and add the first part to the end
5.4	Write a program to Find remainder of array multiplication divided by n
5.5	Write a program to interchange first and last elements in a list
5.6	Write a program to swap two elements in a list
5.7	Write a program to find length of list
5.8	Write a program to check if element exists in list
5.9	Write a program to find sum of elements in list
5.10.	Write a program to find the smallest number in a list
5.11	Write a program to find the largest number in a list
5.12	Write a program to find second largest number and Nth largest number in a list
5.13	Write a program to remove multiple elements from a list in Python
5.14	Write a program to find count occurrences of an element in a list
5.15	Write a program to print duplicates from a list of integers
5.16	Perform the following operations on a list of name - Create a list of 5 names 'Anil', 'Amol', 'Aditya', 'Avi', 'Alka' - Insert the name 'Anuj' before 'Aditya' - Append a name 'Zulu' - Delete 'Avi' from the list - Replace Anil with 'AnilKumar' - Print sorted list - Print Reverse Sorted List
5.17	Write a program to add and multiply two 3 x 3 matrices

5.18	Write a program that implements Matrix class and performs transpose operation on 3 X 3 matrices
5.19	Write a program to Find the size of a Tuple
5.20.	Write a program to find Maximum and Minimum K elements in Tuple
5.21	Write a program for Adding Tuple to List and vice – versa
5.22	Write a program to Order Tuples using external List
5.23	Write a program to create the following 3 lists : - a list of names - a list of roll numbers - a list of marks Generate a print a list of tuples containing name,roll number and marks from the 3 lists. From this list generate 3 tuples - one containing all names, another containing all roll numbers and third containing all marks

5.24	Given the following Dictionary
	marks = {
	'Ram' : {'Maths':82, 'Hindi':81, 'SSt' :95},
	'Krishna' : {'Maths':76, 'Hindi':78, 'SSt' :79},
	'Aditya' : {'Maths':56, 'Hindi':66, 'SSt' :77},
	}
	
	Write a program to perform following operations :
	- Print marks obtained by Aditya in Maths
	- Set Marks obtained by Krishna in Maths to 87
	- Sort the dictionary by name

5.25	Write a program to extract Unique values dictionary values
5.26	Write a program to find the sum of all items in a dictionary
5.27	Ways to remove a key from dictionary
5.28	Ways to sort list of dictionaries by values in Python – Using itemgetter
5.29	Ways to sort list of dictionaries by values in Python – Using lambda function
5.30.	Merging two Dictionaries
5.31	Sort Dictionary key and values List
5.32	Sort Python Dictionaries by Key or Value
5.33	Write a program using lambda to obtain a list that contains elements obtained by adding following two list

	l1=[1,2,3,4,5,6] l2=[6,5,4,3,2,1]
5.34	A list contains names of employees. Write a program to filter out those names whose length is more than 8 characters
Module - 6 : Exception Handling	
6.1	Write a program that infinitely receives a positive integer as input and prints its square. If a negative number is entered then raise an exception, display a relevant error message and make a graceful exit.
6.2	Write a program that receives an integer as input. If a string is entered instead of an integer, then report an error and give another chance to user to enter an integer. Continue this process till correct input is supplied
Module -7 File Handling	
7.1	Write a program to read the contents of file 'messages' one character at a time. Print each character that is read
7.2	Write a program to get the number of characters, words , spaces and lines in a file
7.3	Write a program to eliminate repeated lines from a file
7.4	Write a program to merge two files into a third file
7.5	Append content of one text file to another
7.6	Write a program to obtain the line number in which given word is present
Module - 8 Regular Expressions	
8.1	Write a Program to Check if String Contain Only Defined Characters using Regex
8.2	Write a program to Count Uppercase, Lowercase, special character and numeric values using Regex
8.3	Write a Program to find the most occurring number in a string using Regex
8.4	Write a Regex to extract maximum numeric value from a string
8.5	Write a Program to put spaces between words starting with capital letters using Regex
8.6	Write a regex to find sequences of one upper case letter followed by lower case letters
8.7	Write a Program to Remove duplicate words from Sentence
8.8	Write a Program to Remove all characters except letters and numbers
8.9	Write a Program to Check if email address valid or not
8.10	Write a program to find files having a particular extension using RegEx
8.11	Write a program to check the validity of a Password

8.12	Categorize Password as Strong or Weak using Regex in Python
Module - 9 Search and Sort Algorithms	
9.1	Write a Python program for Insertion Sort
9.2	Write a Python program for Quick Sort
9.3	Write a Python Program for Merge Sort
9.4	Write a Python program for Linear Search
9.5	Write a Python program for Binary Search

INSTRUCTIONAL STRATEGY

Teachers should lay emphasis on practicals. List of Practicals can be increased by the instructor as per the need. Experts from industries may be invited to deliver lectures and share experiences with the students.

MEANS OF ASSESSMENT

- Assignments and quiz/class tests, mid-term and end-term written tests
- Software installation, operation, development
- Actual laboratory and practical work exercises
- Viva-voce

RECOMMENDED BOOKS

1. Let Us Python by Yashavant Kanetkar BPB Publications.
2. Core Python Programming by Dr R. Nageshwara Rao
3. Learning Python by Mark Lutz; Pratham Books, Bangalore
4. Dive Into Python by Mark Pilgrim; Pratham Books, Bangalore
5. Think Python by Allen B. Downey; O'Reily Media
6. Python Programming For Beginners: A Must Read Introduction to Python Programming by Robert Richards; Pratham Books, Bangalore
7. e-books/e-tools/relevant software to be used as recommended by BTE/NITTTR, Chandigarh.

Websites for Reference:

https://docs.python.org/

http://swayam.gov.in

SUGGESTED DISTRIBUTION OF MARKS

Topic No.	Time Allotted (Periods)	Marks Allotted (%)
1.	06	08
2.	10	10
3.	08	12
4.	08	10
5.	16	22
6.	08	08
7.	10	08
8.	10	10
9.	08	12
Total	**84**	**100**

1.4 DATA SCIENCE - TOOLS AND TECHNIQUES

L T P
2 - 4

RATIONALE

This subject aims to cover the technical as well as soft skills of the students required for professional world. Organizations bring data scientists to augment their IT infrastructure by adding a humanized, unique angle and specialized skill sets. Therefore, in addition to mastering technical skills, data scientists must polish their soft skills to achieve their career potential. Students need to learn fundamental tools for data analysis, aspiring data analysts also need excellent communication, teamwork, and leadership skills.

