Course Name - Python for Data Science

Aim

At the entry-level, your primary aim is to build a strong analytical foundation. Goals should focus on developing technical proficiencies in programming languages such as Python or R, understanding data manipulation and visualization, and gaining experience with machine learning algorithms.

Vision:

The data revolution has created novel challenges and unprecedented opportunities. The vision of the Department of Data Science is to create a first-class academic department that trains the next generation of students as data scientists who will solve these grand challenges and innovate through world-class research to take advantage of these opportunities.

Mission

- Educate students in a field that has ushered in a once-in-a-generation revolution, comparable to the industrial revolution and the original computing revolution.

- Provide an environment for leading edge research that has a strong and rapid impact on the economy and that reestablishes New Jersey as a world leader in technological advancement.

- Create a source of scholarship on the many technical, ethical, and privacy issues that ubiquitous data creation is constantly confronting us with.

- Establish a center of technology knowledge and a "go to organization" to service data creators, providers, managers, curators, and users of the State and the Nation.

Objectives: -

1. At the entry-level, your primary aim is to build a strong analytical foundation.

2. Goals should focus on developing technical proficiencies in programming languages such as Python.

3. Understanding data manipulation and visualization.
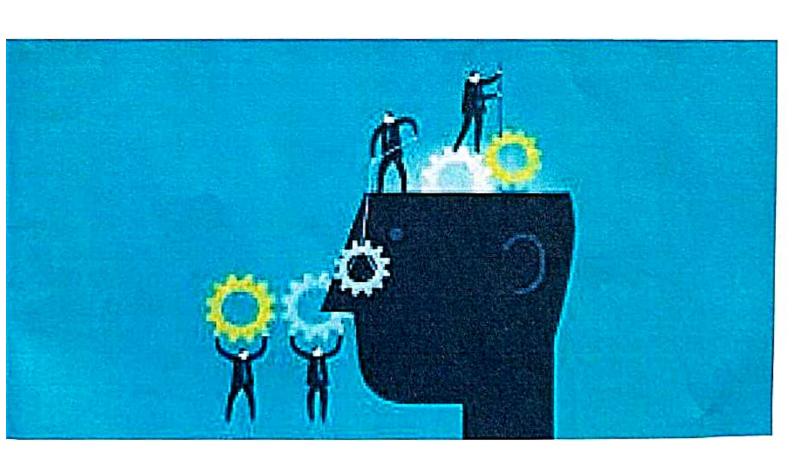
4. Gaining experience with machine learning algorithms.

Syllabus

Unit	Content	No. of Lectures
I	**Introduction to Data Science**: Evolution of Data Science – Data Science life cycle, Data Science Roles – Stages in a Data Science Project – Applications of Data Science in various fields – Data Security Issues. **Data Collection Strategies** – Introduction to Collection of Data, Primary and Secondary Data, Methods of Collecting Primary Data, Methods of Secondary Data, Statistical Errors, Rounding off Data. **Data Pre-Processing Overview** – Data Cleaning – Data Integration and Transformation – Data Reduction – Data Discretization, Outlier analysis, Testing and Training **Model design and Development**: Model Evaluation using Visualization – Residual Plot – Distribution Plot –, Measures for In-sample Evaluation – Prediction and Decision Making. Generalization Error – Out-of-Sample Evaluation Metrics – Cross Validation – Over fitting – Under Fitting and Model Selection – Prediction by using Ridge Regression – Testing Multiple Parameters by using Grid Search.	15
II	**Introduction to NumPy**: Arrays and Vectorized Computation-The NumPy ndarray- Creating ndarrays- Data Types for ndarrays-Arithmetic with NumPy Arrays- Basic Indexing and Slicing - Boolean Indexing-Transposing Arrays and Swapping Axes. **Universal Functions**: Fast Element-Wise Array Functions-Mathematical and Statistical Methods-Sorting Unique and Other Set Logic. **Introduction to pandas**: Series, Data Frame, Creating Data Frame from an Excel, .csv file, python Dictionary, python List and Tuples, Operations on Data Frames: Dropping Entries Indexing, Selection, and Filtering- Function Application and Mapping- Sorting and Ranking. Summarizing and Computing Descriptive Statistics-Unique Values, Value Counts, and Membership. Reading and Writing Data in Text Format. **Data visualization using Matplotlib and Seaborn**: Basic plot-Line, Bar, scatter, subplot, Statistics plots-Box, Histogram, errorbar and pie, 3D plots-scatter, surface, triangular, Heatmap, Density Plots, CatPlot (Categorical Plot), Joint Distribution Plot.	15

Course Outcome: -

After the completion of the course, the students will-

1. Know basic data types in Python.
2. Know operators, and how to clean and merge datasets.
3. Know pandas library, the main methods for Data Frames.
4. Know how to import data in Python.
5. Know how to work in Spyder.

।।न हि ज्ञानेन सद्दशं । पवित्रमिह विद्यते ।।

Shriram Shikshan Sanstha's
Shriram Institute of Information Technology, Paniv
Ta. – Malshiras, Dist. – Solapur, [MH] 413113
NAAC Accredited with 'B' Grade

Admission Form of Certification Course

Course Name -
☐ English for Personality Development
☐ Knowledge Engineering
☐ Yoga Education
☐ Advanced MS-Office
☐ Python for Data Science

Name of Students - _____

Gender - ☐ Male   ☐ Female

Academic Program - ☐ B.C.A.   ☐ B.Sc.(ECS)   ☐ M.Sc.(Comp. Sci.)

Part - ☐ I   ☐ II   ☐ III

Academic year - _____

Mobile No - _____

Category - SC / OBC / NT / Open / ST

Address - _____
_____
_____

Student Signature

।। न हि ज्ञानेन सद्दशं । पवित्रमिह विद्यते ।।

Shriram Shikshan Sanstha's
Shriram Institute of Information Technology, Paniv
Ta. – Malshiras, Dist. – Solapur, [MH] 413113
NAAC Accredited with 'B' Grade

Faculty Members

Institute Facilites

1) Qualified staff.

2) Computer Lab with Internet Facilities.

3) 1CT enabled classrooms.

4) Library with reading room

5) Campus Interview.

6) Purified Drinking Water.

7) Green Campus.

8) Girls And Boys hostel.

9) Sports and cultural activities.

10) Mess, Canteen facilities.

Our Institute Programs

1) **Bachelor of Science (Entire Computer Science)(B.Sc.(ECS))**

2) **Bachelor of Computer Application (B.C.A.)**

3) **Master of Science(Computer Science) (M.Sc.(Comp.Sci.))**

NAAC Co-Ordinator   Vice - Principal

Asst. Prof. Nale V. D.   Asst. Prof. Dawkare R.R.

Pune   135 km   Indapur   Tembhurni   Solapur →

28 km

27 km

Akluj

Shriram Shikshan Sanstha Paniv   8km

7.5 km

8km

Phaltan   Malshiras   Khudus   Pandharpur