Information Booklet cum Syllabus

Of

Data Science and ML using Python

March 2025

National Institute of Electronics and Information Technology
An Autonomous Scientific Society under
Ministry of Electronics and Information Technology, Government of India

NIELIT Gorakhpur
M.M.M.U.T. Campus, Deoria Road
Gorakhpur (U.P.) -273010

CONTENTS

Sl. No.	Title	Page No.
1.	About the course	2
2.	NIELIT	2
3.	Objective of Course	2
4.	Job Role of Course	3
5.	Eligibility	3
6.	Duration of Course	3
7.	Registration	3
8.	Certification	3-5
9.	Faculty for the course	5
10.	Course Details	6-8
11.	Reference Books / Study Material	8-9

1. About Course

This Data Science and Machine Learning using Python course provides a practical introduction to data manipulation, visualization, and machine learning techniques using NumPy, Pandas, Matplotlib, Seaborn, and Scikit-learn. It covers essential topics such as data preprocessing, exploratory data analysis (EDA), regression, classification, clustering, and model evaluation. Through real-world projects like house price prediction, customer segmentation, learners will gain hands-on experience in applying machine learning to solve real problems. This course is ideal for beginners and intermediate learners looking to build a strong foundation in Data Science and ML with Python.

2. NIELIT

National Institute of Electronics and Information Technology, NIELIT, (Erstwhile DOEACC Society) is an autonomous scientific society of the Ministry of Electronics & Information Technology, Government of India. The Society is registered under the Societies Registration Act, 1860. NIELIT was set up to carry out Human Resource Development and related activities in the area of Information, Electronics & Communications Technology (IECT). NIELIT is engaged both in Formal & Non-Formal Education in the areas of IECT besides development of industry oriented quality education and training programmes in the state-of-the-art areas. NIELIT has endeavored to establish standards to be the country's premier institution for Examination and Certification in the field of IECT. It is also one of the National Examination Body, which accredits institutes/organizations for conducting courses in IT and Electronics in the non- formal sector.

3. Objective of Course

- To introduce Polytechnic Diploma students to the fundamentals of Data Science and Machine Learning using Python.
- To teach data handling and preprocessing techniques using NumPy and Pandas for efficient data manipulation.
- To develop skills in data visualization using Matplotlib and Seaborn for insightful analysis.
- To provide hands-on experience with machine learning algorithms such as regression, classification, and clustering using Scikit-learn.
- To enable students to build and evaluate ML models for real-world applications like house price prediction and customer segmentation.
- To prepare students for industry applications and further studies in Data Science and Machine Learning.

4. Job Roles of Course

After successful completion of the qualification the candidates shall be employed in the industries for following occupations:

- Assistant Programmer

- Data Analyst
- Career in Machine Learning/AI

5. Eligibility

- 3 Year Polytechnic Deploma (Pursuing/Passout).
- Graduation (Pursuing/Passout)

6. Total duration of the Course

60 Hours (Theory: 24 Hrs, Practical/Tutorial: 36 Hrs)

7. Registration: Interested candidates can get registration done on following web site: *https://regn.nielitvte.edu.in/*

8. **Certification:** At the end of the course, online assessment will be done. After assessment, Certificate will be provided to successful candidates.\

9. Faculty for the course:

Shri Neeraj Kumar
M.Tech (Data Science) JNU, NEW DELHI
Senior Technical Assistant (STA),
NIELIT Gorakhpur
Mobile/Whatsapp: 8127019816
E-Mail: neerajkumar@nielit.gov.in

10. Course Details

7.1. Course Outline and Objective of Each Unit

S. No.	Unit Name	Duration (Theory) in Hours	Duratio n (Practic al) in Hours	Total Learni ng Hrs.	Learning Objectives
1	Introduction to Python Programming, Operators, Expression in Python,	2	2	4	After completion of this unit of module, Learner will be able to • Understand Basic of Python Programming Compare it's with other programming language understand Data type. • Installation of Python. • Able to make

					Program by accepting user input. • Understand the Operators and expressions in Python Programming.
2	Python Statements (Conditional statement, looping statement, break statement, continue statement & pass statement)	3	3	6	• Able to make Program by using conditional Statement • Understand the concept of Loop, Nested Loop and able to make program • Based on loop & nested loop.
3	Sequence Data Types (List, Tuple, Set & Dictionary), String Handling in Python	3	6	9	After completing this unit, Learner will be able to understand • Know very well what Sequence Data type is and how to use it. Able to make program by using Sequence Data type, String in Python.
4	Functions (Types of functions, return statement, Types of arguments, concepts of local & global variable, recursion), Built-in Modules, Import statement, Packages, OOPs Concepts	3	5	8	After completing this unit, Learner will be able to understand • User Defined Function and Lambda Function, Understand the concept of arguments. • Understand the concept of local & global variables. • Able to make program by using recursion. • Understand the

					concept of Built in Module. • Able to use different inbuilt module • Concept OOPs.
5	Introduction to NumPy, Introduction to pandas	2	1	3	After completing this unit, Learner will be able to understand • Understand the concept of NumPy. • Able to use inbuilt methods of NumPy. • Understand the concept of Pandas • Able to handle csv, xls data
6	Data Exploration and Visualization, Exploratory Data Analysis (EDA) process: Descriptive statistics (mean, median, variance), Visualization techniques (histograms, scatterplots, boxplots)	4	6	10	After completing this unit, Learner will be able to understand • Visualization of data in the graph, also calculation of mean, median, mode, variance,
7	Data Cleaning and Preprocessing, Handling missing values, outliers, and data transformations, Use Pandas to fill in or drop missing values, Scaling data (MinMaxScaler, StandardScaler), one-hot encoding categorical variables	2	4	6	After completing this unit, Learner will be able to understand • Cleaning, preprocessing, handling missing data, outlier, noise, • which of the algorithm can be used to reduce the noise, redundence
8	Introduction to Machine Learning, Simple Linear Regression, Cross-Validation and Model Splitting	2	2	4	After completing this unit, Learner will be able to understand • Introduction of machine learning, types of ML

9	Linear vs. Polynomial regression, Ridge, Lasso, Regularization Techniques, Classification Models, Logistic Regression	2	2	4	After completing this unit, Learner will be able to understand • Understand the difference between Linear and non-linear data, types of regression models • How to optimize cost of the model
10	Decision Trees, and Random Forests, Unsupervised Learning and Clustering	2	4	6	After completing this unit, Learner will be able to understand • To understand the data using decision tree model, and random forest • How to handle unsupervised data and which types of algorithm be used
		25	35	60	

11. Reference Books/Study Material

- VanderPlas, J. (2016). Python Data Science Handbook: Essential Tools for Working with Data. O'Reilly Media.
- Géron, A. (2022). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (3rd ed.). O'Reilly Media.
- Mueller, C., & Massaron, L. (2018). Machine Learning with Python Cookbook: Practical Solutions from Preprocessing to Deep Learning. O'Reilly Media.