Course Title: Python for Basic Data Science

(i) Online Course Design

This course introduces the fundamental concepts of Data Science and unveils how the Python programming language can be used in Data Science. Emphasis is given to data handling and solving primary tasks of data science. An experience on Data Visualization and building simple learning models in python are the highlights of the course. The Course is designed in such a way that students will be able to apply data science concepts using Python programming language and the course can be beneficial to students who are eager to know data science.

Upon completion, students will be able to perform data handling tasks and build simple learning models using Python programming language.

ii Learning objectives, assessment, and instructional strategies or activities

Learning Objectives

- To understand the basics of python programming language and get introduced to data science
- To explore python packages used for data handling and solve simple tasks of data science
- To perform data visualization and interpret various charts using python.
- To apply the concepts of data science and build simple learning models using python

Assessment:

Digital Assignments that require students to submit the python code for the task given, and mandates the submission within the deadlines. The Assignments will be shared to students on a regular basis and the same will be evaluated by the course teacher. Ten Assignments will be distributed during the duration of the course and each assignment will be graded for 10 marks. The Assessments will be consolidated for 100 marks.

10 Assignments * 10 Marks each = 100 Marks.

Instructional strategies or activities:

Presentations, Lectures with real-time examples, Demos, Discussions, Live Quizzes and Brain Storming sessions.

(iii) Course Syllabus

A. Introduction to Data Science

- The Multidisciplinary nature of Data Science
- What Deep Learning?
- Artificial Intelligence
- Data Science Life cycle
- Application of data science
- Issues in data science

B. Introduction to Python Programming Language

- Features of Python
- Setting Up a Python Environment
- Interpreters
- Program Execution,
- Statements,
- Expressions,
- Flow Controls
- Functions
- Lists
- Tuples
- Dictionaries
- Packages in Python

C. Python for Data Science

- Why Python for Data Science?
- NumPy
- Pandas
- Data Loading, Storage, and File Formats
- Data Frame and Data Access
- Data Operations
- Slicing and Dicing
- Data Cleaning and Preparation
- Data Wrangling: Join, Combine, and Reshape

D. Data Visualization with Matplotlib

- Bar Chart
- Line Chart
- Histogram
- Scatter Plot
- Pie Chart

E. Model Building using Python

- Scikit-learn
- Core and Advanced APIs
- Scikit-learn Examples
- Case Study Using Free Datasets
     Regression Models
     Classification Models

References:

- Practical Machine Learning with Python, Dipanjan Sarkar. Raghav Bali, Tushar Sharma, Apress, ISBN:978-1-4842-3206-4

- Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter 3rd EditionWes McKinney, Third edition, O'Relly Media,2022. ISBN: 978-1098104030
- Data Mining: Concepts and Techniques, Fourth Edition, Jiawei Han, Jian Pei, Hanghang Tong, Morgan Kaufmann, 2022

- Data Science From Scratch: First Principles with Python, Second Edition, Joel Grus, 2019,O'Relly Media,ISBN: 9781492041139

iv. Classroom management

The Course will be conducted fully **online** and in **Synchronous** mode. The classroom will be managed through **Zoom** Platform. The Students can be communicated offline through Gmail for the purpose of conveying information during non-class hours.