

Machine Learning Roadmap with Python

Three phases from programming basics to guided models and evaluation

FA-0791 · Approximately 8 days (3 + 2 + 3)

Data & Analytics · Software Development

Introduction

Progress through Python foundations, data analysis and introductory machine learning in three sequential phases. Use prepared examples to understand modelling, predictions and evaluation. The final three-day phase is a broad guided survey: selected algorithms receive hands-on practice and others are demonstrations, not independent mastery of every method.

Learning outcomes

- Translate simple problems into algorithms and Python code.
- Use functions, modules, collections, exceptions and a notebook/editor environment.
- Acquire, clean, manipulate and visualise data with NumPy, pandas, Matplotlib and Seaborn.
- Compare supervised, unsupervised, recommendation, bandit and neural-network examples.
- Prepare data without leaking held-out information and evaluate a guided project.

Prerequisites

- Ability to use a computer
- Ability to use the internet
- Manage files and understand basic file systems in the computer
- Ability to communicate in the English language

A suitable computer, browser, internet and access to the prepared compatible Python/library environment.
No prior Python knowledge is required for Phase I; later phases build on earlier practice.

Training outline

1. Phase I — Python development foundations (3 days)

Programming concepts

- What Is Algorithm
- What Is Programming
- What Is Program
- Machine Language
- Programming Language
- What Is Translator

Python introduction

- Brief History of Python
- Python Versions
- Installing Python
- IDEs
- Environment Variables
- Python Documentation
- Hello world with python
- Modes of Programming

Google Colab notebooks

What Is Colab Notebook

- Why Colab Notebook Is Important
- Accessing hosted Colab and selecting a runtime
- Main Components of Colab Notebook
- Modes
- Exporting Notebook Documents

Basic Python syntax

- Identifiers

- Reserved Words
- Lines and Indentation
- Comments

Variables and collections

- Numbers
- Python Lists
- Python Tuples
- Python Dictionaries
- Python Sets
- Copying
- Python Strings
- String formatting
- Regular Expressions

Python operators

- Arithmetic Operators
- Comparison (Relational) Operators
- Assignment Operators
- Logical Operators
- Membership Operators
- Operators Precedence

Control flow

- Decision Making
- The if Statement
- The if else Statement
- For Loop
- While Loop
- Break And Continue

Functions

- Defining Your Own Functions

- Parameters
- Function Documentation
- Passing Collections to a Function
- Variable Number of Arguments
- Scope
- Map
- Filter
- Lambda

Modules and packages

- What Are Modules
- Importing Modules
- Aliasing
- Importing Set Of Element From A Module
- Namespace
- What Are Packages
- dir function
- help function

Exceptions

- What Is Exception
- Exception Types
- Try Except Component
- Handling General Exception
- Handling Specific Exception
- Raise

2. Phase II — Data workflows and analytics (2 days)

- External-library use in notebook, desktop, server and remote contexts.
- NumPy and pandas for acquisition, pipelining, ingestion, cleaning, conformance and manipulation.
- Files, encoding, extraction and conversion.
- Matplotlib and Seaborn for visual and preliminary analysis.
- Introduction to prediction problems in preparation for Phase III.

3. Phase III — ML process and preprocessing (shared 3-day phase)

- Recursion
- Memoization
- Selected advanced library concepts

Fundamentals of Machine Learning

- What is Machine Learning?
- Process of Machine Learning
- Life Cycle of Machine Learning
- Application working in Machine Learning
- Types Of Machine Learning
- Setting up the development environment lab

Data Preprocessing

- Importing Libraries
- Importing Dataset
- Taking care of Missing Data
- Encoding Data: Categorical Data
- Splitting the dataset into the Training set and Test set
- Feature Scaling

Split training/test data before fitting imputation, encoding or scaling; learn transformations from training data and apply them consistently to held-out data.

4. Phase III — Supervised learning examples

- What is Supervised Learning?
- Types of Supervised Learning

Regression

- Simple Linear Regression
- Multiple Regression
- Polynomial Regression

Decision Tree

- Implementation of Decision Tree
 - Random Forest Regression
 - Implementation of Random Forest Regression
- Classification
 - Implementation of Logistic Regression
 - Implementation of Naive Bayes
 - Implementation of Support Vector Machine (SVM)

Introduce XGBoost as gradient boosting for supervised learning, not as a neural-network architecture.

5. Phase III — Clustering, associations and recommendations

- K-means Clustering
- K-means Selecting the Number of Clusters
- Implementation of k-means clustering
- Implementation of Hierarchical Clustering

Apriori association-rule mining is separate from clustering.

- User based
- Item based
- Implementation

6. Phase III — Bandits, neural networks and NLP

Introductory bandit examples

- Implementation of Upper Confidence Bound (UCB)
- Implementation of Thompson Sampling

Neural-network demonstrations

- Implementation of Artificial Neural Networks (ANN)
- Implementation of Convolutional Neural Network (CNN)

Dimensionality reduction

- Implementation of Principal Component Analysis

- Implementation of Linear Discriminant Analysis
- Implementation of Kernel PCA

Linear discriminant analysis uses class labels for supervised reduction, unlike PCA.

NLP and project

- Implementation of Natural Language Processing in Python

Hands on Project

- Evaluation

Discuss potential business uses and the relationship among AI, ML and deep learning; evaluate examples rather than guaranteeing accurate future predictions.