LEARNING OUTCOMES

After undergoing the subject, the students will be able to:

- Use Google Suite for office data management tasks
- Demonstrate the ability to apply application software in an office environment.
- Use Excel for Data Analysis
- Use Advance Tools for Data Analysis
- Store code's history in Git and collaborate with others in GitHub.
- Develop soft skills required for professional world

DETAILED CONTENTS

1. Basic Collaboration Tools (04 Periods)

Creating , saving , downloading , sharing files/folders from Google drive creating and sharing Google docs, import and export docs, creating and sharing Google sheet, import and export Google sheet, Google forms and form responses ,creating Google slides to present your ideas. MS Word concepts, MS Excel Concepts, MS Power Point Concepts.

2. Introduction to Data Analysis Using Excel (06 Periods)

Introduction to Spreadsheets- Reading Data into Excel, Basic Data Manipulation in Excel, Basic Functions in Excel
Spreadsheet Functions to Organize Data- IF, nested IF, VLOOKUP and HLOOKUP Functions in Excel
Introduction to Filtering, Pivot Tables, and Charts- Data Filtering in Excel, Use of Pivot Tables in Excel, Introduction to Charts in Excel

Advanced Graphing and Charting- Line Graphs, Bar Graphs and Pie Charts, Pivot Charts, Scatter Plots, Histograms

3. Advance Tools for Data Analysis (06 Periods)

BIGML -Introduction to the BigML Interface, Resource, List View, Resource View, Searching and Deleting Resources, Changing Resource metadata
Natural Language Toolkit - Introduction and Installation, Tokenizing text, Wordnet.
Talend-Introduction, installation, Creating Project, Importing Project, components for data Integration
Tableau- Start Page, Connecting to Excel Files, Connecting to Text Files, Connect to Microsoft SQL Server, Creating and Removing Hierarchies, Joining Tables.

4. Code Development with Git and GitHub (06 Periods)

Introduction to Version Control-Version Control, Keeping Historical Copies, Version Control and Automation, What is Git?, Installing Git
Using Git Locally-Deleting and Renaming Files, Undoing Changes Before Committing, Amending Commits, Rollbacks.
Working with Remotes-Introduction to github, Working with Remotes, Fetching New Changes, Updating the Local Repository, Best Practices for Collaboration

5. Professional Soft Skills (06 Periods)

Resume Writing- Introduction to the Resume, Types of Resumes, Address Section, Summary Section, Education Detail, Technical Skills, Projects Detail, Strengths, Extra Curricular, Personal Detail, Conclusion
Email Writing-Organization, Style & Editing Basics, Common Errors in Punctuation, Effective Subject Lines, Well-organized Email Text, Overview and key language of Announcement Emails.
Mock Interview-Introduction, Benefits of Mock Interviews, Interview Preparation-Research the Employer , Research the Position, Know Yourself, Know the Interview Format, Make a List of Questions
Group Discussion- Introduction, Types of GD, Evaluation Criteria, Do's and Don'ts

LIST OF PRACTICALS

1. Creating a document using different font, changing font size and color, changing the appearance through bold/italic/underline.
2. Creating a document using subscript and superscript, justification of the document.
3. Create a document using Bullets and Numbering.
4. Create a document using page number, header and footer.
5. Creating table, formatting cells, use of different border styles, shading in tables, merging of cells, and partition of cells, inserting and deleting a row in a table in MS word document.

6. Create a spreadsheet with LOOKUP/VLOOKUP features.
7. Create different charts in excel and implement formulas(automatic and use defined).
8.Create and share files/folders in Google drive
9. Create and share Google docs.
10. Create and share Google sheets.
11. Create and share Google Forms.
12. Create and share Google slides.
13. Perform Data Manipulation in Excel
14. Perform Arithmetic Manipulation in Excel
15. Use The "IF" Command in Excel Using Numerical Data
16. Create Bar Graphs and Pie Chart in excel
17. Use a table to filter, sort and see totals.
18. Use multiple pivot tables and pivot charts to create dashboard.
19. Use calculated fields on top of the pivot table to calculate profitability and find anomalies.
20. Installing Git on Windows and Practical Application of diff and patch
21. Write a resume including all the sections in proper format
22. Write an email to a hotel manager to make a reservation
23. Write an e-mail to the supervisor intimating about your absence

INSTRUCTIONAL STRATEGY

As the subject is practice oriented, more stress should be given to students to do the work practically. The features of software packages G-Suite/Excel and other analytical tools to be demonstrated in class using LCD projector.

MEANS OF ASSESSMENT

− Assignments and quiz/class tests, mid-term and end-term written tests

− Actual laboratory and practical work, exercises and viva-voce

− Software installation, operation, development and viva-voce

RECOMMENDED BOOKS

1. Working in Microsoft Office - Richard Mansfield - Tata McGraw Hill Education.
2. Microsoft Excel 2019 Data Analysis and Business Modeling- L. Winston Wayne
3. Excel Data Analysis For Dummies-Nelson Stephen L.
4. Introduction to Data Science-Laura Igual Santi Seguí
5. Essentials of Data Science and Analytics-Dr. Amar Sahay
6. Beginning Git and GitHub-Mariot Tsitoara
7. Introduction to Git and GitHub-Bobby Iliev
8. How to Book of Writing Skills-J.H Hood

Websites for Reference

• http://office.microsoft.com/en-us/training/CR010047968.aspx

• https://gsuite.google.com/learning-center

SUGGESTED DISTRIBUTION OF MARKS

Topic No.	Time Allotted (Periods)	Marks Allotted (%)
1	04	15
2	06	25
3	06	20
4	06	25
5	06	15
Total	**28**	**100**

1.5 SEMINAR AND CASE STUDIES

L	T	P
-	-	6

RATIONALE

This subject aims to cover the various Case Studies of various industries who are using Data Science and Machine Learning to boost their production, make smarter decisions and develop innovative products. There are many case studies that prove that data science has boosted the performance of industries and has made them smarter and more efficient.

