Industrialisation et Déploiement de Solutions IA

INFORMATIONS GÉNÉRALES

Type de formation : Formation continue

Domaine : IA, Big Data et Bases de données

Filière : IA

Rubrique : Certification ATLAS : CISIA (Actions co.)

Éligible au CPF : Non

Action collective : Oui
Code ACO : CISIA

Code de formation : CISIA-IAR

€ Tarifs

Prix public : 3000 €

Tarif & financement :
Financement possilble via les Actions Collectives ATLAS ou le Plan de Formation.

Lieux & Horaires

Campus : Ensemble des sites

Durée : 22 heures

Rythme : 9h30-12h30 et 14h-17h

Délai d'accès :
Jusqu'a 8 jours avant le début de la formation, sous condition d'un dossier d'insciption complet

Distanciel possible : Oui

Prochaines sessions

Cliquez sur la date choisie pour vous inscrire :

- 23 / 06 / 2025
  : Ensemble des sites
  : Distanciel possible
  : 22 heures
  : 3 jours
- 23 / 06 / 2025
  : Ensemble des sites
  : Distanciel possible
  : 22 heures
  : 4 jours

PRÉSENTATION

Objectifs & compétences

Maîtriser l'architecture et les services IA proposés par les hyperscalers (AWS, Azure, GCP).
Mettre en œuvre les bonnes pratiques de déploiement et d'intégration des modèles IA.
Utiliser les frameworks MLOps pour l'automatisation du cycle de vie des modèles.
Optimiser l'entraînement des modèles IA et assurer leur mise en production efficace.

Public visé

Data scientists, ingénieurs IA, architectes cloud et développeurs ayant des bases en machine learning et infrastructure cloud.
Développeurs IA, data scientists, et ingénieurs IT ayant des bases en entraînement de modèles d'IA et déploiement.

Pré-requis

Connaissance des bases en intelligence artificielle et déploiement de modèles IA.
Familiarité avec le cloud computing est un atout.
Connaissances de base en machine learning, familiarité avec des frameworks comme TensorFlow, Keras, ou PyTorch.

PROGRAMME

Modules

- Maîtriser les Architectures et Services IA des Hyperscalers – 4H00 – 2
- Optimiser et Monitorer l'Entraînement des Modèles IA – 2H00 – 1,2
- Maîtriser les Pratiques Avancées de Développement et de Déploiement de Solutions d'IA – 16H00 – 1,2, 3

Outilis utilisés : Docker, Kubernetes, Azure, AWS, Grafana.

Mots clés : Déploiement de modèles IA, Docker, Kubernetes, MLOps (MLflow, Kubeflow), API IA (FastAPI, Flask), CI/CD pour IA, Conteneurisation, Déploiement sur cloud (AWS, Azure, GCP), Microservices IA, Gestion des modèles en production.

CESIA-IAR01 - Maîtriser les Architectures et Services IA des Hyperscalers
Durée : 4 heures

Objectifs pédagogiques globaux :

1. **Comprendre les grands principes d'architectures et l'état de l'art sur le packaging de modèles IA** (C6, C7)
2. **Acquérir des notions sur les services IA proposés par les hyperscalers (AWS, Azure, Google Cloud)** (C6, C7, C9)

Programme détaillé
1. Développer des Compétences en Architectures pour l'IA (2 heures)

- **Objectif :** Comprendre les différents concepts d'architecture pour intégrer et déployer

Informations non contractuelles - PDF généré le 12/08/2025

des modèles IA.
- **Contenu :**
  - Exploration des principaux éléments d'architectures IA (microservices, containers, pipelines de déploiement) (C6, C7).
  - Principes de packaging de modèles IA pour leur déploiement (Docker, Kubernetes, Serverless) (C6).
  - Étude des meilleures pratiques pour assurer des performances et des coûts optimisés (C7, C9).
- **Activité :** Atelier d'analyse d'architectures existantes avec cas d'usage concret d'une solution IA.

2. Utiliser les Services des Hyperscalers pour l'IA (2 heures)

- **Objectif :** Acquérir une compréhension des services IA fournis par AWS, Azure, et Google Cloud pour optimiser les pipelines d'IA.
- **Contenu :**
  - Panorama des services cloud spécialisés pour l'IA : Amazon SageMaker, Azure Machine Learning, Google AI Platform (C6, C7).
  - Intégration des solutions d'IA en cloud dans les pipelines de déploiement (C6, C7).
  - Outils d'optimisation, gestion des coûts et monitoring des modèles en production (C9).
- **Activité :** Démonstration et mise en pratique des services hyperscalers pour le packaging et la gestion d'un modèle IA.

CESIA-IAR02 - Optimiser et Monitorer l'Entraînement des Modèles IA
Durée : 2 heures

Objectifs pédagogiques globaux :

3. **Utiliser des outils de monitoring d'entraînement de type Tensorboard** (C5, C9)
4. **Utiliser des bibliothèques d'optimisation de modèles de type Optuna** (C4, C8)
5. **Acquérir des notions d'architecture d'un système d'information intégrant l'IA** (C6, C7)

Programme détaillé
1. Monitorer l'Entraînement avec TensorBoard (45 min)

- **Objectif :** Comprendre et mettre en œuvre des techniques de monitoring pour suivre l'évolution de l'entraînement des modèles IA.
- **Contenu :**
  - Introduction à TensorBoard pour visualiser les métriques d'entraînement (C5).
  - Suivi des performances des modèles pendant et après l'entraînement (C8).
  - Analyse des courbes de perte et d'accuracy pour affiner les modèles (C5, C9).
- **Activité :** Mise en place d'un tableau de bord TensorBoard pour un modèle en cours d'entraînement.

