Information Booklet cum Syllabus

Of

Machine Learning using Python

Revision-I

May 2025

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

NIELIT Gorakhpur Deoria Road Gorakhpur (U.P.) -273010	NIELIT Extension Centre Lucknow NIELIT Lucknow Sumit Complex, A-1/9, Vibhuti Khand, Gomti Nagar, Lucknow – 226010

CONTENTS

1. About Course

Machine Learning is broad and fast-growing sub-field of Artificial Intelligence. This course introduces students to the basic concepts and techniques of Machine Learning. And this course also contains Basic of Python Programming to contain control structure, conditional statement, function Sequence Data type and numpy that make students to more skilled.

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

Machine Learning is broad and fast growing sub-field of Artificial Intelligence. This course introduces students to the basic concepts and techniques of Machine Learning. The objective of this course is to develop the skills required for Machine Learning Technologies with use of Python to analyze data and solving ML problems like Regression and Classification using machine learning algorithms.

After completing the module, the learner will be able to:

- Understand the basic concepts of Python language.
- Understand the basics of Machine Learning & their types.
- Understand various learning models, methods and applications under supervised and unsupervised learning.
- Understand data preprocessing for Machine Learning.
- Solve real world problems through machine learning implementation leading to predictions.

4. Job Roles of Course

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

- Machine Learning Developer
- Machine Learning Quality/Test Engineer
- Machine Learning Product Manager

5. Eligibility

BE / B.Tech. (Any Branch), 3-Year Diploma (Computer Science / Electronics/ IT), NIELIT O/A Level, 12^{th} with knowledge of programming skills.

6. Total duration of the Course

90 Hours (Theory: 40Hrs, Practical/Tutorial: 50Hrs)

7. Course Details

7.1 Course Outline and Objective of Each Unit

S. No.	Unit Name	Duration (Theory) in Hours	Duration (Practical) in Hours	Total Learning Hrs.	Learning Objectives
1	Introduction of Python	12	20	32	After completion of this module, the candidate will be able to : • How to install python software and IDEs. • How to use various data types like List, Tuple, and Dictionary etc. • How to use various loops and conditional statements. • Understand concept of Object Oriented • Connect database like MySQL using python
2	Data analysis and Exploration	08	12	20	After completion of this module students will be able to • Data Manipulation using Numpy and Pandas • Data Visualization using Matplotlib. • Preprocessing of machine learning. • Clean data for model training.
3	Machine learning & its Application.	16	16	32	After completing this unit, Learner will be able to understand • After attending this module the participants will be able to • Implementation of various machine learning algorithms • Evaluation of various machine learning algorithm

4	Mini Project/prediction	04	02	06	After completion of the project students will be • Able to apply machine learning algorithm on given data. • Evaluation and visualization of model performance. • Make predictions using machine learning algorithms.

7.2 Detailed Syllabus

Unit Name	Contents	Hrs.
Introduction of Python	• Python Installation with various IDE's • Python data Types • Control Structure • Functions • File handling • Introduction of OOP's • Connecting database like MySQL and accessing relational data	32
Data analysis and Exploration	• Data Analysis & visualization – using • numpy, • pandas • matplotlib, • seaborn • Data Cleaning	20
Machine learning & Its Application.	• Introduction to machine learning. • Supervised machine learning • Unsupervised machine learning • Study of various machine learning algorithms including Classification, Regression, • Using sklearn library to implement ML Algorithm • KNN, • K Means, • Logistic Regression, • Support Vector Machines (SVM), • Decision Tree, • Naïve Bayes, • Ensemble Methods, Random Forest etc • Gradient Boosting: An optimization technique	32
Mini Project/Prediction	• Mini Project/Prediction	06

8 Reference Books/Study Material

1. Python Programming- A modular Approach (with Graphics, database, Mobile and Web Applications by Sheetal Taneja and Naveen Kumar, Pearson.
2. Beginning Programming with Python Dummies by John Paul Meuller.
3. Machine Learning an algorithmic Perspective by Stephen Marshland
4. Introduction to Machine Learning with python by Andreas C Muller, Sarah Guido.

