O-I Training for Advanced Users

The estimated session duration is 2 hours, final sessions might take up to 3 hours.

1 .Introduction to Python

- Introduction to Python and Jupyter notebook
- Conditionals (if statement, if expression)
- Loops (list comprehension)
- Basic Python functions
- Defining functions in Python
- Creating and importing modules and packages
- Workshop: Building list comprehensions and if expressions

2. Working with data in Python

- Reading files with Python
- Introduction to Pandas and NumPy
- Creating a Pandas dataframe from files
- Introduction to descriptive analytics
- Understanding different datatypes and type parsing
- Data transformations with Pandas
- Workshop: Manipulating data with Pandas

3. Exploratory Data Analysis

- Data visualizations with Python
- Interpreting charts, graphs and tables
- Defining objectives and questions
- Breaking problems down
- Generating hypothesis
- Workshop: Making an EDA for a dataset
- Milestone: Select problem and write proposal

4. Information gathering, data wrangling and cleaning

- Different data sources at OI
- Extracting information and web scraping
- Data preparation and data cleaning
- Information sufficiency
- Workshop: Gathering and preparing information for project

5. Power BI

- Loading data into Power BI
- Basic data transformations with M
- Basic DAX functions
- Workshop: Making different plots with Power BI
- Milestone: Completed data gathering and preparation for project

6. Supervised Machine Learning

- Supervised Machine Learning Algorithms introduction and use cases
- Classification Algorithms Overview
- Regression Algorithms Overview
- Algorithm evaluation
- Workshop: Creating a linear and logistic regression
- Milestone: Completed exploratory data analysis for project

7. Unsupervised Machine Learning

- Unsupervised Machine Learning Algorithms introduction and use cases
- Clustering overview
- Dimensionality reduction overview
- When to use an algorithm
- Workshop: application of cluster algorithms

8. Modern Machine Learning

- Modern Machine Learning Algorithms
- Hyper parameter tuning
- Deep learning and Neural Networks
- Tensorflow applications
- Model deployment for production
- Workshop: Cross validation and comparing model performance

10. Project presentation

- Milestone:
  - Document containing:
    - Introduction
    - Research question
    - Scope
    - General and specific objectives
    - State of the art (Contextualization and summary of key information)
    - Model implementation
    - Conclusions
    - Data sources
  - Jupyter notebook with the code (HTML and ipynb)
  - Datasets used
  - Final model and implementation code