Concevoir et implémenter une solution d'Intelligence Artificielle pour les Data scientists

15 j (105 heures) Ref : IADS

Public

Professionnels dont les missions, en termes de traitement de données, sont plus avancées (notamment l'exploration de données avancée, la modélisation statistique et l'utilisation de techniques d'apprentissage automatique et d'intelligence artificielle pour résoudre des problèmes complexes) et qui disposent de compétences déjà solides en mathématiques et statistiques.

Pré-requis

- Exercer un emploi comportant des activités orientées vers la création d'infrastructures, de logiciels, vers le traitement et l'analyse de données ou encore vers la maintenance des systèmes ;
- Des connaissances en mathématiques et statistiques, en particulier dans l'exploitation de données (statistiques descriptives, probabilités, statistiques inférentielles, analyse exploratoire des données, modélisation statistique) ;
- Justifier d'une expérience en programmation (1er niveau de maîtrise des langages de programmation, connaissances algorithmes et structures de données, architecture logicielle, systèmes d'exploitation).

Moyens pédagogiques

Formation réalisée en Présentiel, Distanciel ou Dual-Learning selon la formule retenue.
Mise à disposition de nos environnements de visio sur nos formations en Distanciel Remise d'une documentation pédagogique numérique pendant la formation La formation est constituée d'apports théoriques, d'exercices pratiques et de réflexions Dans le cas d'une formation sur site Entreprise, le client s'engage à avoir toutes les ressources pédagogiques nécessaires (salle, équipements, accès internet, TV ou Paperboard...) au bon déroulement de l'action de formation conformément aux prérequis indiqués dans le programme de formation

Modalités de suivi et d'évaluation

Auto-positionnement des stagiaires avant la formation
Émargement des stagiaires et formateur par 1/2 journée
Exercices de mise en pratique ou quiz de connaissances tout au long de la formation permettant de mesurer la progression des stagiaires
Auto-évaluation des acquis de la formation par les stagiaires
Questionnaire de satisfaction à chaud et à froid à l'issue de la formation

Formation finançable par votre OPCO
Spécificité OPCO ATLAS : cette formation est 100% financée par ATLAS dans le cadre du campusAtlas pour la branche Bureau d'Etude. Sous réserve de validation de votre dossier par ATLAS.

Objectifs

Documenter les jeux de données (datasheet, documentation technique)

Documenter le flux de traitement des données (donnée source jusqu'à l'exploitation, chaine d'approvisionnement des données)

Documenter le cycle de vie de la donnée Prendre en compte les techniques de génération de données (données synthétiques, confidentialité différentielle, etc.)

Maitriser les techniques d'augmentation de données

Comprendre les impacts de la solution par rapport au destinataire direct et indirect

Connaître, identifier et intégrer les différents risques éthiques et sociétaux associés à l'utilisation de l'IA en fonction du cas d'usage

Connaître la réglementation en lien avec la confidentialité et utilisation des données et de l'IA pour les usages sensibles

Choisir et documenter le modèle de stockage adapté en fonction du cas d'usage et des données sources Maitriser les techniques de versionnage et de mise en production des modèles

Maitriser les techniques de Mlops et frameworks associés (MLFlow, …), conteneurisation (docker) et les articuler avec les techniques de Devops

Développer des tableaux de bord pour le suivi des performances et les métriques associées Intégrer les retours utilisateurs et les limites d'utilisation (retour du contrôle vers un humain)

Connaître les menaces qui pèsent sur l'élaboration d'une solution d'IA et comprendre l'état de l'art des mécanismes d'atténuation (adversarial example) et évaluer les risques résiduels

Comprendre les modèles d'IA et leurs champs d'application

Maîtriser les outils (Tensorflow, Keras, ...) pour faire de L'IA

Compréhension, prise de recul par rapport des cas d'usage Intégrer les contraintes opérationnelles (accès aux données, SI, etc…)

Sensibilisation Écoconception (ex-gestion de flux de données vertueuse, code optimisé, …)

Connaissance des différents modèles d'IA et de leurs spécificités (ex. contraintes liées aux différentes formes d'apprentissage

Optimiser les méthodes d'apprentissage au regard du jeu de données

Elargir les connaissances sur les bibliothèques du marché.

Bonnes pratiques Mlops, versionning, dév. Continu, CI/CD…

Savoir « exposer » l'IA (API, web service, events, etc) Développement d'IHM basiques pour première retitution MVP (Tkinter, Flask, Django, …)

Avoir une bonne connaissance des principaux éléments d'architectures impliquées pour l'IA

Être capable de challenger les propositions techniques sous les contraintes performance, économique, écologique, …)

Programme détaillé

DOCUMENTATION

Durée : 8 heures

Conception de jeu de donnée

TECHNIQUE DE TRAITEMENT DE DONNEE

Durée : 12 heures

Technique de traitement de donnée

Evaluation de connaissance

ADAPTATION DE LA SOLUTION AUX ENJEUX SOCIETAUX ET BESOINS CLIENTS

Durée : 14 heures

Identifier les risques éthiques et sociétaux

Identifier et corriger les dérives du modèle

Analyser et réévaluer de manière périodique les indicateurs de performance

Intégrer les retours utilisateurs et les limites d'utilisation

MESURE ET SUIVI DE PERFORMANCE

Durée : 16 heures

Analyser et réévaluer de manière périodique les indicateurs de performance

Comprendre les objectifs et les domaines d'applications de l'IA

L'automatisation des corrections des modèles

Outils de monitoring d'entrainement

Evaluation de connaissance

MENACES

Durée : 8 heures

Introduction aux menaces

Attaques adversariales

Empoisonnement des données

Stockage de la donnée

Fuites d'informations

Évaluer les risques résiduels après l'application des mécanismes de défenses

Evaluation de connaissance

CONNAISSANCE GENERALES LIES AUX MODELES IA

Durée : 14 heures

Introduction aux différents modèles d'IA

Développer des modèles d'IA en python

Développer des réseaux de neurones

Evaluation de connaissance

MODELISATION IA

Durée : 8 heures

Prise de recul par rapport à des cas d'usage

Sensibilisation à l'ecoconception et au code optimisé

Apprendre les bonnes pratiques pour gérer un projet d'IA

Maîtriser les bases et les bibliothèques de Python pour l'IA

Comprendre les objectifs et les domaines d'application de l'IA

Intégration d'agents dans la conception

Evaluation de connaissance

LES METHODES D'APPRENTISSAGE

Durée : 8 heures

Maîtriser les environnements de développement

L'optimisation du code

Optimiser les méthodes d'apprentissage en fonction du jeu de donnée

Evaluation de connaissance

L'INDUSTRIALISATION ET ARCHITECTURE

Durée : 16 heures

Introduction aux bonnes pratiques MLOps

Exposer un modèle IA

Maîtriser les techniques de versionnage et de mise en production des modèles

Optimisation du cycle de vide des composents

Intégrer les retours utilisateurs et les limites d'utilisations

Évaluation de la qualité et de la pertinence des donnés Intégration et déploiement continue

Evaluation de connaissance

CERTIFICATION CONCEVOIR ET IMPLEMENTER UNE SOLUTION D'INTELLIGENCE ARTIFICIELLE