# Formation à l’Intelligence Artificielle : introduction et approfondissement, créations de modèles
IA



Orano Projets SAS 23 place Wicklow - 78180 Montigny le Bretonneux Siège social : 125 Ave de Paris 92320 Châtillon
www.orano.group

## 1. Objectif

L'objectif de la prestation proposée par Orano Projets est d'initier l'auditoire aux connaissances de base des principes de conception d'une Intelligence Artificielle de manière à leur permettre par la suite d'avoir une certaine autonomie pour créer des fonctions de transfert basées sur l'Intelligence Artificielle.

## 2. Données d'entrée d'une formation

Le niveau requis des connaissances des personnes formées est une formation d'ingénieur ou de technicien, avec une connaissance basique des notions de programmation. Aucun pré-requis spécifique au langage Python ou au machine learning n'est demandé.

Sont à la charge des personnes formées :

- Un PC par stagiaire, capable de faire fonctionner un exécutable Python Windows
- Une connexion WIFI

Le programme de formation est dispensé sur 3 journées.

Les formations peuvent se faire dans nos locaux ou les vôtres selon vos préférences.

## 3. Contenu de la formation

Dans le cadre de la formation il est prévu d'aborder et présenter les points suivants :

**Introduction et concepts généraux (~jour 1) :**

- Ce qu'est l'IA (et ce que n'est pas l'IA)
- Les différents types d'IA : supervisée, non supervisée et auto supervisée, reinforcement learning, statique et dynamique, classification et régression, transfer learning
- Biais et Variance, Underfitting et Overfitting
- Combien de données faut-il pour entraîner un modèle ?
- Exemple d'applications Orano – Procédé
- Principes généraux de fonctionnement/apprentissage d'une IA
- Bases de l'optimisation/minimisation d'erreur pour l'apprentissage d'une IA
- L'IA explicable : pourquoi et comment o Entraînement et inférence
- Notions de hardware : différence entre CPU et GPU, les forces et faiblesses, le calcul parallèle
- L'IA créative
- Zoom sur l'IA supervisée : modèles prédictifs
- Zoom sur l'IA non supervisée : clustering et détection d'anomalies : K-Means, mélange gaussien, DBSCAN, point d'inflexion, Isolation Forest, Auto-Encodeur

**Approfondissement (~jour 2) :**

- Bases de données : comment l'assembler, comment la créer, quelle doit être sa structure, plans d'expérience, etc.
- Analyse et visualisation de données en vue de leur réconciliation et préparation
- Analyse en Composantes Principales
- Normalisation ou standardisation des données (Inputs et/ou outputs ? Comment et pourquoi ?)
- Encoder/numériser/binariser les données qualitatives : comment et pourquoi ?
- Données déséquilibrées : que faire, et comment ? Upsampling, downsampling, différentes méthodes
- Feature Engineering en amont pour améliorer la précision du machine learning
- Comment valider un modèle de machine learning : les différentes stratégies, leurs forces et faiblesses, validation (croisée ou non ?), training set, validation set, test set, split aléatoire ou non, etc.
- Les différents modèles de machine learning, leur forces et faiblesses (modèle «idiot», linéaire, polynomial, kNN, splines, arbre de décision, forêt aléatoire, bagging, boosting, stacking, processus gaussien, réseau de neurones, les différents types de réseau de neurones, etc.)
- Comment optimiser l'entraînement d'un modèle et ses hyperparamètres ? Grid Search, Random Search, Smart Search, Ratio Validation/Training, méthodes de régularisation
- AutoML : le machine learning automatisé
- Comment analyser l'erreur d'un modèle, et que faire de cette erreur une fois l'IA déployée en production ? Erreur moyenne, absolue moyenne, relative, écart-type, quartiles, graphes de parité, histogramme, fitting gaussien, matrices de confusion, précision et rappel, f1 score, métriques
- Zoom sur le machine learning en dynamique (transitoire) : quelles différences avec le machine learning statique en termes d'outils, de méthodes, de complexité, vrais et « faux » dynamiques, etc.
- Retour d'expérience industriel Orano : ce qui marche, ce qui ne marche pas, les erreurs qu'on a commises, comment les éviter, etc.
- Extrapolation : peut-on extrapoler, doit-on extrapoler, comment savoir si on est en train d'extrapoler, capacités des modèles à extrapoler ou non ?
- TPs avec utilisation de scripts Python (~jour 3) de machine learning en régression et classification avec explicabilité des modèles. Ces scripts permettent à Inevo, par simple copier-coller, d'entraîner et d'optimiser les modèles d'IA en régression et en classification, de voir ce que les modèles ont "compris" d'un point de vue physique et comment ils se composent, et de faire des prédictions sur de nouveaux cas pour déploiement industriel en conditions réelles. Ces scripts seront fonctionnels tels quels, et seront détaillés ligne par ligne afin qu'Inevo puisse les adapter à sa convenance.