Data Science has not only accelerated the performance of companies but has also made it possible for them to manage & sustain their performance with ease.

LEARNING OUTCOMES

After undergoing the subject, the students will be able to:

- Relate the real world case studies with concepts of Data Science and Machine Learning
- In depth learning process by studying a particular Case Study
- Apply fundamental and disciplinary concepts and methods in ways appropriate to their areas of study
- Demonstrate skill and knowledge of current information and technological tools and techniques specific to the professional field of study.
- Represent their understanding in the form of Seminar

Seminar and Case Studies aims to expose the students to various case studies of industries and research being done in Data Science and Machine Learning. It is expected from them to get acquainted with Data Science ecology. Depending upon their interest students can select a case study/research.

After completing their Case Study , students need to present their learning in the form of a Seminar, in which their approach towards the subject shall be evaluated along with presentation skills.

Students are to be instructed to prepare the presentation of current research work of the concerned field as suggested by the mentor/supervisor/faculty member. The following parameters could be taken into account in determining the effectiveness of the research presentation. Before the presentation, the topic/content has to be approved by the mentor/supervisor/faculty member.

1. **Presentation**

- Clarity and understanding of the topic
- Communication skill
- Speed/pace of presentation (too fast? too slow? too static? varied?)
- Use of effective tools/animation/handouts
- Appropriate size and finalization within the given time frame

2. **Content**

- Relevancy of the topic and its usefulness
- familiarity with the topic, including current research findings
- appropriateness of the level of content
- use of relevant examples and illustrations (where appropriate)
- appropriateness of the content of aids (handouts, overheads, etc)

3. **Organization and Integration**

- Brief introduction of the proposal
- logic and coherence of material
- Pointwise illustration of the content
- Different parts of the presentation are connected and balanced Links to other areas (if appropriate)
- Findings and future scope are to be included in the conclusion

Broad areas of Data Science and Machine Learning

- Medical Services
- Language Processing
- Business Management
- Image Recognition

- Speech Recognition
- Face Detection
- Video Games
- Computer Vision
- Banking
- E-commerce/ Finance
- Manufacturing
- Automation of Transport System
- Weather Forecasting

Note: The teachers may guide /help students to identify their Case Study and check out their plan of action well in advance.

MEANS OF ASSESSMENT

The teachers conduct internal performance assessments of students. Criteria for assessment will be as follows:

SNo	Criteria	Weightage
(a)	Relevancy	15%
(b)	Able to apply Data Science Concept with the Case Study	25%
(c)	Clarity of Presentation	20%
(d)	Case Study Report	30%
(e)	VIVA	10%

RECOMMENDED BOOKS

1. Python Machine Learning Case Studies by Danish Haroon , Apress.

2.1 ADVANCED PROGRAMMING FOR DATA SCIENCE

L	T	P
6	-	6

RATIONALE

This course introduces Python libraries like NumPy, Pandas, Matplotlib and Seaborn. Upon completion of this course, the student will be able to do mathematical calculations and operations on Data by using NumPy and Pandas. Apart from mathematical and statistical operations students will be able to visualize the data more effectively by using Matplotlib and Seaborn libraries.

LEARNING OUTCOMES

After undergoing the course, the students will be able to:

- Perform mathematical operations by using NumPy.
- Arrange and manipulate the data using Pandas Series
- Import the data from csv file into Pandas Dataframe.
- Learn how to do various mathematical and statistical operations on Panadas DataFrame.
- Learn how to visualize the data using matplotlib library
- Learn how to visualize and draw different plots using Seaborn

DETAILED CONTENTS

1. NumPy (16 Periods)

Introduction, Ndarray Object, Creating ndarrays, DataTypes for ndarrays, array attributes of numpy, Indexing and Slicing, Iterating over Array, Array Manipulation, Mathematical Functions, Arithmetic Operations, Statistical Functions, Sort Search and Counting Functions, Matrix Library, numpy.linalg module.

2. Pandas Introduction and Series (18 Periods)

Introduction, features and installation, Data Structures in Pandas : Series, DataFrame and Panel, Descriptive Statistics

Series : Defining series from list,dictionary and numpy arrays, Indexing in series, Creating Series from Dataset with the pd.read_csv() and accessing rows using .head(), .tail(), Attributes

in Series object (.values, .index, .size, .name, and .is_unique.dtype), Methods on a Series Objects (.sum(), .product(), .mean(), .idxmax(), .idxmin(), .value_counts(), .apply(), .map()), Sorting of series by value and index.

3. DataFrames in Pandas (20 Periods)

Introduction to DataFrames, Data Frame Attributes (.index, .values, .shape, .ndim, and .dtypes, .columns and .axes), One Column Selection from a DataFrame, Two or More Column Selection from a DataFrame, New Column Addition, Broadcasting (.add(), .sub(), .mul() and .div()), Data Frame methods (.value_counts() , .dropna() , .fillna() , .astype() , .rank()), Sorting in DataFrame with respect to index and values

Filtering the Data : Filtering of rows in DataFrame based on one or more conditions, Methods for filtering (isin(), .isnull(), .notnull(), .between(), .duplicated(), .drop_duplicates(), .unique() and .nunique())

Data Extraction : Customize and Reset the index (set_index and reset_index), Row retrieval by loc Accessor and iloc Accessor, Methods for Data Extraction (.drop(), .pop(), .sample(), .nsmallest(), .nlargest(), .where(), .query(), .apply(), .copy())

MultiIndex, GroupBy, Merging Joining and Concatenating Data Frames, Importing and Exporting Excel/CSV files in Pandas

4. Matplotlib (16 Periods)

Introduction to Matplotlib, Parts of Figure, Types of input to plotting functions, Introduction to Pyplot, Figure Class, Axes Class, Different types of plots (Line Graph, Bar Chart, Histogram, Scatter Plot, Pie Chart, Box Chart), Multiplots, The Lifecycle of Plot, Annotations

Working with images : Importing image data into numpy arrays, plotting numpy arrays as image

