



## Artificial Intelligence Training: introduction and in-depth training, AI model creation

*This training is available only in French.*



## 1. Objectif

The objective of the service offered by Orano Projects is to introduce the audience to the basic knowledge of the design principles of an Artificial Intelligence so as to allow them thereafter to have a certain autonomy to create transfer functions based on Artificial Intelligence.

## 2. Basic training knowledge

The required level of knowledge of the people trained is an engineer or technician degree, with a basic knowledge of programming concepts. No prerequisites specific to Python or Machine Learning are required.

Trainees are responsible for:

- One PC per trainee, capable of running a Python Windows executable
- WIFI connection

The training program is delivered over 3 days.

The training can be done in our premises or yours according to your preferences.

## 3. Training content

As part of the training, it is planned to address and present the following points:

### Introduction and general concepts (~day 1):

- What is AI (and what is not AI)
- The different types of AI: supervised, unsupervised, and self-supervised, reinforcement learning, static and dynamic, classification and regression, transfer learning
- Bias and Variance, Underfitting and Overfitting
- How much data does it take to train a model?
- Example of Orano applications – Process
- General principles of operation/learning of an AI
- Basics of optimization/minimization of error for learning an AI o Explainable AI: why and how
- Training and inference
- Notions of hardware: difference between CPU and GPU, strengths and weaknesses, parallel computing
- Creative AI
- Focus on supervised AI: predictive models
- Focus on unsupervised AI: clustering and anomaly detection: K-Means, Gaussian mixture, DBSCAN, inflection point, Isolation Forest, Auto-Encoder

### Deepening (~day 2):

- Databases: how to assemble it, how to create it, what should be its structure, experimental plans, etc.
- Data analysis and visualization for reconciliation and preparation
- Principal Component Analysis
- Normalization or standardization of data (Inputs and/or outputs? How and why?)
- Encode/digitize/binarize qualitative data: how and why?
- Imbalanced data: what to do, and how? Upsampling, downsampling, different methods
- Upstream Feature Engineering to improve the accuracy of machine learning
- How to validate a machine learning model: the different strategies, their strengths and weaknesses, validation (crossed or not?), training set, validation set, test set, random split or no, etc
- The different machine learning models, their strengths and weaknesses (“idiot” model, linear, polynomial, kNN, splines, decision tree, random forest, bagging, boosting, stacking, Gaussian process, neural network, the different types of neural network, etc.)
- How to optimize the training of a model and its hyperparameters? Grid Search,
- Random Search, Smart Search, Ratio Validation/Training, regularization methods
- AutoML: Automated Machine Learning
- How to analyze the error of a model, and what to do with this error once the AI is deployed in production? Mean error, mean absolute, relative, standard deviation, quartiles, parity plots, histogram, Gaussian fitting, confusion matrices, precision, and recall, f1 score, metrics
- Focus on dynamic (transient) machine learning: what are the differences with static machine learning in terms of tools, methods, complexity, real and “false” dynamics, etc.?
- Orano industrial experience feedback: what works, what does not work, the mistakes we have made, how to avoid them, etc.
- Extrapolation: can we extrapolate, should we extrapolate, how do we know if we are extrapolating, capacity of the models to extrapolate or not?
- TPs with use of Python scripts (~day 3) of machine learning in regression and classification with explain ability of models. These scripts allow Inevo, with simple copy and paste, to train and optimize AI models in regression and classification, to see what the models have “understood” from a physical point of view and how they component, and to make predictions on new cases for industrial deployment in real conditions. These scripts will be functional as is, and will be detailed line by line so that Inevo can adapt them to its convenience.