

# **Formation IA**

## **Rapport sur les contributions**

ID de Contribution: 1 Code de contribution: **FormIA/intro/B01/S01**Type: **Non spécifié**

## Introduction: IA, probabilité et statistique

*lundi 19 mai 2025 09:00 (3 heures)*

Ce cours aborde d'abord les notions générales liées à l'IA et met en place le vocabulaire et les définitions permettant de se repérer dans le domaine du Machine Learning. Il donne des éléments méthodologiques qui seront appliqués au cours des différentes sessions de la semaine de formation. Dans un second temps, il propose des rappels sur les concepts probabilistes et statistiques, à la base des approches par apprentissage automatique. Il montrera l'application directe de ces concepts sur les modèles les plus fondamentaux du Machine Learning, les régressions linéaires et logistiques.

**Authors:** Dr LOMET, Aurore (CEA); Dr DANIEL, Geoffrey (CEA/DES/ISAS/DM2S/SGLS/LIAD)

**Orateurs:** Dr LOMET, Aurore (CEA); Dr DANIEL, Geoffrey (CEA/DES/ISAS/DM2S/SGLS/LIAD)

**Classification de Session:** Introduction générale

ID de Contribution: **2** Code de contribution: **FormIA2025/intro/B01/TP01**  
**spécifié**

Type: **Non**

## Travaux pratiques

*lundi 19 mai 2025 13:30 (3 heures)*

Ce TP met en oeuvre des algorithmes de régression linéaire avec descente de gradient. Un exercice est aussi prévu sur un cas de classification en appliquant les concepts probabilistes vus au cours de la séance du matin.

L'objectif de ce TP est de se refamiliariser avec Python tout en mettant en application des premiers algorithmes de Machine Learning.

**Authors:** Dr LOMET, Aurore (CEA); Dr DANIEL, Geoffrey (CEA/DES/ISAS/DM2S/SGLS/LIAD)

**Orateurs:** Dr LOMET, Aurore (CEA); Dr DANIEL, Geoffrey (CEA/DES/ISAS/DM2S/SGLS/LIAD)

**Classification de Session:** Introduction générale

ID de Contribution: 3 Code de contribution: **FormIA2025/introML/B01/S01**  
**spécifié**

Type: **Non**

## **Support vector machines, arbres de décision, ensembles**

*mardi 20 mai 2025 09:00 (3 heures)*

Cette session spécifique dédiée au Machine Learning supervisé explore les algorithmes utilisés à la base pour des problèmes de classification. Les Support Vector Machines (SVMs), les arbres de décision et méthodes d'ensemble associées sont des classes d'algorithmes intéressantes par leur explicabilité et interprétabilité. Ce cours décrit leur fonctionnement et leur utilisation.

**Authors:** Dr LOMET, Aurore (CEA); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Orateurs:** Dr LOMET, Aurore (CEA); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Classification de Session:** Introduction au Machine Learning

ID de Contribution: 4 Code de contribution: **FormIA2025/introML/B01/TP01**  
**spécifié**

Type: **Non**

## Travaux pratiques

*mardi 20 mai 2025 13:30 (3 heures)*

Ce TP met en application les méthodes présentées la matinée : SVM, arbres de décisions et XGBoost. Il permet de se familiariser avec des bibliothèques Python très utilisées en Machine Learning, en particulier scikit-learn.

**Authors:** Dr LOMET, Aurore (CEA); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Orateurs:** Dr LOMET, Aurore (CEA); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Classification de Session:** Introduction au Machine Learning

ID de Contribution: 5 Code de contribution: **FormIA2025/introML/B02/S01**  
**spécifié**

Type: **Non**

## Réduction de la dimensionnalité

*mercredi 21 mai 2025 09:00 (1h 30m)*

Cette session aborde un dernier algorithme supervisé, les K plus proches voisins (KNN), et met en lumière la problématique liée à la malédiction de la dimensionnalité. Le cours introduit ensuite les premières notions liées au Machine Learning non supervisé, et en particulier la réduction de la dimensionnalité. Elle montre des méthodes permettant de mettre en évidence des relations entre les données, comme l'Analyse en Composantes Principales, ainsi que des algorithmes de représentation de données dans des espaces à faible dimension à travers la méthode t-SNE.

**Authors:** Dr LOMET, Aurore (CEA); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Orateurs:** Dr LOMET, Aurore (CEA); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Classification de Session:** Introduction au Machine Learning

ID de Contribution: **6** Code de contribution: **FormIA2025/introML/B02/TP01**  
**spécifié**

Type: **Non**

## Travaux pratiques

*mercredi 21 mai 2025 10:30 (1h 30m)*

Ce TP propose d'appliquer les méthodes abordées dans la première partie de la matinée : KNN, PCA et t-SNE.