9 Practical Assignments

Assignment 1. Create a numpy array with following columns: hindi, english, science, math and commerce with data type int32.

i. Insert at least 10 rows in the above array.
ii. Display size and shape of the array.
iii. Print sum of each column.
iv. Print maximum element from each column.
v. Print sum of 1,4,5 row.

Assignment 2.

1. Create two array of size (3, 3) and print their sum and multiplication.
2. Create an array of size 10 and calculate square root and standard deviation.
3. Print size and dimension of above arrays.

Assignment 3.

1. Write a Python program to create and display a series of data using Pandas module.
2. Create a pandas series of 10 elements and specify their index as 101 to 110.
3. Print bottom 5 elements of the series created in question 2.
4. Insert 3 new elements in above series on index 111, 112 and 113.
5. Delete the elements at index- 103, 104,107,111 in above list.

Assignment 4.

Write a Pandas program to create and display a DataFrame from a specified dictionary data which has the index labels. Sample Python dictionary data and list labels:

1. exam_data = {'name': ['Ankita', 'Dia', 'Kapil', 'Jayesh', 'Esha', 'Mayank', Ravi, 'Lata', 'Kamal', 'Jatin'],
2. 'score': [12.5, 9, 16.5, 15, 9, 20, 14.5, 17.5, 8, 19],
3. 'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
4. 'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}

Assignment 5.

Create a data frame using dictionary.

1. Dictionary (‘id’:[P101,P102,P103,P104,P105], ‘Price’:[256, 340, 540, 260, 470])
2. Print the price of product id – p102.
3. Print values of Price column.
4. Rename the column id to Product_Id and Price to Base_Price.

Assignment 6.

Create a new data frame with three columns – Product_Name, Cost, Sales.

1. Add 10 values in data frame.
2. Add a new column named quantity with 10 values.
3. Add a new column named: Profit and total_profit and fill values.
4. Insert a new column named location after Product_Name column with 10 cities.

(New Delhi, Lucknow, Kolkata, Lucknow, New Delhi, Bengaluru, Chennai, Chennai, Kolkata, Bengaluru)

Assignment 7.

Solve sample Machine Learning Regression problem.

Assignment 8.

Solve sample Machine Learning classification problem.

10 Sample Questions

Q1. What is Machine learning?

a) The autonomous acquisition of knowledge through the use of computer programs.
b) The autonomous acquisition of knowledge through the use of manual programs.
c) The selective acquisition of knowledge through the use of computer programs.
d) The selective acquisition of knowledge through the use of manual programs.

Q2 Machine learning is a subset of which of the following.

a) AI
b) Deep Learning
c) Data Learning
d) None of the above

Q3 Among the following option identify the one which is not a type of learning

e) Semi Unsupervised Learning
f) Supervised Learning
g) Unsupervised learning
h) None of the above

Q3 Identify the type of learning in which labeled training data is used.

a) Reinforcement learning
b) Supervised Learning
c) Unsupervised Learning
d) None of the above

Q4 What is the term known as on which the machine learning algorithms build a model based on sample data?

a) Data Training
b) Training Data
c) Transfer Data
d) None of the above

Q5 What is unsupervised learning?

a) features of group explicitly stated
b) number of groups may be known
c) neither feature & nor number of groups is known
d) none of the mentioned

Q6 The goal of _____is to restructure the input data into new features or a group of objects with similar patterns.

a) Reinforcement learning
b) Supervised Learning
c) Unsupervised Learning
d) None of the above

Q 7 Unlabelled data means there is_____.

a) Only input data but no output data.
b) Both input and output data.
c) Only output data but no input data.
d) None of the above

Q8 Which is a method of grouping the objects.

a) Reinforcement
b) Association
c) Clustering
d) None of the Above

Q9 Which of the following is not a machine learning algorithm?

a) a) SVM
b) b) SVG
c) c) Random Forest
d) d) None of the Above

Q10. The output of training process in machine learning is.

a) Machine Learning Model
b) Machine Learning Algorithm
c) Null
d) Accuracy