5. Seaborn (14 Periods)

Introduction and features of Seaborn, Seaborn vs Matplotlib, Figure Aesthetic, Color palette, Seaborn Plotting functions : relplot, catplot, Histogram, Bar Plot,Box Plot, Facet Grid, Pair Grid, HeamMap

LIST OF PRACTICALS

SNo	Objective
Module -1 : NumPy	
1.1	Write a NumPy program to create a numpy array and test whether none of the elements of a given array is zero
1.2	Write a NumPy program to create an element-wise comparison (greater, greater_equal, less and less_equal) of two given arrays.
1.3	Write a NumPy program to create an array of 10 zeros,10 ones, 10 fives
1.4	Write a NumPy program to create an array of all the even integers from 50 to 90
1.5	Write a NumPy program to create a 3x3 identity matrix
1.6	Write a NumPy program to generate an array of 25 random numbers from a standard normal distribution
1.7	Write a NumPy program to create a vector with values from 0 to 20 and change the sign of the numbers in the range from 9 to 15
1.8	Write a NumPy program to create a 3x4 matrix filled with values from 10 to 21
1.9	Write a NumPy program to create a 5x5 zero matrix with elements on the main diagonal equal to 1, 2, 3, 4, 5
1.10.	Write a NumPy program to compute sum of all elements, sum of each column and sum of each row of a given array
1.11	Write a NumPy program to sort a given array by row and column in ascending order.
1.12	Write a NumPy program to replace all numbers in a given array which is equal, less and greater to a given number
1.13	Write a NumPy program to multiply two given arrays of same size element-by-element
1.14	Write a NumPy program to compute the multiplication of two given matrixes
1.15	Write a NumPy program to compute the cross product of two given vectors
1.16	Write a NumPy program to compute the determinant of a given square array
1.17	Write a NumPy program to compute the eigenvalues and right eigenvectors of a given square array
1.18	Write a NumPy program to compute the inverse of a given matrix.
1.19	Write a NumPy program to compute the sum of the diagonal element of a given array

1.20.	Write a NumPy program to create a 3x3x3 array with random values.
1.21	Write a NumPy program to create a 5x5 array with random values and find the minimum and maximum values
1.22	Write a NumPy program to create a structured array from given student name, height, class and their data types. Now sort by class, then height if class are equal.
1.23	Write a NumPy program to calculate inverse sine, inverse cosine, and inverse tangent for all elements in a given array
1.24	Write a NumPy program to compute the mean, standard deviation, and variance of a given array along the second axis
1.25	Write a NumPy program to compute the covariance matrix of two given arrays
Module -2 Pandas Introduction and Series	
2.1	Write a program to create a series from a list, numpy array and dict?
2.2	Write a program to get the items of series A not present in series B?
2.3	Write a program to get the items not common to both series A and series B?
2.4	Write a program to get frequency counts of unique items of a series?
2.5	Write a program to Compute the minimum, 25th percentile, median, 75th, and maximum of Series
2.6	Write a program to get the positions of items of series A in another series B?
2.7	Write a program to compute the difference of differences between consecutive numbers of a series?
2.8	Write a program to filter words that contain at least 2 vowels from a series?
2.9	Write a Pandas program to sort a given Series
2.10.	Write a Pandas program to create the mean and standard deviation of the data of a given Series
Module -3 DataFrames in Pandas	
3.1	Write a Pandas program to create and display a DataFrame from a specified dictionary data which has the index labels
3.2	Given a 2D List, Write a Pandas program to convert it into corresponding dataframe and display it
3.3	Wrire a program to combine many series to form a dataframe?
3.4	Given a CSV file, Write a program to read it into a dataframe and display it
3.5	Write a program to import only every 10th row from a csv file to create a dataframe
3.6	Write a program to import only roll_number, percent_marks columns from a student.csv file having following structure : student.csv [roll_number,name,class,percent_marks,hometown]

3.7	Write a program to check if a DataFrame is having any missing value
3.8	Given is a dataframe showing name, occupation, salary of people. Find the average salary per occupation
3.9	Given a dataframe with NaN Values, fill the NaN values with 0
3.10.	Given is a dataframe showing Company Names (cname) and corresponding Profits (profit). Convert the values of Profit column such that values in it greater than 0 are set to True and the rest are set to False.
3.11	Write a Pandas program to join the two given data frames along rows and assign all data Test Data: Test Data: Student_data1: student_id name marks 0 S1 Danniella Fenton 200 1 S2 Ryder Storey 210 2 S3 Bryce Jensen 190 3 S4 Ed Bernal 222 4 S5 Kwame Morin 199 student_data2: student_id name marks 0 S4 Scarlette Fisher 201 1 S5 Carla Williamson 200 2 S6 Dante Morse 198 3 S7 Kaiser William 219 4 S8 Madeeha Preston 201

3.12	Write a Pandas program to join the two given data frames along rows and merge with another dataframe along the common column id. Test Data: Student_data1: student_id name marks 0 S1 Danniella Fenton 200 1 S2 Ryder Storey 210 2 S3 Bryce Jensen 190 3 S4 Ed Bernal 222 4 S5 Kwame Morin 199 student_data2: student_id name marks 0 S4 Scarlette Fisher 201 1 S5 Carla Williamson 200 2 S6 Dante Morse 198 3 S7 Kaiser William 219 4 S8 Madeeha Preston 201 exam_data: student_id exam_id 0 S1 23 1 S2 45 2 S3 12 3 S4 67 4 S5 21 5 S7 55 6 S8 33 7 S9 14 8 S10 56 9 S11 83 10 S12 88 11 S13 12
3.13	Write a Pandas program to merge two given dataframes with different columns Test Data: data1: key1 key2 P Q 0 K0 K0 P0 Q0 1 K0 K1 P1 Q1 2 K1 K0 P2 Q2 3 K2 K1 P3 Q3 data2: key1 key2 R S 0 K0 K0 R0 S0 1 K1 K0 R1 S1 2 K1 K0 R2 S2 3 K2 K0 R3 S3

