Machine Learning (Python)

Machine learning has become an essential component in many applications and projects that involve data. With the power of Python and the scikit-learn package, this exciting field is no longer exclusive to large companies with extensive research teams. If you use Python, even as a beginner, machine learning applications are limited only by your imagination.

During this workshop, we will take a hands-on approach to learning about machine learning algorithms. Topics include: regression, classification, outlier detection, dimensionality reduction, and clustering. During two days, we'll explore various algorithms such as linear regression, logistic regression, decision trees, neural networks, and many more.

By the end of this workshop you'll confidently select and employ machine learning algorithms using Python and scikit-learn. You'll have gained a new understanding of the inner workings of machine learning algorithms and know how to leverage them to produce valuable results and insights and build new insights from your data.

What you'll learn

- The fundamental concepts behind machine learning

- An overview of various machine learning algorithms

- How to use JupyterLab, Python, and the scikit-learn package to perform machine learning

- How to apply supervised machine learning, such as regression and classification

- How to apply unsupervised machine learning, such as dimensionality reduction, clustering, and outlier detection

Schedule

Day 1:

- Machine learning fundamentals
  - Features and labels
  - Training and testing
  - Types of machine learning

- Scikit-learn API
  - Overview of modules and classes
  - Common process

- Unsupervised machine learning
  - Outlier detection
  - Clustering

- Feature engineering
    - Principal Components Analysis
    - Feature selection
    - One-hot encoding
    - Bag-of-words representation
    - TF-IDF

Day 2:

- Classification
    - K-Nearest Neighbour
    - Decision Tree Classifier
    - Random Forest
    - Neural Network

- Regression
    - Linear Regression
    - Polynomial Regression
    - Support Vector Regression

- Model evaluation
    - Measuring performance
    - Overfitting and underfitting
    - Cross validation
    - Model selection
    - Pipelines and grid search

- Where to go from here?

Prerequisites: You're expected to have some experience with programming in Python. Roughly speaking, if you're familiar with the following Python syntax and concepts, then you'll be fine:

- assignment, arithmetic, boolean expression, tuple unpacking

- bool, int, float, list, tuple, dict, str, type casting

- in operator, indexing, slicing

- if, elif, else, for, while

- range(), len(), zip()

- def, (keyword) arguments, default values

- import, import as, from import ...

- lambda functions, list comprehension

- JupyterLab or Jupyter Notebook

Trainer: Jeroen Jannsen, PhD

This course will be held in English.

Please note:

This workshop cannot be credited as qualification measures at the Department of Mechanical Engineering. However, you are welcome to participate.