**Authors:** Dr LOMET, Aurore (CEA); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Orateurs:** Dr LOMET, Aurore (CEA); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Classification de Session:** Introduction au Machine Learning

ID de Contribution: 7 Code de contribution: **FormIA2025/introML/B02/S02**  
**spécifié**

Type: **Non**

## Clustering

*mercredi 21 mai 2025 13:30 (1h 30m)*

Ce cours continue d'explorer les méthodes non supervisées en abordant les problématiques de clustering. Ce type d'approche a pour objectif de mettre en évidence des groupes de données présentant des similarités, et il rassemble un grand nombre de méthodes, dont les principales seront présentées dans le cadre de cette session.

**Authors:** Dr LOMET, Aurore (CEA); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Orateurs:** Dr LOMET, Aurore (CEA); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Classification de Session:** Introduction au Machine Learning



ID de Contribution: **8** Code de contribution: **FormIA2025/introML/B02/TP02**  
**spécifié**

Type: **Non**

## Travaux pratiques

*mercredi 21 mai 2025 15:00 (1h 30m)*

Cette séance de TP met en application des méthodes de clustering : les K-Means, DBSCAN et les Gaussian Mixtures.

**Authors:** Dr LOMET, Aurore (CEA); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Orateurs:** Dr LOMET, Aurore (CEA); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Classification de Session:** Introduction au Machine Learning

ID de Contribution: **9** Code de contribution: **FormIA2025/introIA/B01/S01**  
**spécifié**

Type: **Non**

## Introduction aux réseaux de neurones

*jeudi 22 mai 2025 09:00 (1h 30m)*

Cette seconde partie de la semaine se focalise sur les approches de Deep Learning. Ce cours introduit les modèles à la base de l'apprentissage profond : les réseaux de neurones. Il montre les principes mathématiques à la base de ces modèles et leur apprentissage. Enfin, il aborde leur utilisation dans le cadre de l'analyse de données de type image à travers les réseaux de neurones convolutifs.

**Authors:** Dr DANIEL, Geoffrey (CEA/DES/ISAS/DM2S/SGLS/LIAD); KUCHER, Inna (CEA-ISAS)

**Orateurs:** Dr DANIEL, Geoffrey (CEA/DES/ISAS/DM2S/SGLS/LIAD); KUCHER, Inna (CEA-ISAS)

**Classification de Session:** Introduction au Deep Learning

ID de Contribution: **10** Code de contribution: **FormIA2025/introlA/B01/TP02**  
**spécifié**

Type: **Non**

## TP réseaux de neurones (suite)

*jeudi 22 mai 2025 15:00 (1h 30m)*

Cette séance de TP propose de continuer l'implémentation des réseaux de neurones vues dans la matinée, en particulier en abordant les réseaux de neurones convolutifs. Pour celles et ceux qui seront suffisamment avancées et à l'aise avec les premiers exercices, des TP supplémentaires sont proposées pour aborder des concepts plus avancés, comme les auto-encodeurs, la quantification d'incertitudes ou encore la robustesse des modèles.

**Authors:** Dr DANIEL, Geoffrey (CEA/DES/ISAS/DM2S/SGLS/LIAD); KUCHER, Inna (CEA-ISAS)

**Orateurs:** Dr DANIEL, Geoffrey (CEA/DES/ISAS/DM2S/SGLS/LIAD); KUCHER, Inna (CEA-ISAS)

**Classification de Session:** Introduction au Deep Learning

ID de Contribution: **11** Code de contribution: **FormIA2025/introIA/B02/S01**  
**spécifié**

Type: **Non**

## Techniques avancées pour les réseaux de neurones

*vendredi 23 mai 2025 09:00 (2 heures)*

Cette session s'intéresse à des modèles avancés fondés sur les réseaux de neurones. Il aborde en particulier les Graph Neural Networks (GNNs), vus comme une généralisation des réseaux de neurones convolutifs pour des données qui peuvent se représenter sous forme de réseau. Puis, il introduit les modèles Transformers, à la base des algorithmes de traitement du langage naturel, en plein essor ces dernières années avec les grands modèles de langage comme GPT, Bard, Deep Seek, Le Chat de Mistral AI...