3.14	Write a Pandas program to split a dataset, group by one column and get mean, min, and max values by group. Using the following dataset find the mean, min, and max values of purchase amount (purch_amt) group by customer id (customer_id). Test Data: ord_no purch_amt ord_date customer_id salesman_id 0 70001 150.50 2012-10-05 3005 5002 1 70009 270.65 2012-09-10 3001 5005 2 70002 65.26 2012-10-05 3002 5001 3 70004 110.50 2012-08-17 3009 5003 4 70007 948.50 2012-09-10 3005 5002 5 70005 2400.60 2012-07-27 3007 5001 6 70008 5760.00 2012-09-10 3002 5001 7 70010 1983.43 2012-10-10 3004 5006 8 70003 2480.40 2012-10-10 3009 5003 9 70012 250.45 2012-06-27 3008 5002 10 70011 75.29 2012-08-17 3003 5007 11 70013 3045.60 2012-04-25 3002 5001
Module - 4 : Matplotlib	
4.1	Write a Python program to draw a line with suitable label in the x axis, y axis and a title
4.2	Write a Python program to plot two or more lines with legends, different widths, colors and style
4.3	Write a Python programming to display a bar chart and pie chart of the popularity of programming Languages Sample data: Programming languages: Java, Python, PHP, JavaScript, C#, C++ Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7
4.4	Write a Python programming to display a horizontal bar chart of the popularity of programming Languages. (Use different color for each bar) Sample data: Programming languages: Java, Python, PHP, JavaScript, C#, C++ Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7

4.5	Write a Python program to create bar plots from a DataFrame. Sample Data Frame: a b c d e 2 4,8,5,7,6 4 2,3,4,2,6 6 4,7,4,7,8 8 2,6,4,8,6 10 2,4,3,3,2
4.6	Write a Python program to draw a scatter plot comparing two subject marks of Mathematics and Science. Use marks of 10 students. (Math Marks should be of color Red, and Science Marks should be of color Green) Sample data: Test Data: math_marks = [88, 92, 80, 89, 100, 80, 60, 100, 80, 34] science_marks = [35, 79, 79, 48, 100, 88, 32, 45, 20, 30] marks_range = [10, 20, 30, 40, 50, 60, 70, 80, 90, 100]
Module - 5 : Seaborn	
5.1	Write a program to draw Histogram using Seaborn for a sample data
5.2	Draw vertical and horizontal Bar Plot using Seaborn
5.3	Create a program using Box Plot to visualize the Data
5.4	Create a program using Heal Map to visualize the Data

INSTRUCTIONAL STRATEGY

Teachers should lay emphasis on practicals. List of Practicals can be increased by the instructor as per the need. Experts from industries may be invited to deliver lectures and share experiences with the students.

MEANS OF ASSESSMENT

- Assignments and quiz/class tests, mid-term and end-term written tests
- Software installation, operation, development
- Actual laboratory and practical work exercises
- Viva-voce

RECOMMENDED BOOKS

1. Mastering python for data science, Samir Madhavan.

2. Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython By Wes McKinney
3. Python Data Analytics: With Pandas, NumPy, and Matplotlib
4. Python for Data Science For Dummies By Luca Massaron and John Paul Mueller
5. Learning Pandas – Python Data Discovery and Analysis Made Easy By Michael Heydt
6. Learning the Pandas Library by By Matt Harrison
7. Pandas for Everyone: Python Data Analysis By Daniel Y. Chen
8. Hands-On Data Analysis with NumPy and Pandas By Curtis Miller
9. e-books/e-tools/relevant software to be used as recommended by BTE/NITTTR, Chandigarh.

Websites for Reference:

https://numpy.org/doc/
https://pandas.pydata.org/docs/
https://matplotlib.org/
https://seaborn.pydata.org/
http://swayam.gov.in

SUGGESTED DISTRIBUTION OF MARKS

Topic No.	Time Allotted (Periods)	Marks Allotted (%)
1.	16	20
2.	18	20
3.	20	24
4.	16	18
5.	14	18
Total	**84**	**100**

2.2 EXPLORATORY DATA ANALYSIS AND VISUALIZATION METHODS

L T P
6 - 6

RATIONALE

Exploratory data analysis (EDA) is used by data scientists to analyze and investigate data .Data scientists can use exploratory analysis to ensure the results they produce are valid and applicable to any desired business outcomes and goals.

LEARNING OUTCOMES

After completion this course, the students will be able to:

- Understand the basic data science concepts and methods of EDA
- to see what data can reveal
- Visualizing the data to get a better perspective.
- determine if the statistical techniques used for data analysis are appropriate
- Identify obvious errors, better understand patterns within the data, detect outliers or anomalous events, and find interesting relations among the variables.

DETAILED CONTENTS

1. INTRODUCTION TO EDA (14 Periods)

Data Analysis, Types of data analysis; - descriptive, predictive and perspective, Exploratory Data Analysis and inferential with example, Definition of Exploratory Data Analysis, objectives of EDA, Advantages of EDA, Steps involved in EDA.

2. DATA SOURCING (20 Periods)

Definition, Data Source Nomenclature, Data Source Types-Machine Data Sources, File Data Sources, Purpose of Data Source, Data Sources- Private Data, Public Data, Main Data Types-Primary, Secondary, Prioritizing Data Quality, Relying on Data Providers, Data Sourcing Challenges and Concerns, Integration Of Data Resources

3. DATA CLEANING (22 Periods)

Introduction to Data Cleaning, The importance of integrity of data, Regular Expressions, Tools and Techniques for Data Cleaning, Effective data cleaning techniques: Remove duplicates, Remove irrelevant data, Standardize capitalization, Convert data type, Clear formatting, Fix errors, Language translation, Handle missing values, Optimize the data-cleaning process, Data Cleaning using Python

4. EXPLORATORY ANALYSIS (14 Periods)

Univariate Analysis: Frequency Distribution Tables, Bar Charts, Histograms, Pie Charts, Frequency Polygons
Bivariate Analysis: Numeric-Numeric Analysis-Scatter Plot, Numeric - Categorical Analysis , Categorical -Categorical Analysis, Linear Correlation, Bivariate Analysis of two categorical Variables
Multivariate Analysis: Cluster Analysis, Principal Component Analysis (PCA), Correspondence Analysis.

