



CLUSTER UNIVERSITY SRINAGAR
SYLLABUS (FYUP UNDER NEP 2020)
Offered By Department of INFORMATION TECHNOLOGY
Semester 3rd Skill Enhancement Course (SEC)

Course Title: Basics of Software Development Using Python-III

Course Code: UGICT22S301	Max. Marks 100
Credits: 4 (Theory: 1, Practical: 3)	Theory External: 15; Min Marks: 06
Contact Hrs: 105 (Theory: 15, Practical: 90)	Theory Internal (Continuous Assessment): 10 Marks, Min Marks: 04
	Practical Experimental Basis= 45 Marks, Min. Marks: 18
	Practical Internal (Continuous Assessment): 30 Marks, Min. Marks: 12

Course Objectives: This course aims to introduce fundamental concepts of data Analysis and machine learning using Python. The course outcome shall enable learners analyse different types of data using Python NumPy and Pandas and implement basic supervised and unsupervised machine learning algorithms.

Course Outcomes: After completion of this skill course the learner should be able to:

1. Understand the fundamental concepts of machine learning and data, and perform data pre-processing, processing, and data visualization to get insights from data.
2. Create and manipulate arrays using NumPy.
3. Use pandas to create and analyse data sets.
4. Use matplotlib and seaborn libraries to create beautiful data visualization.
5. Implement Supervised and unsupervised machine learning algorithms and evaluate performance metrics using scikit-learn.

UNIT-I

Understanding data, structured and unstructured data, types of data: numerical, categorical, and cardinal data. Data Pre-processing: Cleaning and preparation of data, Identify and handling missing Values. Introduction to Machine Learning: Need for Machine Learning, Machine learning Process, Types of Machine Learning: Supervised Learning, Unsupervised Learning and Reinforcement Learning, Applications of Machine Learning.

UNIT-II

Working with NumPy and Pandas library

Python NumPy: Introduction to NumPy, how to install and import NumPy library, Understanding and creating one-dimensional and n-dimensional arrays, Basic indexing and slicing, Boolean indexing, adding, removing, and sorting arrays, Basic array operations: addition, subtraction, multiplication, division, working with matrices, basic matrix operations, Data processing using arrays, File handling using NumPy.

Python Pandas: Introduction to Pandas, how to install and import Pandas library, series and data frames, creating series, accessing data from series, basic series functions. Creating data frames from scratch, reading, and writing data from csv files, basic data frame operations, data frame slicing, selecting, extracting, grouping, and sorting. Pandas built-in functions for basic operations, working with text data.

UNIT-III

Data Visualization using Matplotlib and Seaborn in Python

Introduction to Matplotlib, how to install and import Matplotlib, pyplot in Matplotlib, subplot in Matplotlib, matplotlib grids, formatting axes, legends, appearance, colours, markers, line styles, and range.

Types of plots in Matplotlib: Line plot, Histogram, Bar Chart, Scatter plot, Pie charts, Boxplot etc.

Introduction to Seaborn: how to install and import Seaborn, loading dataset, Seaborn plotting functions, styling and themes in Seaborn, Seaborn colour palette, types of plots in Seaborn: Distribution Plots, Pie Chart, Bar Chart, Scatter Plots, Pair Plots, Heat maps.

UNIT-IV

Introduction to Machine Learning in Python

Introduction to scikit-learn, how to install and import scikit-learn, understanding classes in scikit-learn, using scikit-learn to implement Machine Learning algorithms.

Regression Algorithms: Linear Regression, Logistic Regression.

Introduction to Clustering: Types, Clustering algorithms: K-means algorithm.

Classification Algorithm: Random Forest classifier, SVM, Naïve Bayes;

Model evaluation methods: accuracy score, Precision, recall, F1 -score, confusion matrix.

Recommended Books:

1. Python: The Complete Reference, Martin C Brown, McGraw Hill.
2. Machine Learning with Python, Abhishek Vijayvargia, BPB Publications.
3. Data Analysis with Python: A Modern Approach, David Taieb, Packt Publishing
4. Python for Data Analysis, Wes Mckinney, Publisher O'Reilly Media