2. Optimiser les Modèles avec Optuna (45 min)

- **Objectif :** Utiliser des outils d'optimisation hyperparamétrique pour améliorer les performances des modèles IA.
- **Contenu :**
  - Introduction à Optuna pour la recherche d'hyperparamètres (C4).
  - Implémentation de l'optimisation dans un pipeline d'entraînement (C4, C9).
  - Évaluation des résultats d'optimisation et ajustement des modèles (C8).
- **Activité :** Utilisation d'Optuna pour optimiser les hyperparamètres d'un modèle en temps réel.

3. Notions d'Architecture pour l'Intégration de l'IA (30 min)

- **Objectif :** Comprendre les composants essentiels d'une architecture système intégrant des solutions IA.
- **Contenu :**
  - Éléments d'architecture d'un système d'information avec IA (C6).
  - Intégration des briques IA dans un système existant (API, monitoring, déploiement) (C6, C7).
  - Optimisation des architectures pour des critères de performance et de coûts (C7, C8).
- **Activité :** Analyse et conception d'une architecture simple pour intégrer un modèle IA.

CESIA-IAR03 - Maîtriser les Pratiques Avancées de Développement et de Déploiement de Solutions d'IA
Durée : 16 heures

Objectifs pédagogiques globaux :

6. **Appliquer les bonnes pratiques MLOps (versioning, développement continu, CI/CD)** (C5, C6, C9)

Informations non contractuelles - PDF généré le 12/08/2025

7. **Exposer les modèles IA via des API, services web, et événements** (C6, C8)
8. **Développer des interfaces utilisateur basiques pour des MVP d'IA (Tkinter, Flask, Django)** (C4, C6)
9. **Acquérir une bonne connaissance des éléments d'architecture impliqués dans les projets IA** (C7, C8)
10. **Challenger les propositions techniques sous contraintes de performance, coût, et écologie** (C7, C8, C9)

Programme détaillé

1. Mettre en œuvre les pratiques MLOps pour les projets IA (4h)

- **Objectif :** Comprendre et appliquer les pratiques de MLOps pour améliorer le cycle de vie des modèles IA.
- **Contenu :**
  - Gestion des versions de modèles IA (versioning) (C5, C9).
  - Automatisation du développement avec CI/CD pour le déploiement de modèles (C6, C9).
  - Monitoring des modèles et amélioration continue (C9).
- **Activité :** Mise en place d'un pipeline CI/CD pour un modèle IA, suivi du versioning du modèle.

2. Exposer et intégrer l'IA via des API et Web Services (4h)

- **Objectif :** Apprendre à déployer et rendre accessible un modèle IA via des API et services web.
- **Contenu :**
  - Création et déploiement d'une API pour l'inférence IA (C6).
  - Utilisation de frameworks (Flask, FastAPI) pour exposer les modèles (C6).
  - Gestion des événements et intégration dans des systèmes existants (C8).
- **Activité :** Développer et exposer un modèle IA via une API avec Flask.

3. Développer des interfaces utilisateurs basiques pour IA (3h)

- **Objectif :** Développer une interface utilisateur simple pour visualiser et exploiter les résultats IA.
- **Contenu :**
  - Introduction aux frameworks d'IHM (Tkinter, Flask, Django) pour le développement d'un MVP (C4, C6).
  - Création d'une interface utilisateur pour interagir avec les modèles IA (C6).
  - Intégration des modèles IA dans une application web (C6).
- **Activité :** Création d'une IHM basique avec Flask pour visualiser les résultats d'un modèle IA.

4. Concevoir l'architecture technique des projets IA (3h)

- **Objectif :** Connaître les principaux éléments d'architecture pour les projets IA et les adapter selon les besoins.
- **Contenu :**
  - Principes de base de l'architecture logicielle pour les systèmes IA (C7).
  - Infrastructure cloud et sur site pour héberger et déployer les solutions IA (C6, C7).
  - Gestion des ressources pour l'IA (CPU, GPU, etc.) (C7, C8).
- **Activité :** Conception d'une architecture technique pour déployer un modèle IA dans un environnement cloud.

5. Challenger les Propositions Techniques sous Contrainte (2h)

- **Objectif :** Être capable d'évaluer et de challenger des propositions techniques selon les critères de performance, coût, et écologie.
- **Contenu :**
  - Analyse des performances des solutions IA (temps de réponse, efficacité) (C8).
  - Optimisation des solutions pour un coût et une consommation énergétique réduits (C7, C9).
  - Prise en compte des enjeux écologiques dans le développement des systèmes IA (C9).
- **Activité :** Étude de cas pour évaluer les performances d'un modèle IA selon plusieurs critères (coût, efficacité, écologie).

Informations non contractuelles - PDF généré le 12/08/2025

MODALITÉS

Modalités

L'ensemble du parcours est accessible en présentiel, à distance ou mode hybride.

Présentation théorique en présentiel.
Atelier pratique avec exercices en ligne et en présentiel.
Études de Cas : Analyse d'applications réelles des techniques de génération et d'augmentation.
Discussion Interactive : Échange sur les meilleures pratiques, les défis rencontrés et les retours d'expérience.

CERTIFICATIONS

A l'issus du parcours (10 modules), les candidats pourrons passer le jury de certifcation ATLAS :
Concevoir et implémenter une solution d'IA

Informations non contractuelles - PDF généré le 12/08/2025