

PY-ASL



Applied Supervised & Unsupervised Learning with Python

DURATION	LEVEL	TECHNOLOGY	DELIVERY METHOD	TRAINING CREDITS
5 Days	Advanced	Python	Instructor Led	NA

INTRODUCTION

Applied Supervised Learning with Python provides a rich understanding of how you can apply machine learning techniques to your data science projects using Python. You'll explore Jupyter notebooks, a technology widely used in academic and commercial circles with support for running inline code. With the help of fun examples, you'll gain experience working on the Python machine learning toolkit—from performing basic data cleaning and processing to working with a range of regression and classification algorithms. Once you've grasped the basics, you'll learn how to build and train your own models using advanced techniques such as decision trees, ensemble modelling, validation, and error metrics. You'll also learn data visualization techniques using powerful Python libraries such as Matplotlib and Seaborn.

Applied Unsupervised Learning with Python guides you on the best practices for using unsupervised learning techniques in tandem with Python libraries and extracting meaningful information from unstructured data. The course begins by explaining how basic clustering works to find similar data points in a set. Once you are well versed with the k-means algorithm and how it operates, you'll learn what dimensionality reduction is and where to apply it.

AUDIENCE PROFILE

This course is designed for developers, data scientists, and machine learning enthusiasts who are interested in unsupervised learning.

PREREQUISITES

Some familiarity with Python programming along with basic knowledge of mathematical concepts including exponents, square roots, means, and medians will be beneficial.

Applied Supervised Learning with Python is for you if you want to gain a solid understanding of machine learning using Python. It'll help if you have some experience in any functional or object-oriented language and a basic understanding of Python libraries and expressions, such as arrays and dictionaries.

COURSE OBJECTIVES

After completing this course, students will be able to:

- Understand the concept of supervised learning and its applications
- Implement common supervised learning algorithms using machine learning Python libraries
- Validate models using the k-fold technique
- Build your models with decision trees to get results effortlessly
- Use ensemble modeling techniques to improve the performance of your model
- Apply a variety of metrics to compare machine learning models
- Understand the basics and importance of clustering
- Build k-means, hierarchical, and DBSCAN clustering algorithms from scratch with built-in packages
- Explore dimensionality reduction and its applications
- Use scikit-learn (sklearn) to implement and analyse principal component analysis (PCA) on the Iris dataset
- Employ Keras to build autoencoder models for the CIFAR-10 dataset
- Apply the Apriori algorithm with machine learning extensions (Mlxtend) to study transaction data

COURSE CONTENT

Lesson 1: Python Machine Learning Toolkit

- Supervised Machine Learning
- Jupyter Notebooks
- Pandas
- Data Quality Considerations

Lesson 2: Exploratory Data Analysis and Visualization

- Summary Statistics and Central Values
- Missing Values
- Distribution of Values
- Relationships within the Data

Lesson 3: Regression Analysis

- Regression and Classification Problems
- Linear Regression
- Multiple Linear Regression
- Autoregression Models

Lesson 4: Classification

- Linear Regression as a Classifier
- Logistic Regression
- Classification Using K-Nearest Neighbors
- Classification Using Decision Trees

Lesson 5: Ensemble Modelling

- Overfitting and Underfitting
- Bagging
- Boosting

Lesson 6: Model Evaluation

- Evaluation Metrics
- Splitting the Dataset
- Performance Improvement Tactics

Lesson 7: Introduction to Clustering

- Introduction
- Unsupervised Learning versus Supervised Learning
- Clustering
- Introduction to k-means Clustering

Lesson 8: Hierarchical Clustering

- Clustering Refresher
- The Organization of Hierarchy
- Introduction to Hierarchical Clustering
- Linkage
- Agglomerative versus Divisive Clustering
- k-means versus Hierarchical Clustering

Lesson 9: Neighbourhood Approaches and DBSCAN

- Introduction to DBSCAN
- DBSCAN Versus k-means and Hierarchical Clustering

Lesson 10: Dimension Reduction and PCA

- Overview of Dimensionality Reduction Techniques
- PCA

Lesson 11: Autoencoders

- Fundamentals of Artificial Neural Networks
- Autoencoders

Lesson 12: t-Distributed Stochastic Neighbour Embedding (t-SNE)

- Stochastic Neighbour Embedding (SNE)
- Interpreting t-SNE Plots

Lesson 13: Topic Modelling

- Cleaning Text Data
- Latent Dirichlet Allocation
- Non-Negative Matrix Factorization

Lesson 14: Market Basket Analysis

- Market Basket Analysis
- Characteristics of Transaction Data

Lesson 9: Hotspot Analysis

- Kernel Density Estimation
- Hotspot Analysis

ASSOCIATED CERTIFICATIONS & EXAM

None