5. DATA VISUALIZATION (14 Periods)

Introduction to Data Acquisition – Applications –Visualization-Introduction - Terminology- Basic Charts and Plots, Pie Chart, Bar Chart, Histogram, Gantt Chart, Scatter Plot, Pictogram - Multivariate Data Visualization- Data Visualization Techniques.
Data Visualization Tools– Rank Analysis Tools- Trend Analysis Tools- Multivariate Analysis Tools- Distribution Analysis Tools, Introduction of Data Visualization tools- Tableau, Microsoft Power BI

LIST OF PRACTICALS

1. Write a program to import the necessary libraries for data cleaning.

2. Write a program to print out columns with missing values, and shows its amount.

3. Write a program to find the percentage of missing values in each of the columns.

4.Wririte a program to replace values in the columns

5. Write a program to rename your columns

6. Write a program to remove rows having duplicate values.

7. Write a program to change the data type of a column.

8. Identify Columns That Contain a Single Value

9. Delete Columns That Contain a Single Value

10. Consider Columns That Have Very Few Values

11. Remove Columns That Have A Low Variance

12. Identify Rows that Contain Duplicate Data

13. Delete Rows that Contain Duplicate Data

14. Identify and Delete Zero-Variance Predictors

15. Display the count of non-null values in each column of a DF

16. Display the count of null values in each column of a DF

17. Creating common visualizations (bar charts, line charts etc.) with the data.

18. Assemble a dashboard layout using dashboard filters on Tableau.

INSTRUCTIONAL STRATEGY

The subject is conceptual. Students should be given clear idea about the basic concepts of Exploratory Data analysis and Visualization methods. Student should be asked to explain the working of analysis and its applications in Data Science.

MEANS OF ASSESSMENT

− Assignments and quiz/class tests, mid-term and end-term written tests
− Viva-voce

RECOMMENDED BOOKS

1. Data Science and Analytics with Python Programming by V. K. Jain(Author) , Khanna Book Publication.
2. "Information Dashboard Design: Displaying Data for At-a-glance Monitoring" by Stephen Few
3. "Storytelling With Data: A Data Visualization Guide for Business Professionals" by Cole Nussbaumer Knaflic
4. Hands-On Exploratory Data Analysis with Python: Perform EDA Techniques to Understand, Summarize, and Investigate Your Data

REFERENCE WEBSITES

1. https://javatpoint.com
2. https://www.analyticsvidhya.com
3. https://www.upgrad.com
4. https://iiitl.ac.in

SUGGESTED DISTRIBUTION OF MARKS

Topic No.	Time Allotted (Periods)	Marks Allotted (%)
1.	14	20
2.	20	20
3.	22	20
4.	14	20
5.	14	20
Total	**84**	**100**

2.3 MACHINE LEARNING ALGORITHMS

L T P
6 - 6

RATIONALE

Machine learning algorithms have been powering the world around us. Most industries working with large amounts of data have recognized the value of Machine Learning technology.
This course introduces the basic concept and techniques of Machine Learning. Studying this course will help the students to develop skills in using recent ML software for solving various practical problems. After completing the program, students may apply for entry roles in positions such as Data Scientist, Computer Vision Engineer, Machine Learning Engineer, and NLP Engineer.

LEARNING OUTCOMES

After completing this course, the student will be able to

- Understand the fundamental concepts of machine learning algorithms
- Appreciate the importance of visualization in the data analytics solution
- Apply a mathematical approach to solve various Machine Learning and Data science-based problems.
- Develop various models involved in learning from data.

DETAILED CONTENTS

1. Foundations for Machine learning (20 Periods)

ML Techniques overview, Validation Techniques (Cross-Validations), Feature Reduction/Dimensionality reduction, Principal components analysis (Eigen values, Eigen vectors, Orthogonality).

2. Regression Techniques (20 Periods)

Regression basics: Relationship between attributes using Covariance and Correlation, Relationship between multiple variables: Regression (Linear, Multivariate) in prediction, Residual Analysis, Identifying significant features, feature reduction using AIC, multi-collinearity, Non-normality and Heteroscedasticity, Hypothesis testing of Regression Model, Confidence intervals of Slope, R-square and goodness of fit, Influential Observations – Leverage.

3. Multiple Linear Regression and Non-Linear Regression (20 Periods)

Polynomial Regression, Regularization methods, Lasso, Ridge and Elastic nets, Categorical Variables in Regression, Logit function and interpretation, Types of error measures (ROCR), Logistic Regression in classification.

4. Clustering (20 Periods)

Distance measures, Different clustering methods (Distance, Density, and Hierarchical), Iterative distance-based clustering, Dealing with continuous, categorical values in K-Means, Constructing a hierarchical cluster, K-Medoids, k-Mode and density-based clustering, Measures of quality of clustering.

5. Classification Techniques (20 Periods)

Naive Bayes Classifiers: Model Assumptions, Probability estimation, Required data processing, M-estimates, Feature selection: Mutual information, Classifier.
K-Nearest Neighbors: Computational geometry, Voronoi Diagrams, Delaunay Triangulations, K-Nearest Neighbor algorithm, Wilson editing and triangulations, Aspects to consider while designing K-Nearest Neighbor.
Support Vector Machines: Linear learning machines and Kernel space, Making Kernels and working in feature space, SVM for classification and regression problems.

Advanced Machine Learning topics: Neural Network algorithms, Deep Learning algorithms, Natural Language processing algorithms.

LIST OF PRACTICAL

1. Implement and demonstrate the FIND-S algorithm for finding the most specific hypothesis based on a given set of training data samples. Read the training data from a .CSV file.
2. For a given set of training data examples stored in a .CSV file, implement and demonstrate the Candidate-Elimination algorithm to output a description of the set of all hypotheses consistent with the training examples.
3. Write a program to demonstrate the working of the decision tree based ID3 algorithm. Use an appropriate data set for building the decision tree and apply this knowledge to classify a new sample.
4. Write a program to implement the naïve Bayesian classifier for a sample training data set stored as a .CSV file. Compute the accuracy of the classifier, considering few test data sets.
5. Assuming a set of documents that need to be classified, use the naïve Bayesian Classifier model to perform this task. Built-in Java classes/API can be used to write the program. Calculate the accuracy, precision, and recall for your data set.