**Authors:** KUCHER, Inna (CEA-ISAS); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Orateurs:** KUCHER, Inna (CEA-ISAS); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Classification de Session:** Introduction au Deep Learning

ID de Contribution: **12** Code de contribution: **FormIA2025/introIA/B02/D01**  
**spécifié**

Type: **Non**

## Démon (modèles de langage)

*vendredi 23 mai 2025 11:00 (30 minutes)*

Du fait du coût calculatoire des entraînements de ces modèles avancés, des démonstrations sont proposées à la place des séances de TP. Cette séance de démon s'intéressera plus particulièrement à HuggingFace, qui est une plateforme de modèles de fondation utilisée pour aborder des problématiques variées en IA.

**Authors:** KUCHER, Inna (CEA-ISAS); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Orateurs:** KUCHER, Inna (CEA-ISAS); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Classification de Session:** Introduction au Deep Learning

ID de Contribution: **13** Code de contribution: **FormIA2025/introIA/B02/S02**  
**spécifié**

Type: **Non**

## Introduction à l'IA générative

*vendredi 23 mai 2025 13:00 (2h 30m)*

Cette dernière séance se concentrera sur l'IA générative, qui permet entre autres de générer du contenu réaliste, comme des fausses images. Le cours abordera la problématique de la génération de données à haute dimension et les méthodes, fondées sur les réseaux de neurones, qui ont été développées pour l'aborder, comme les Variational AutoEncoders, les Generative Adversarial Networks, les Normalizing Flows et les modèles de diffusion.

**Authors:** KUCHER, Inna (CEA-ISAS); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Orateurs:** KUCHER, Inna (CEA-ISAS); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Classification de Session:** Introduction au Deep Learning

ID de Contribution: **14** Code de contribution: **FormIA2025/introIA/B02/D02**  
**spécifié**

Type: **Non**

## Démonstration (génération d'images)

*vendredi 23 mai 2025 15:30 (30 minutes)*

Cette séance de démonstration montrera la mise en œuvre de modèles de génération d'images, et le couplage avec des modèles de traitement du langage pour effectuer de la génération conditionnelle.

**Authors:** KUCHER, Inna (CEA-ISAS); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Orateurs:** KUCHER, Inna (CEA-ISAS); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Classification de Session:** Introduction au Deep Learning

ID de Contribution: 15 Code de contribution: **FormIA/Q&A/S01**Type: **Non spécifié**

## Questions et wrap-up

*vendredi 23 mai 2025 16:00 (30 minutes)*

Cette séance de clôture permet d'effectuer ensemble le bilan de la semaine et de revenir sur les points méthodologiques essentiels. Elle sera l'occasion aussi d'échanger sur les dernières questions des participants et participantes.

**Authors:** Dr LOMET, Aurore (CEA); Dr DANIEL, Geoffrey (CEA/DES/ISAS/DM2S/SGLS/LIAD); KUCHER, Inna (CEA-ISAS); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Orateurs:** Dr LOMET, Aurore (CEA); Dr DANIEL, Geoffrey (CEA/DES/ISAS/DM2S/SGLS/LIAD); KUCHER, Inna (CEA-ISAS); FINOTELLO, Riccardo (CEA Paris-Saclay)

**Classification de Session:** Q&A



ID de Contribution: **16** Code de contribution: **FormIA2025/introlA/B01/TP01**  
**spécifié**

Type: **Non**

## TP Réseaux de neurones

*jeudi 22 mai 2025 10:30 (1h 30m)*

Cette session de TP propose de développer des premiers réseaux de neurones appliqués à la reconnaissance d'images sur une base de données très classiquement utilisée en Deep Learning : MNIST pour la classification d'images de chiffres. Elle permet de se familiariser avec une des bibliothèques principales du Deep Learning : TensorFlow, ainsi que la surcouche de haut niveau Keras.

**Orateurs:** Dr DANIEL, Geoffrey (CEA/DES/ISAS/DM2S/SGLS/LIAD); KUCHER, Inna (CEA-ISAS)

**Classification de Session:** Introduction au Deep Learning

ID de Contribution: 17 Code de contribution: **FormIA2025/introlA/B01/S02**  
**spécifié**

Type: **Non**

## Quantification d'incertitudes en Deep Learning

*jeudi 22 mai 2025 13:30 (1h 30m)*

Cette session introduit les notions de quantification d'incertitudes en Deep Learning, fondamentales pour garantir la fiabilité des modèles. Ce sujet reste un problème ouvert dans la littérature sur le Deep Learning, et le cours propose de définir les concepts associés et d'explorer certaines méthodes de l'état de l'art.

**Orateurs:** Dr DANIEL, Geoffrey (CEA/DES/ISAS/DM2S/SGLS/LIAD); KUCHER, Inna (CEA-ISAS)

**Classification de Session:** Introduction au Deep Learning