Introduction to Machine Learning

12 hours of instruction (Python)

Learn how to mine data and uncover patterns within it during this course. Clustering is a foundational unsupervised machine learning technique that is key to discovering latent patterns and trends. By the end of this course, attendees will learn to identify use cases where clustering is relevant, use Python to perform clustering on real-world and evaluate the results.

Objectives	Prerequisites
1. Mine data to find latent patterns and groups in different types of data 2. Evaluate the accuracy and effectiveness of clustering 3. Describe the implications of what clustering and classification methods can and cannot achieve 4. Build predictive classification models and evaluate them	Attendees must be comfortable using Python to manipulate data and must know how to create basic visualizations and import data.

Syllabus & Topics Covered

1. Introduction to clustering
   a. Commercial applications of data mining
   b. Introduction to clustering, the k-means algorithm used on voting data
2. Implementation of clustering
   a. k-means clustering on multi-dimensional data
   b. Evaluating the quality of clustering
   c. Determining the right number of clusters to use
3. Introduction to classification and supervised machine learning
   a. Commercial applications of classification models and predictive analytics
4. Building kNN models and performance metrics
   a. Define kNN technique and application
   b. Apply cross-validation to the kNN algorithm
   c. Identify and evaluate performance metrics
5. Logistic regression
   a. Introduction to logistic regression
   b. Build logistic regression model
   c. Assessing classifier performance

Software Requirements

Python & Anaconda