6. Write a program to construct a Bayesian network considering medical data. Use this model to demonstrate the diagnosis of heart patients using standard Heart Disease Data Set. You can use Java/Python ML library classes/API.
7. Apply EM algorithm to cluster a set of data stored in a .CSV file. Use the same data set for clustering using k-Means algorithm. Compare the results of these two algorithms and comment on the quality of clustering. You can add Java/Python ML library classes/API in the program.
8. Write a program to implement k-Nearest Neighbor algorithm to classify the iris data set. Print both correct and wrong predictions. Java/Python ML library classes can be used for this problem.

MEANS OF ASSESSMENT

- Assignments and quiz/class tests, mid-term and end-term written tests
- Software installation, operation, development
- Actual laboratory and practical work exercises
- Viva-voce

RECOMMENDED BOOKS

- Machine Learning, Tom M. Mitchell, McGraw-Hill
- Introduction to Machine Learning, Alex Smola and S.V.N. Vishwanathan, Cambridge University Press.
- Introduction to Machine Learning (Second Edition), Ethem Alpaydın, The MIT Press Cambridge, Massachusetts London, England
- The Elements of Statistical Learning, by Trevor Hastie, Robert Tibshirani, Jerome H. Friedman (freely available online)
- Pattern Recognition and Machine Learning, by Christopher Bishop (optional)

Web Links:

- Machine Learning Mastery
- Launch Your Career in Data Science (elitedatascience.com)
- Machine Learning, Data Science, Big Data, Analytics, AI - KDnuggets
- Kaggle: Your Machine Learning and Data Science Community

SUGGESTED DISTRIBUTION OF MARKS

Topic No.	Time Allotted (Periods)	Marks Allotted (%)
1.	10	12
2.	20	25
3.	18	20
4.	16	18
5.	20	25
Total	**84**	**100**

2.4 PROJECT WORK

L	T	P
-	-	8

RATIONALE

Project Work aims at developing innovative skills in the students whereby they apply in totality the knowledge and skills gained through the course work in the solution of particular problem or by undertaking a project. The individual students have different aptitudes and strengths. Project work, therefore, should match the strengths of students. For this purpose, students should be asked to identify the type of project work, they would like to execute. It is also essential that the faculty of the respective department may have a brainstorming to identify suitable project assignments for their students. The project assignment can be individual assignment or a group assignment. There should not be more than 5 students if the project work is given to a group. The students should identify themselves or accept the given project assignment at least two to three months in advance. The project work identified in collaboration with industry should be preferred.

LEARNING OUTCOMES

After undergoing the course, the students will be able to:

- Use effectively oral, written and visual communication
- Demonstrate skill and knowledge of current information and technological tools and techniques specific to the professional field of study.
- Identify, analyze and solve problems creatively through sustainment critical investigation.
- Develop, leadership abilities..
- Apply fundamental and disciplinary concepts and methods in ways appropriate to their areas of study.

A suggestive criterion for assessing student performance by the external (personnel from industry) and internal (teacher) examiner is given in table below:

SNo	Performance criteria	Max Marks	Rating Scale				
			Excellent	Very Good	Good	Fair	Poor
1	Selection of project assignment	10	10	8	6	4	2
2	Planning and execution of considerations	10	10	8	6	4	2
3	Quality of performance	20	20	16	12	8	4
4	Providing solution of the problems or production of final product	20	20	16	12	8	4
5	Sense of responsibility	10	10	8	6	4	2
6	Self expression/ communication skills	5	5	4	3	2	1
7	Interpersonal skills/human relations	5	5	4	3	2	1
8	Report writing skills	10	10	8	6	4	2
9	Viva voce	10	10	8	6	4	2
	Total Marks	100	100	80	60	40	20

In order to qualify for the diploma, students must get "Overall Good grade" failing which the students may be given one more chance of undergoing 8 -10 weeks of project oriented work and re-evaluated before being disqualified and declared "not eligible to receive diploma ". It is also important to note that the students must get more than six "goods" or above "good" grade in different performance criteria items in order to get "Overall Good" grade.

Important Notes

1. These criteria must be followed by the internal and external examiner and they should see the daily, weekly and monthly reports while awarding marks as per the above criteria.

2. The criteria for evaluation of the students have been worked out for 100 maximum marks. The internal and external examiners will evaluate students separately and give marks as per the study and evaluation scheme of examination.

3. The external examiner, preferably, a person from industry/organization, who has been associated with the project-oriented professional training of the students, should evaluate the students performance as per the above criteria.

4. It is also proposed that two students or two projects which are rated best be given merit certificate at the time of annual day of the institute. It would be better if specific nearby industries are approached for instituting such awards.

5. The teachers are free to evolve another criterion of assessment, depending upon the type of project work.

6.The students must submit a project report of not less than 50 pages (excluding coding). The report must follow the steps of Data Science and Machine Learning Model Life Cycle Events

10. LAB & OTHER RESOURCE REQUIREMENT

Data Science and Machine Learning Lab
Furniture Requirements

LIST OF EQUIPMENT

S No	Description	Qty	Approx. price (Rs)
10.1 DATA SCIENCE AND MACHINE LEARNING LAB			
1	**Computer Server** (Intel® Core™ i9-12900KS Processor (30M Cache, up to 5.50 GHz) NVIDIA GEFORCE RTX™ 3090 Ti, 16 GB GDDR6, 32 GB, 2 x 16 GB, DDR5, 4800 MHz, dual-channel, 1 TB, M.2, PCIe NVMe, SSD, 43.9-cm display (1920X1080) 360Hz Keyboard , Mouse with Latest Windows Server Operating System/Ubuntu 16.04	1	3,50,000/-
2	Computer Desktop (i7 latest Generation, 1TB Hard disk, 8GB RAM or more, Pre loaded window 10/11/Latest) Preloaded window-10/11, 19-inch standard ratio LCD monitor, Keyboard , Mouse, Basic Connecting Ports (VGA, HDMI, USB)	60	90,000 * 60 = 54,00,000/-
3	Multifunctional Laser/Ink tank Printer	3	90,000/-
4	Multifunctional Printer, A3 size	1	50,000/-
5	Scanner	1	20,000/-
6	Laptop (i7 latest Generation, 1TB Hard disk, 8GB RAM or more, Pre loaded window 10/11/Latest)	1	1,00,000/-
7	Online UPS, 6KVA	2	2,50,000/-
8	Internet Connectivity	60 Nodes	1,00,000/-
9	Smart Board with Projector	1	2,00,000/-
10	Visual Studio Code (Open Source) / Any other Python IDE	-	-
11	Open-Source Tools: Python 3.6.5, Anaconda, TensorFlow, R-Language, OpenCV, H2O.ai, Apache Mahout, Weka, Keras, C, C++, JAVA.	-	-
12	MS Office latest or Libre Office or equivalent FOSS	1	20,000/- Per year
13	Oracle 12.1.0.2.0 with analytics.	1 User LS	
14	Matlab Tool boxes: Statistics and Machine Learning Toolbox Deep Learning Toolbox ,Reinforcement Learning Toolbox Deep Learning HDL Toolbox Text Analytics Toolbox Predictive Maintenance Toolbox	1 User LS	-

15	Web camera, Mike and speakers with audio system	1	50,000/-
16	Air Conditioner 2 ton	2	100,000/-
17	External Hard Disk (Space - 2TB)	1	10,000
18	External DVD Writer	1	5,000
19	Antivirus Software	5 Users	10,000/-
20	Miscellaneous- cables and connectors, computer stationery, printer consumables (inks), toner etc.	LS	30,000/-

Note : System Configuration of the Computer system and other equipments should be adopted to latest available configurations as per the need of the syllabus

10.2 Furniture Requirement

Norms and standards laid down by AICTE be followed for working out furniture requirements for this course.

- Furniture for laboratories 15 lacs

11. Suggestions For Effective Implementation of Curriculum

1. **Eligibility criteria** - This diploma course is a highly specialized course that requires programming and analytical skills of a certain level, so students from the Engineering background are more suitable for this programme. So, on the basis of recommended eligibility criteria the name of the course should be "**Post Graduate Diploma In Data Science and Machine Learning".**
2. A Review committee from higher institutions/industry should be framed for continuous review.
3. Necessary up-gradation in labs is required to provide the practical implementation of the ML concepts.
4. Rigorous Training of Faculty Members is required who are going to undertake this course.
5. Industry collaboration and feedback systems should also be incorporated.
6. Higher institutions (IIT/NIT/STATE ENGINEERING COLLEGES) could be assigned as mentors for this course.
7. Faculty members are to be encouraged to create a research environment in the concerned field.
8. Area specific center of studies could also be developed.
9. Regular visits of students to various industries working in this field should be organized.
10. Faculty coordinators are appointed to evaluate the various pedagogical parameters
11. Regular meetings (Online/Offline) among the faculty members are to be held to evaluate shortcomings and their solutions.
12. Consulting firms can be appointed for the establishment of the course-specific New Lab.
13. Expert Lecture required from the people of the industry.
14. IIT Kanpur, IIT Indore, IIIT Jabalpur and other Institutions can be incorporated as Training Partners.
15. MOU Can be Signed with other institutions and industries.
16. Physical Visits of Laboratory already set up at other institutions and industries.

12. LIST OF MENTORS/EXPERTS/REVIEWERS

1. Subhajit Roy, Associate Professor Department of Computer Science And Engineering, IIT Kanpur
2. Amey Karkare , Associate Professor Department of Computer Science And Engineering, IIT Kanpur .
3. Ashish Tendulkar , Machine Learning Architect and Specialist (GOOGLE INDIA)
4. Prof. Arun Sharma, Dept. of AI & Data Science, IGDTUW, Delhi
5. Prof. Tanveer Ahmad, Dept. Computer Engg., Jamia Millia Islamia, New Delhi
6. Dr. Surya Prakash, Associate Professor, Dept. Computer Science, IIT Indore
7. Dr. Rajib kr. Jha, Associate Professor, Dept. Electrical Engg., IIT Patna
8. Dr. Aparjita Ojha, Professor, Computer Science, IIIT Jabalpur
9. Dr. Sandesh Gupta, Asst. Professor, CSJM University Kanpur
10. Prof. Krishna Kumar Singh,Assistant Professor,RGUKT IIIT Nuzvid, Andhra Pradesh (Govt. Andhra Pradesh)
11. Dr. Chandan Gautam,Data Scientist-I,Institute for Infocomm Research (I2R), Agency for Science, Technology and Research (A*STAR), Singapore
12. Mr Shravan Kumar Yadav, Data Architect, Girl Effect Enterprise India Pvt Ltd

13. LIST OF PARTICIPANTS

The following experts participated in workshop for Developing the Curricula Structure and Contents of various Diploma Programmes at IRDT, Kanpur :

1. Mr F R Khan , Principal (Headquarter) DTE Kanpur
2. Mr Janbeg Loni , Principal, Government Polytechnic, Gaziabad
3. Mrs Rashmi Sonkar, Join Secretary , Board of Technical Education UP Lucknow
4. Mr Sheetanshu Krishna, Lecturer Computer, Govt Girls Polytechnic, Amethi
5. Mrs Swati Verma, Lecturer Computer, Km Mayawati Govt Girls Polytechnic, Badalpur
6. Dr. Vivek Singh Verma, Lecturer Computer , Govt. Poly. Shahabad Rampur
7. Mr Vishal Aggrawal, Lecturer Computer IRDT UP Kanpur (Coordinator)

Name of Course - Post Graduate Diploma in Data Science and Machine learning

Staff Requirement

Sr. No.	Name of Post	Post	Remarks
1	Principal	01	Common
2	Head of Department Computer Science	01	Common with Computer Science / Information Technology or Relevant Branch as per AICTE rules
3	Lecturer Computer Science	03	Computer Science/Information Technology or Relevant Branch as per AICTE rules

Note:

1. Services of other discipline staff of the Institute may be utilized if possible.
2. Qualifications of Staff: as per service rule.
3. AICTE APH-2023-24 or latest Guidelines to be followed.